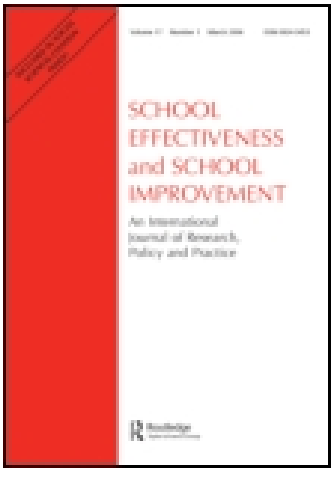


School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nses20

A longitudinal study of principals' activities and student performance

Henry May [a], Jason Huff [b] & Ellen Goldring [c]
[a] University of Pennsylvania, Philadelphia, PA, USA
[b] University of Tennessee, Knoxville, TN, USA
[c] Vanderbilt University, Nashville, TN, USA
Published online: 22 Jun 2012.

To cite this article: Henry May, Jason Huff & Ellen Goldring (2012) A longitudinal study of principals' activities and student performance, School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 23:4, 417-439, DOI: 10.1080/09243453.2012.678866

To link to this article: http://dx.doi.org/10.1080/09243453.2012.678866

School Effectiveness and School Improvement
Vol. 23, No. 4, December 2012, 417–439

A longitudinal study of principals' activities and student performance

Henry May[a]*, Jason Huff[b] and Ellen Goldring[c]

[a]*University of Pennsylvania, Philadelphia, PA, USA;* [b]*University of Tennessee, Knoxville, TN, USA;* [c]*Vanderbilt University, Nashville, TN, USA*

Although a substantial amount of research on school leadership has focused on what principals may do to improve teaching and learning, little of this research has explored how principals' time spent on leadership activities may relate to and possibly affect student performance. This article presents results from a 3-year longitudinal study of principal activities and student performance. A 3-level HLM growth model (with test scores nested within students, and students nested within schools) was employed to determine the degree to which principals' activities were associated with student performance at baseline, and changes in student performance over time. Results suggest that principals' activities are remarkably variable over time, that specific leadership activities are more prevalent in some school contexts, and that specific changes in leadership activities over time (e.g., increasing time on instructional leadership) do not predict changes in student performance in a consistent manner across schools.

Keywords: principals; achievement; longitudinal

Introduction

With the movement toward increased accountability for student performance, there has been a simultaneous call for principals to take more responsibility for improving teaching and learning in their schools (Hallinger, 2005; Leithwood & Riehl, 2003; Louis, Leithwood, Wahlstrom, & Anderson, 2010; Robinson, Lloyd, & Rowe, 2008; Stein & Spillane, 2003). There is widespread agreement between researchers and policymakers alike that principals' instructional leadership is key to increased student achievement; school leaders are central to focusing their schools on improved teaching and learning (for a recent discussion see Louis, Leithwood, et al., 2010). And yet, existing research reveals a complex relationship between principals and student achievement – principals' influences on student learning outcomes are often indirect, mediated through factors such as the quality of teacher pedagogy and a school's organization (Hallinger & Heck, 1996b; Heck & Hallinger, 2009; Supovitz, Sirinides, & May, 2010). Most recently, Louis, Leithwood, et al. (2010) found that leaders' impact on student achievement is mediated through their influence on teachers' motivation and working conditions. Despite the substantial amount of research on school leadership that has verified the complex mediated relationships between principal leadership and student achievement, little of this research has

*Corresponding author. Email: hmay@gse.upenn.edu

ISSN 0924-3453 print/ISSN 1744-5124 online

http://dx.doi.org/10.1080/09243453.2012.678866
http://www.tandfonline.com

explored just what kinds of principal leadership activities relate to and possibly affect student performance.

The purpose of this article is to examine leadership practice over time to address the question, what are the relationships between principals' activities and student achievement? The article presents findings from a 3-year, longitudinal study of all school principals in one school district in the southeastern United States. During this time, principals were asked to complete a series of daily logs recording their activities, and data from these activity logs were linked to individual-level student achievement data for these 3 years. Our analyses of the logs and the student achievement data allowed us to examine whether (1) specific principal activities are more or less prevalent in high- or low-performing schools and (2) whether changes in the amount of time a principal spends on each activity are associated with changes in student performance.

We first discuss recent research that examines the relationship between leadership and student achievement, and we highlight areas of research that remain for the field. We then present our analyses and findings and discuss how they help to address outstanding research issues.

Linking principal leadership with student achievement

While researchers often agree that strong leadership is important for successful schools, explicating the links between leadership and student achievement has proven to be challenging. Research in the 1970s and 1980s reported that effective schools often had strong leaders who were integral to their successful organization and focused on teaching and learning. Yet, multiple studies during this same timeframe failed to show significant relationships between leadership and achievement (see, e.g., Glasman, 1984; O'Day, 1983; Van de Grift, 1990). Critics pointed out that the literature often ignored the more complex contexts of schools and how leaders addressed these conditions (Hallinger, Bickman, & Davis, 1996; Hallinger, Leithwood, & Murphy, 1993; Hallinger & Murphy, 1986); still others questioned the empirical validity of direct connections between principals and student achievement (Miskel, 1982; Rowan, Dwyer, & Bossert, 1982). Various "direct-effects models" that used primarily bivariate analyses such as correlations, *t* tests, or chi-square did not provide consistent evidence of leadership effects on student outcomes (see, e.g., Braughton & Riley, 1991; Cantu, 1994; Cheng, 1994). Hallinger and Heck's widely cited syntheses of this research (1996a, 1996b, and 1998) emphasized the need for analyses that would capture the indirect relationships between school leaders and student performance.

More complex analytical models including mediating variables have shown greater promise in capturing leaders' influence on student outcomes through indirect paths mapping the effects of leadership on those people and conditions more closely connected to students' engagement and achievement, such as school culture, teachers' instructional practices, or faculty commitment to the school (Hallinger & Heck, 1998; Leithwood, 1994). Evidence for leaders' indirect effects on student achievement has come from studies in a number of countries. Multiple papers have, for instance, presented results connecting leadership to the health of the school community or its organization (e.g., Silins, Mulford, & Zarin's 2002 examination of schools in Australia and Krüger, Witziers, & Sleegers' 2007 examination of Dutch schools). Leithwood and Jantzi's (2006) analysis of a transformational leadership

model implemented in England produced evidence of leadership's effects on teachers' work settings, motivation, and classroom practices. Day et al.'s. (2009) later analysis of English school data found evidence for leaders' influences on the level of collaborative culture in school faculty, and Louis, Dretzke, and Wahlstrom's (2010) examination of US schools suggested that instructional leadership in schools relate strongly to the levels of professional commitment that teachers report. Finally, in their examination of longitudinal data from Chicago, IL, public schools, Bryk, Sebring, Allensworth, Luppescu, and Easton (2010) offer evidence that principals play key roles in promoting the complex organizational changes that are essential to school improvement. More recent research continues to confirm the complex, indirect nature of the relationship between leadership and student achievement, with studies demonstrating that leaders influence student learning through their work to build trust with their faculties (Louis, et al., 2010; Supovitz et al., 2010). Other important factors such as promoting teacher collaboration (Goddard, Hoy, Woolfolk Hoy, 2010; Supovitz et al., 2010) also influence the relationship between principal leadership and student achievement

Analytical methods such as structural equation modeling and hierarchical linear modeling have advanced to enable researchers to control for multiple factors that affect student achievement and account for the multilevel structure of schools, and these analyses have been much more successful in identifying the influences that leaders can have on student achievement (Day et al., 2009; Heck & Hallinger, 2010; Krüger et al., 2007; Louis, Leithwood, et al., 2010; Supovitz et al., 2010). In addition, recent meta-analyses have shown leadership effects on school conditions and/or student achievement ranging from negligible (Witziers, Bosker, & Krüger, 2003) to moderate or strong (Marzano, Waters, & McNulty, 2005; Robinson et al., 2008).

As these models and analyses have evolved, findings have shown that principals' engagement in a number of key activities helps to explain much of the relationship between leadership and student achievement. These activities fall into five main categories: (a) principals' involvement in framing and sustaining their schools' vision or mission and planning specific goals and strategies for school improvement (Hallinger & Heck, 1998; Leithwood, Seashore Louis, Anderson, & Wahlstrom, 2004); (b) instructional leadership functions such as monitoring instruction and providing feedback, analyzing student data, and supporting teachers' professional development (Hallinger, 2005; Knapp, Copland, Plecki, & Portin, 2006); (c) their work to enhance the organizational and social structures in their schools (through such actions as developing teacher leadership, collaboration in decision-making) (Hallinger & Heck, 1998; Leithwood et al., 2004); (d) their efforts to improve the culture or climate in their schools (in areas such as student and teacher expectations and communication between school members) (Hallinger & Heck, 1998; Knapp et al., 2006); and (e) their investment in their personnel by hiring and retaining qualified teachers (Harris, Rutledge, Ingle, & Thompson, 2010; Leithwood et al., 2004). These results point to the importance of leaders' actions to shape school conditions that improve student learning.

One shortcoming of much of the research on leadership effects is that it is based primarily on cross-sectional data; researchers have only recently analyzed longitudinal data to examine principal leadership over time (Hallinger & Heck, 1998; Heck & Hallinger, 2010; Jacobsen, 2011; Robinson et al., 2008). While some previous longitudinal studies have included leadership alongside other school-level variables to examine changes in student achievement, few if any have included

extensive measures of leaders' activities in their examinations, and these have yielded limited evidence of the impact of leadership on achievement (see, e.g., D'Agostino, 2000). While recent longitudinal examinations such as Sammons, Gu, Day, and Ko (2011) that have analyzed more robust measures of leadership have produced stronger evidence of leaders' indirect influence on achievement.

Heck and Hallinger (2010) highlight the importance of implementing longitudinal research designs:

> examining processes in organizations at any one point in time provides a very limited snapshot of conditions – as if nothing is "in motion." In those circumstances, it is not possible to represent the separate trajectories that different organizational processes are following over time or to determine the relationships among the variables at other points in time (i.e., before or after the snapshot) ... Moreover, at any point in time it is possible that extraneous variables (e.g., idiosyncratic turnover, community problems, funding changes) may exert an impact on leadership or other variables. Longitudinal data collection and analyses can facilitate the separation of extraneous factors from intended effects. (pp. 880–881)

Longitudinal examinations thus better capture the dynamic nature of leadership and whether or not *changes* in leaders' practices and activities across multiple years have an impact on student outcomes or other school conditions.

However, many longitudinal models fail to take full advantage of the power of longitudinal data to control for time-invariant confounds. Although a model may use change in an outcome as a dependent variable, the independent variables may still be strictly cross-sectional. The most powerful longitudinal designs are those in which both dependent and independent variables are measured at multiple points in time (Heck & Hallinger, 2010). This allows one to test whether changes in an independent variable (e.g., a principal's activities) are predictive of simultaneous or subsequent changes in a dependent variable (e.g., student performance). Under this approach, we can distinguish whether specific leadership activities are more or less prevalent in high- or low-performing schools, and whether increases or decreases in specific leadership activities are predictive of subsequent changes in student performance. The purpose of this study is to implement such a longitudinal design that explores both changes in principal leadership activities and their relation to changes in student achievement.

Specifically, our research questions are:

(1) Is the amount of time a principal spends on specific leadership activities related to the average achievement of students in his or her school?
(2) Are increases or decreases in the amount of time a principal spends on specific leadership activities related to the schools' value added[1] to student performance?

Methods

During a period spanning 3 school years from the spring of 2005 to the spring of 2007, principals from 39 elementary and middle schools in a southeastern urban school district participated in this study. All elementary and middle schools in the district were included in this study, while high schools were excluded due to limited data on student achievement. Of the students enrolled in the study schools, 74%

were from a minority race group, 29% were eligible for free or reduced price lunch, and 2% had limited English proficiency. Variation in school-level demographics across this sample was substantial, ranging from 24% to 100% minority enrollment, from 7% to 45% free and reduced lunch eligibility, and from 0% to 12% limited English proficiency.

Each principal in the 39 participating schools was asked to participate in up to seven, separate, week-long data collection periods during which they completed daily logs documenting their activities. The end of day (EOD) log is a web-based instrument that captured principals' engagement in leadership functions and activities for a single school day. Principals were instructed to keep a record of their daily activities on a small calendar note card given to them by the research team. Then, respondents were asked to complete the logs online at the end of the school day while memories of the day were still fresh. The first page of the log presents principals with a calendar that captures an hour-by-hour account of how principals allocate their time across different categories of leadership practice between the hours of 6 a.m. and 7 p.m. Principals were asked to complete logs for 6 consecutive days in the spring 2005, for 5 days each in the fall of 2005, winter of 2006, and spring of 2006, and again three times (15 total days) during the fall, winter, and spring of the 2006–2007 school year. Thus principals could have log data for up to 36 days. The response rates for these seven waves of data collection were 93%, 78%, 80%, 70%, 84%, 80%, and 65%, respectively. The accuracy and validity of this EOD instrument has been demonstrated in previous research using data from this same project in which log estimates were compared to (a) researchers' observations and interviews with principals regarding their daily work (Spillane & Zuberi, 2009) and (b) the results of an experience sampling instrument that prompted principals to report their activities while they were actually working (Camburn, Spillane, & Sebastian, 2006). Results from that study suggested remarkable congruence between log data, observational data, and data from the experience sampling instrument.

The activity logs required that principals document their activities in 15-minute intervals each day by classifying their primary activity at each point in time into one of nine categories as follows.

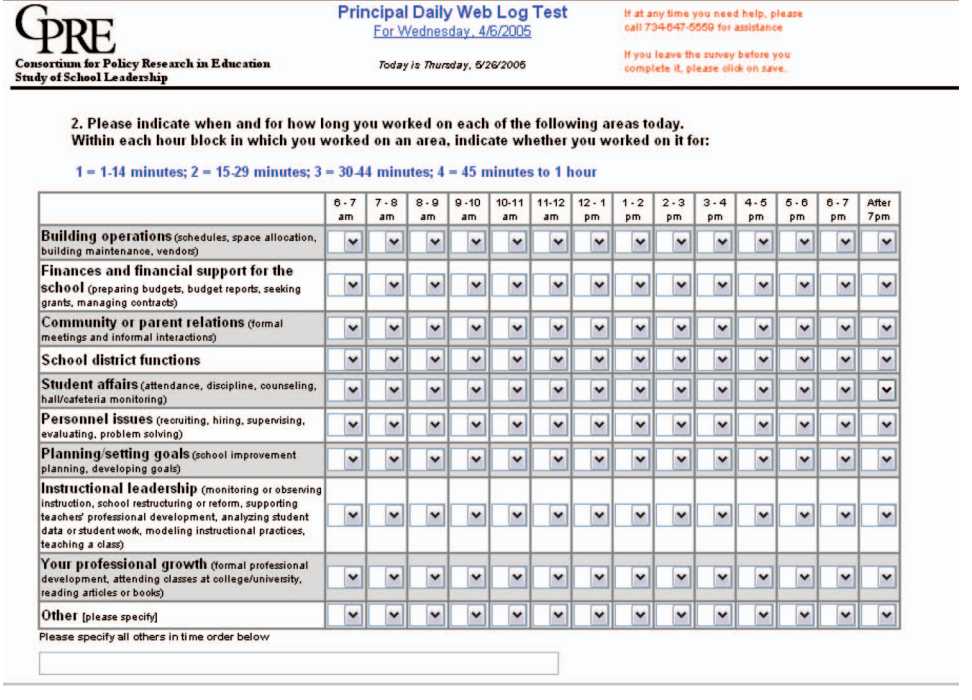
(1) *building operations*: schedules, space operations, building maintenance;
(2) *finances and financial support for the school*: budgets, budget reports, seeking grants;
(3) *community or parent relations*: formal meetings and information interactions;
(4) *school district functions*: meetings;
(5) *student affairs*: student attendance, discipline, counseling, hall/cafeteria monitoring;
(6) *personnel issues*: recruiting, hiring, supervising, evaluating, problem solving;
(7) *planning/setting goals*: school improvement planning, developing goals;
(8) *instructional leadership*: monitoring/observing instruction, supporting teachers' professional development, analyzing student data or work, modeling instructional practices;
(9) *principal professional growth*: formal professional development, attending classes at a college/university, reading books or articles.

The development of the EOD log and the identification and definitions of these nine categories followed a review of major frameworks classifying principals' activities

and responsibilities (Drake & Roe, 2003; Hallinger & Murphy, 1985; Heck & Marcoulides, 1992; Larsen & Hartry, 1987; Martin & Willower, 1981; Peterson, 1977; Pitner & Hocevar, 1987). Figure 1 contains a screenshot of the log they completed each day.

Principals' response data from the principal activity logs were aggregated to produce an activity profile for each principal, with the time spent on each type of the nine realms of responsibility coded as a proportion of the total time documented by the principal. This was done separately for each year (so that longitudinal changes in activities reflect annual changes, with 6 to 15 days contributing to the principal's average for each year), and the activity data were then linked to achievement data from individual students for the 3 years.

Achievement scores from the state assessment in English Language Arts (ELA) and Mathematics were obtained for 38,510 students enrolled in Grades 1 through 8 during the 3-year period including the 2004–2005, 2005–2006, and 2006–2007 school years. Given the natural matriculation of students into first grade and out of eighth grade, along with relatively high rates of mobility, only 16% of students in the sample had scores for all 3 years, 28% had scores for 2 years, and 56% had scores for only 1 year. The state assessment was not vertically scaled across grades, and changes in the test ruled out direct comparison of test scores from one year to the next. But even though the scores cannot be analyzed to determine absolute levels of student learning across the 3 years (i.e., the scores cannot be subtracted to estimate annual "gains"), they can be used to gauge shifts in the relative performance of students and groups of students (e.g., schools), relative to the larger population,


CPRE
Consortium for Policy Research in Education
Study of School Leadership

Principal Daily Web Log Test
For Wednesday, 4/6/2005

Today is Thursday, 5/26/2006

If at any time you need help, please call 734-647-5559 for assistance

If you leave the survey before you complete it, please click on save.

2. Please indicate when and for how long you worked on each of the following areas today.
Within each hour block in which you worked on an area, indicate whether you worked on it for:

1 = 1-14 minutes; 2 = 15-29 minutes; 3 = 30-44 minutes; 4 = 45 minutes to 1 hour

	6 - 7 am	7 - 8 am	8 - 9 am	9 -10 am	10-11 am	11-12 am	12 - 1 pm	1 - 2 pm	2 - 3 pm	3 - 4 pm	4 - 5 pm	5 - 6 pm	6 - 7 pm	After 7pm
Building operations (schedules, space allocation, building maintenance, vendors)														
Finances and financial support for the school (preparing budgets, budget reports, seeking grants, managing contracts)														
Community or parent relations (formal meetings and informal interactions)														
School district functions														
Student affairs (attendance, discipline, counseling, hall/cafeteria monitoring)														
Personnel issues (recruiting, hiring, supervising, evaluating, problem solving)														
Planning/setting goals (school improvement planning, developing goals)														
Instructional leadership (monitoring or observing instruction, school restructuring or reform, supporting teachers' professional development, analyzing student data or student work, modeling instructional practices, teaching a class)														
Your professional growth (formal professional development, attending classes at college/university, reading articles or books)														
Other [please specify]														

Please specify all others in time order below

Figure 1. Screenshot of the Principal Daily Web Log.

from one year to the next. In other words, test scores can be compared across years to determine which students and schools have improved or declined in performance relative to other students and schools in the district. To support this analysis, test scores were first standardized to a mean of zero and standard deviation of one within each year and grade to reflect student performance relative to the district average.

A three-level value-added HLM model (with test scores nested within students, and students nested within schools) was employed to determine the degree to which principals' activities were associated with student performance averaged across the 3-year period, and also changes in student performance over time. Each activity variable, denoting average hours per week spent on a specific realm of responsibility, was included in the model as both an overall school average for the 3-year period and as a time-varying covariate representing relative emphasis in each year. For example, a principal might have spent an average of 30% of her time on instructional leadership during the 3-year period (indicated by a value of .30 for the school average activity variable), but she might have spent the most time on instructional leadership, say 45% of her time, during the 2006–2007 school year (indicated by a value of + .15 for the time-varying activity variable during Year 3).[2] This approach allows us to see how (1) specific principal activities are more or less prevalent in high- or low-performing schools and (2) whether changes in the amount of time a principal spends on each activity are associated with changes in student performance.

We included control variables for students at the individual and school levels for characteristics where previous research has identified gaps in student performance, namely gender, race, poverty, and English proficiency (Porter, 2005). At the student level, we used dummy variables to capture these characteristics (male, minority, Free/Reduced Lunch status, Limited English Proficiency), while at the school level our controls used the percentage of students who were male, minority, Free/Reduced Lunch, or Limited English Proficiency.[3]

The specification of the value-added statistical model is complex. The model is similar to the Tennessee Value-Added Assessment model described by Sanders, Sexton, and Horn (1997), except that our model focuses on estimating school-level value-added estimates (as opposed to teacher-level estimates). Our model also includes student and school characteristics as explicitly controlled covariates, and our model uses a simpler autoregressive structure for the repeated measures for one subject at a time.[4] Our model is only one from a diverse class of value-added models, but it is likely the model best suited for our data and research questions for several reasons. First, the model uses a repeated measures structure to estimate annual gains for individual students in the presence of missing data. By leveraging multiple years and multiple cohorts of data, the model is able to produce value-added estimates that utilize all of the available data (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004), whereas other models might delete students who had 1 or more years of missing test scores. Second, by using 3 years of data, the value-added estimates are not based on simple gains or even residualized gains (as is the case in regression and ANCOVA) but instead are indicative of the gain specific to one year that is uncorrelated with the performance in either of the other 2 years (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006). For example, the current year value added may be interpreted as gains above and beyond those that would be expected given performance last year and the year before. Lastly, regression to the mean and bias due to unreliability in test scores are less of a problem in our model than in

traditional regression and ANCOVA (Littell et al., 2006; Sanders, 2006). Although initial criticism of this type of value-added model was grounded largely in the lack of research on such models, a flurry of publications in recent years have supported its use and, in many cases, confirmed its superiority to alternative models (Briggs & Weeks, 2011; Lockwood & McCaffrey, 2007; Lockwood, McCaffrey, Mariano, & Setodji, 2007; McCaffrey et al., 2004; Sanders, 2006; Tekwe et al., 2004).

Because the variables representing time spent on the nine leadership activities are multicollinear (i.e., spending more time on one activity results in less time on others), a separate model is estimated for each activity. The general mathematical form of the model is:

$$Y_{tij} = \beta_o + \beta_1\left(MeanActivity_j\right) + \beta_1\left(Activity_{tj} - MeanActivity_j\right) + \sum \beta_k X_{k(i,j)} + \gamma_{tj}\left(Year_t * School_j\right) + \varepsilon_{tij}$$

where,

Y_{tij} = the test score of student i from school j during year t;
β_0 = the model intercept;
β_1 = the regression coefficient for the average time spent on a specific leadership activity;
β_2 = the regression coefficient for time-varying variable for time spent on a specific leadership activity;
β_k = is a set of regression coefficients for student and school-level control variables;
γ_{tj} = is the value-added random effect for school j during year t;
ε_{tij} = is the error term for the test score of student i from school j during year t, which is assumed to have an autoregressive structure Σ, where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & \\ \rho & \sigma_2^2 & \\ \rho^2 & \rho & \sigma_3^2 \end{bmatrix},$$

with σ_t^2representing the residual variance of test scores in year t.

The key benefit of this modeling strategy over traditional repeated measures analyses is that every student in the sample having at least one test score can be included in the analysis. The implication is that for students who have fewer than three test scores, it is assumed that their unobserved scores can be accurately predicted using the test scores that were actually observed along with student and school demographic variables included in the model.[5] This modeling approach also has an advantage over HLM growth curve models in that there is no functional form imposed on students' growth trajectories (e.g., linear, quadratic, etc.). Instead, year is treated as a discrete variable and is allowed to vary randomly for each school. This allows the model to handle the lack of vertical scaling in the test scores, given that each student's relative performance, along with school value added, may go either up or down in any given year.

We next present our results with the following qualification. Because this study is observational, the relationships we observed do not allow us to draw strong conclusions about the causal relationships between principals' activities and student

performance. For example, a positive relationship between a specific activity and student achievement might be interpreted with opposite causal direction – the principal's activity might lead to higher student performance, or higher performance might allow the principal to devote more time to the activity.

To begin to address this issue of causality, our second research question focuses on simultaneous changes in principals' activities and changes in value added to student achievement. More specifically, we wish to evaluate whether shifts in a principal's activities from one year to the next are associated with increases or decreases in the schools' value added to student performance for that school year. In essence, we seek to make two distinct comparisons in our analyses and results. The first is a comparison of principals' activities across low- to high-performing schools. The second is a comparison of each principals' activities during the years in which their school produced lesser or greater value added to student achievement.

Results

Table 1 presents descriptive statistics for our measures of principals' activities. The most prevalent activity across the sample of principals was Student Affairs, which accounted for 23.3% of principals' time on average across the 3-year study period. A close second was Instructional Leadership, which accounted for 19.3% of principals' time on average. Four other activities (Building Operations, Parent Relations, Personnel Issues, and Planning & Setting Goals) each accounted for approximately 10% of principals' time on average. The remaining three activities (Finance, District Functions, and the Principal's Professional Growth) each accounted for approximately 5 to 7% of principals' time on average.

The standard deviation of each activity was substantial, ranging from 5 percentage points to almost 15 percentage points. This suggests a large amount of variation, either across schools or across years, in how principals allocate their time. The final two columns in Table 1 show how that variation is attributable to differences between schools, or differences across years within schools. The ratios of between- and within-school variation for each activity are not extreme. Building Operations has the least extreme variance partition, with a nearly even 1 to 1 split for between- and within-school variance. Even the most extreme split in the variance

Table 1. Descriptive statistics for nine principal leadership activities.

	Percentage of Time Spent on Leadership Activity			
	Mean	Standard Deviation	Percent Variance Between Schools	Percent Variance Within Schools, Over Time
Building Operations	9.1	8.0	51.1	48.9
Finance	4.8	5.0	43.1	56.9
Parent Relations	10.1	6.5	61.3	38.7
District Functions	6.6	8.3	27.2	72.8
Student Affairs	23.3	13.7	68.3	31.7
Personnel Issues	10.6	9.2	36.5	63.5
Planning & Setting Goals	9.2	6.8	36.0	64.0
Instructional Leadership	19.3	14.7	76.3	23.7
Principal's Professional Growth	6.9	9.3	37.3	62.7

partition is surprisingly even – Instructional Leadership shows only a 3 to 1 ratio of variance between schools (76.3%) to variance within schools (23.7%). This suggests that principals' activities are variable over time but may be partially predicted by school and/or principal characteristics. Furthermore, the distributions of time and variability in time allocations across the set of nine activities suggest that each activity demands a significant amount of a principal's time, that principals from different schools spend different amounts of time on different activities, and that a principal from a single school is likely to change substantially his or her allocations of time across leadership activities from one year to the next. These longitudinal findings differ with Horng, Klasik, and Loeb's (2010) observations of principal activities over a single day where they found that principals spent most of their time on administrative activities such as fulfilling compliance requirements or managing school schedules. However, as we discuss later, such 1-day findings may easily fall within the wide range of time allocations that principals choose from one week or year to the next – our results capture a wider picture of what principals do over time, and these differences may very well be the result of examining principals' activities over a larger number of days within each year.

Table 2 presents partial correlations, controlling for year, among the leadership activity measures and school-level covariates. Correlations between most pairs of variables are low and insignificant, suggesting that increased time on one activity does not generally correspond to increased or decreased time on another specific activity. However, there are a number of moderate correlations that suggest inverse relationships between pairs of activities. Two remarkable examples of this suggest that increased time on student affairs or personnel issues is associated with decreased time spent on instructional leadership. Lastly, the very high correlation between percent minority and percent free/reduced lunch suggests possible collinearity between these variables, but since they are used only as controls, and one variable remains highly significant (as presented later in the results section), both variables were included in each model.

In the next stage of our analyses, we estimated multilevel models examining time spent on leadership activities and its relationship to schools' value added to student achievement. Estimates of both fixed and random effects from the value-added models are presented in Tables 3 and 4 for ELA and Math, respectively. Each model is based on a sample of 61,859 test scores across 3 years, nested within 38,510 students, nested within 42 schools. Models were estimated using PROC MIXED in SAS version 9.22.

Results across the two subjects are very consistent. First, there are no significant differences in average performance across years (as indicated by non-significant fixed effects for Year 2 and Year 3), but there is consistently significant variance in school effects in each of the 3 years (as indicated by the highly significant random effects for Years 1, 2, and 3). This is not at all surprising, given that the test scores were standardized by year and grade to reflect relative performance over time, which has the effect of removing district-wide average annual differences, leaving only school-to-school differences in relative performance. The random effect estimates reveal that unmeasured school-level factors account for 3% to 6% of the variability in student performance in the two models. When compared to the random effect estimates from an unconditional model without any control variables (i.e., 8% to 10%), we see that approximately half of the school-level variability in student performance is explained by the student- and school-level variables included in the models.

Table 2. Partial correlations (controlling for year) among nine principal leadership activities and school-level covariates.

	A	B	C	D	E	F	G	H	I	J	K	L	M
A. Building Operations	–	0.11	−0.01	−0.16	−0.13	−0.19	−0.14	−0.20*	0.05	−0.09	−0.07	−0.10	0.00
B. Finance		–	−0.02	−0.01	0.09	0.01	−0.18	−0.31**	−0.08	−0.10	−0.06	−0.10	0.08
C. Parent Relations			–	−0.14	−0.21*	−0.05	−0.21*	−0.10	0.05	0.05	−0.01	−0.03	0.10
D. District Functions				–	−0.08	−0.10	−0.15	−0.04	−0.37**	0.05	−0.30**	−0.24*	0.10
E. Student Affairs					–	0.03	−0.20*	−0.43**	−0.17	0.17	0.10	0.14	−0.15
F. Personnel Issues						–	−0.19*	−0.45**	−0.10	−0.12	0.07	−0.04	−0.10
G. Planning & Setting Goals							–	0.06	0.00	0.01	0.22*	0.24*	0.00
H. Instructional Leadership								–	−0.09	−0.01	0.01	0.07	0.08
I. Principal's Professional Growth									–	−0.03	−0.05	−0.08	−0.03
J. Percent Male										–	0.07	0.18	0.09
K. Percent Minority											–	0.87**	−0.21*
L. Percent Free/Reduced Lunch												–	−0.17
M. Percent ELL													–

Note: *$p < .05$, **$p < .01$.

Table 3. Parameter estimates from multilevel models of student achievement in ELA by nine principal leadership activities.

	Principal Leadership Activity								
Model Parameters	Building Operations	Finance	Parent Relations	District Functions	Student Affairs	Personnel Issues	Planning & Setting Goals	Instructional Leadership	Principal's Professional Growth
Fixed Effects									
Intercept	1.02***	0.99***	1.05***	0.97***	1.11***	1.00***	1.20***	1.17***	1.15***
Year 2	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
Year 3	−0.01	0.00	0.00	0.00	−0.01	−0.01	0.01	−0.01	−0.01
Average Proportion of Time on the Leadership Activity	0.04	0.18~	0.05	0.09	0.01	0.08~	−0.14*	−0.06*	−0.08
Changes in Proportion of Time on the Leadership Activity	0.02	0.00	0.00	0.00	−0.01	0.01	−0.01	0.00	0.00
Student-Level Controls									
Male	−0.29***	−0.29***	−0.29***	−0.29***	−0.29***	−0.29***	−0.29***	−0.29***	−0.29***
Minority	−0.33***	−0.33***	−0.33***	−0.33***	−0.33***	−0.33***	−0.33***	−0.33***	−0.33***
Free/Reduced Lunch Eligibility	−0.15***	−0.15***	−0.15***	−0.15***	−0.15***	−0.15***	−0.15***	−0.15***	−0.15***
Limited English Proficiency	−0.54***	−0.54***	−0.54***	−0.54***	−0.54***	−0.54***	−0.54***	−0.54***	−0.54***
School-Level Controls									
Percent Male	−0.37	−0.38	−0.45	−0.50	−0.48	−0.43	−0.47	−0.42	−0.42
Percent Minority	−0.86***	−0.90***	−0.89***	−0.78***	−0.89***	−0.90***	−0.84***	−0.85***	−0.89***
Percent Free/Reduced Lunch	0.02	0.02	0.00	0.01	−0.01	−0.01	0.05	−0.01	0.00
Percent LEP	0.95	0.72	0.81	0.79	0.85	1.02	0.77	0.95	0.84

(*continued*)

Table 3. (*Continued*).

Model Parameters	Principal Leadership Activity								
	Building Operations	Finance	Parent Relations	District Functions	Student Affairs	Personnel Issues	Planning & Setting Goals	Instructional Leadership	Principal's Professional Growth
Random Effects									
Variance of School Effects – Year 1	0.06***	0.05***	0.06***	0.06***	0.06***	0.05***	0.06***	0.05***	0.06***
Variance of School Effects – Year 2	0.06***	0.05***	0.06***	0.06***	0.06***	0.06***	0.05***	0.05***	0.06***
Variance of School Effects – Year 3	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***
Correlation of Year 1 & 2 Sch. Effects	0.96***	0.96***	0.96***	0.96***	0.96***	0.96***	0.96***	0.96***	0.96***
Correlation of Year 2 & 3 Sch. Effects	0.91***	0.90***	0.90***	0.91***	0.91***	0.90***	0.91***	0.89***	0.91***
Correlation of Year 1 & 3 Sch. Effects	0.84***	0.83***	0.85***	0.85***	0.85***	0.83***	0.84***	0.83***	0.85***
Autocorrelation of Annual Test Scores	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***	0.75***
Residual Variance	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***

Note: $^{\sim}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Each model is based on a sample of 61,859 test scores across 3 years, nested within 38,510 students, nested within 42 schools. Each model also includes fixed effects for the eight grade levels, but these parameters are not shown in this table.

Table 4. Parameter estimates from multilevel models of student achievement in math by nine principal leadership activities.

	Principal Leadership Activity								
Model Parameters	Building Operations	Finance	Parent Relations	District Functions	Student Affairs	Personnel Issues	Planning & Setting Goals	Instructional Leadership	Principal's Professional Growth
Fixed Effects									
Intercept	1.22***	1.18***	1.25***	1.25***	1.29***	1.11***	1.41***	1.40***	1.33***
Year 2	−0.01	0.00	0.00	−0.01	−0.01	−0.01	0.00	0.00	0.00
Year 3	−0.05	−0.04	−0.04	−0.06	−0.06	−0.04	−0.03	−0.02	−0.05
Average Proportion of Time on the Leadership Activity	0.05	0.13	0.05	0.06	0.02	0.08~	−0.11*	−0.06*	−0.07
Changes in Proportion of Time on the Leadership Activity	0.03~	0.02	−0.02	0.01	−0.01	−0.01	0.00	0.01	0.00
Student-Level Controls									
Male	−0.08***	−0.08***	−0.08***	−0.08***	−0.08***	−0.08***	−0.08***	−0.08***	−0.08***
Minority	−0.44***	−0.44***	−0.44***	−0.44***	−0.44***	−0.44***	−0.44***	−0.44***	−0.44***
Free/Reduced Lunch Eligibility	−0.14***	−0.14***	−0.14***	−0.14***	−0.14***	−0.14***	−0.14***	−0.14***	−0.14***
Limited English Proficiency	−0.28***	−0.28***	−0.28***	−0.28***	−0.28***	−0.28***	−0.28***	−0.28***	−0.28***
School-Level Controls									
Percent Male	−0.70	−0.63	−0.76	−0.85~	−0.79	−0.56	−0.84~	−0.77	−0.71
Percent Minority	−0.91***	−0.95***	−0.96***	−0.86***	−0.93***	−0.96***	−0.89***	−0.97***	−0.93***
Percent Free/Reduced Lunch	−0.17	−0.14	−0.15	−0.16	−0.17	−0.14	−0.12	−0.11	−0.17
Percent LEP	0.69	0.56	0.66	0.79	0.69	0.62	0.70	0.50	0.67
Random Effects									
Variance of School Effects – Year 1	0.04***	0.04***	0.04***	0.04***	0.04***	0.04***	0.04***	0.03***	0.04***
Variance of School Effects – Year 2	0.05***	0.04***	0.05***	0.05***	0.05***	0.05***	0.04***	0.04***	0.05***
Variance of School Effects – Year 3	0.05***	0.04***	0.04***	0.05***	0.05***	0.04***	0.05***	0.04***	0.05***
Correlation of Year 1 & 2 Sch. Effects	0.93***	0.92***	0.93***	0.93***	0.93***	0.93***	0.92***	0.92***	0.93***
Correlation of Year 2 & 3 Sch. Effects	0.89***	0.88***	0.87***	0.88***	0.89***	0.87***	0.88***	0.87***	0.88***
Correlation of Year 1 & 3 Sch. Effects	0.89***	0.87***	0.88***	0.88***	0.89***	0.88***	0.88***	0.87***	0.88***
Autocorrelation of Annual Test Scores	0.77***	0.77***	0.77***	0.77***	0.77***	0.77***	0.77***	0.77***	0.77***
Residual Variance	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***	0.85***

Note: ~$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. Each model is based on a sample of 61,859 test scores across 3 years, nested within 38,510 students, nested within 42 schools. Each model also includes fixed effects for the eight grade levels, but these parameters are not shown in this table.

Across the models for ELA scores, four of the nine leadership activities are related to student performance; however, annual changes in principals' activities are not related to annual changes in school-level value added for any of the nine activities. This suggests that while some leadership activities may be more heavily emphasized by principals in lower or higher performing schools, changes in principals' leadership activities across the 3-year period are not related to changes in their schools' value added to student achievement. Of the four activities that are related to overall student achievement, two activities are positively related and two activities are negatively related to student performance. A 10-percentage-point increase in the percent of time devoted to Finance issues is associated with a .18 standard deviation increase in student achievement. A 10-percentage-point increase in the percent of time devoted to Personnel issues is associated with a .08 standard deviation increase in student achievement. A 10-percentage-point increase in the percent of time devoted to Planning and Setting Goals is associated with a .14 standard deviation decrease in student achievement. Lastly, a 10-percentage-point increase in the percent of time devoted to Instructional Leadership is associated with a .06 standard deviation decrease in student achievement. It is important to note that the causal direction of these relationships cannot be established. It is possible that the activities lead to differences in performance, but it is perhaps more likely that differences in performance lead to differences in leadership activities. For example, a decrease in student achievement in a school may cause a principal to focus more time on instructional leadership activities to help teachers address their students' learning needs.

Across the models for Math scores, three of the nine leadership activities are related to student performance, and one time-varying indicator of changes in principals' activities is related to annual changes in school-level value added. The three activities that are related to overall student achievement are a subset of the same variables found significant in the ELA models (i.e., Personnel Issues, Planning & Setting Goals, and Instructional Leadership), and the coefficients for these three variables are remarkably similar to those from the ELA models. The one time-varying covariate that is related to changes in school value added is the proportion of time spent on building operations – a 10-percentage-point increase in the percent of time a principal spent on building operations is associated with a .03 standard deviation increase in student performance. Upon visual inspection of the data for building operations, it was apparent that this relationship was likely driven by one school in which the principal spent an unusually large amount of time on building operations (i.e., over 20%) in one particular year.

Highly significant coefficients are observed for all four student-level covariates and one school-level covariate. All four student-level variables (male, minority, free lunch, and LEP) are associated with lower test scores. At the school level, a higher proportion of minority students is associated with lower test scores. In ELA, we also see a weak negative relationship between achievement and the proportion of boys at a school. These effects are consistent with the literature on achievement gaps by gender, race, poverty, and English proficiency (Porter, 2005).

Another important result from this analysis is reflected in the very high correlations between school value-added estimates over time. These correlations for adjacent years are near or above .90, suggesting that school effectiveness is very consistent from one year to the next. Correlations in the low to mid-80s between Year 1 and Year 3 school effects suggest that changes in school effectiveness happen very slowly.

The final result from these analyses reveals a relatively high correlation between individual student test scores from one year to the next, although this is not surprising. The adjacent year correlation is estimated as .75 for ELA and .77 for math, suggesting that a student's test scores from one year are highly predictive of his or her test scores for the next year.

Discussion and conclusion

We set out in this study to examine the relationships between principals' leadership activities and student achievement in their schools. While previous studies have used observations to examine principals' activities over a number of days within a year (Peterson, 1977; Pitner, 1982) or over the course of single days (Horng et al., 2010), our end of day log data allowed us to analyze principals' activities across many days and multiple school years. With these data, we have been able to examine not only principals' use of time within years but also (a) changes in their activities across school years and (b) differences in their activities in relation to their students' achievement. We discuss the significance of each of these below.

First, what did we learn in terms of how principals spend their time? Our descriptive results for year-to-year variations underscore the wide range in principals' activities over time. While we found that on average they focused primarily on activities such as student affairs (23.3%) and instructional leadership (19.3%), the standard deviations for these averages along with the ratios of percent variance between and within schools (across years) indicate that principals across the schools were likely to change their practices from year to year. Simply put, a principal who emphasized student affairs or instructional leadership in one year may not engage in that activity to the same degree in the next one. These variations over time add a new dimension to the traditional picture of principals' actions as often unpredictable or fragmented (Peterson, 1977; Pitner, 1982). Although our analyses do not identify the sources of this variance (i.e., unanticipated issues or needs that change over the course of the school year, such as testing), we see strong evidence of the dynamic set of activities, even when viewing them from year to year. On the surface, these findings contrast with Horng et al. (2010), who recently reported that over the course of a single day principals spent more time on administrative tasks (about 30%) and organization management (about 20% of the time). However, as we emphasized earlier, our study measures activities over several days at three time periods for 3 consecutive years. Definitions of activities may also explain the differing results. For example, Horng et al. (2010) included "managing student discipline" under "administration", while we included the activity under "student affairs".

Second, what did we learn about how principals' activities are related to student achievement? When looking at principals' activities in relation to student outcomes, our results indicated that while principals who spent relatively more time on finance and personnel issues tended to work in schools with higher test scores, principals who spent relatively more time on planning and setting goals and instructional leadership tended to work in schools with lower test scores.

Of those leadership activities that we found to be significantly related to student performance (i.e., the link between high performance and finance/personnel activities and the link between low performance and instructional leadership activities), we

believe that the more plausible causal relationship is that school context drives principals' activities. This would support Hallinger and Heck's (1998) discussion of the important reciprocal effects between leaders and school context. It would also underscore our previous findings regarding the importance of context to explain principal practice (Goldring, Huff, May, & Camburn, 2008) as well as Louis, Leithwood, et al.'s (2010) report that contextual factors such as school size, level, and poverty level influence what principals do and accomplish. Principals in higher performing schools may simply have more time and capacity to focus on finance and personnel issues because they do not feel the same level of urgency and pressure to set goals or drive instructional reform as principals in the lowest performing schools. On the other hand, as lower performing schools face greater sanctions under current policies, principals may feel the need to engage more often in addressing instructional needs with their teachers. It is only logical to expect that school leadership is driven not only by the principal's preexisting philosophy about schooling and desire to change a school in certain ways but also by the principal's reaction to conditions in the school, which may change from year to year. It is also possible that principals in lower performing schools are forced to be more reactive more often to the schools' current circumstances than principals in higher performing schools. This may explain why principals from lower performing schools in our study spent more time on instructional leadership activities. This would suggest that the job of a principal may often be more reactive than proactive.

Third, we did not find evidence that changes in a principal's activities were associated with changes in a schools' value added to student achievement. As researchers increasingly use longitudinal data to examine principal leadership, these results illustrate the importance of looking more closely at how individuals' practices vary *across* schools and *over time*. As discussed above, this study has provided a more systematic examination of principal practices across schools and years, and the findings demonstrate the need to capture more closely individual principals' actions as they relate to their unique school conditions and changes over time – while our previous cross-sectional analysis (Goldring et al., 2008) showed that school context related more strongly to leaders' practices than did their individual characteristics, we still do not know if these relationships hold across multiple years.

As the field continues to examine leaders' effects on student achievement, reciprocal longitudinal models may capture more accurately the relationships between school leadership and student performance. Such analyses may help to reveal the causal direction of these relationships. Furthermore, the inclusion of teacher-level variables and effects in our model may have revealed mediated relationships in which effects of leadership on achievement occur through effects on instruction (as in Supovitz et al., 2010). Unfortunately, teacher-student links were available only for a single year and only for elementary and middle schools. Using those links would have reduced our sample dramatically and, more importantly, eliminated our ability to look at longitudinal changes in practices and outcomes.

Our results may be explained by considering the direct effects perspective implemented in this article and by considering the importance of context in school leadership. In terms of direct effects of leadership on student outcomes, there still seems to be support for a direct effects perspective on leadership and student outcomes, as a meta-analysis of 70 studies found effect sizes ranging from 0.16 to 0.33 between teacher perception of school leadership and student achievement

(Waters, Marzano, & McNulty, 2003). Similarly, Robinson et al. (2008) analyzed 22 studies and concluded the average effect of instructional leadership is 0.44 on student outcomes. However, as noted above, many recent studies have found significant mediated effects of principal leadership on student outcomes, such as teacher professional community. Our findings suggest a need to continue to explore both direct and indirect effects of leadership on student outcomes in longitudinal studies. There are just not sufficient longitudinal studies to understand how mediation of leadership activities and practices might influence student outcomes.

One mediated factor that we suggest deserves more attention in longitudinal studies is the notion of relational trust between teachers and principals. Researchers have identified trust in schools as essential for the principal's role in school improvement (Kruse, Louis, & Bryk, 1994). Bryk and Schneider (2003) explain that trust is important because of the mutual dependence that exists between school members: "All participants remain dependent on others to achieve desired outcomes and feel empowered by their efforts" (p. 41). Studies demonstrate that leaders influence student learning through their work to build trust with their faculties (Louis, Leithwood, et al., 2010; Supovitz et al., 2010). As noted by Bryk and colleagues (2010), "regardless of which of the essential supports local leaders might emphasize (enhancing parent outreach, professional capacity building, improving the quality of the student learning environment or the instructional guidance system), trust facilitates the initiation of these improvement efforts" (p. 139) and "Relational trust creates a motivating force for taking up the difficult work for school reform" (p. 140).

Another important mediator we propose is academic press, or the extent to which teachers focus on academic excellence and professional and academic standards that support learning (Hoy & Hannum, 1997, p. 294). A principal may have excellent leadership skills, yet if there is no change in academic press, students may not be directed in ways that lead to higher achievement. Although academic press is not the only mediator in school improvement, we focus on it because the findings of its mediating impact on student achievement are robust (e.g., Alig-Mielcarek, 2003; Goddard et al., 2000; Louis, Leithwood, et al., 2010; Shouse, 1996), and it encompasses the aspects of a normative, collective school climate that can provide a path for principals to influence student achievement by establishing a clear sense of purpose and goals. Furthermore, Hoy, Hannum, and Tshannen-Moran (1998) suggest that academic press is an aspect of leadership that shows sustained influence on student outcomes over time. Louis, Leithwood, et al. (2010) indicate that leaders should focus attention on a small number of priorities that are squarely focused on the ultimate outcomes, in this case, student achievement.

In terms of the importance of context in understanding our results, undeniably, the results suggesting that school effects are highly consistent from year to year are sobering, perhaps disappointing, but not surprising. Schools are complex organizations, which typically serve hundreds of students and often face substantial cultural and economic difficulties. To expect that large numbers of schools can dramatically change their outcomes in a short period of 1–3 years is quite optimistic. Fullan (2005, 2006) has discussed the challenges of "turning around" schools at length, and has commented that often "what looks like apparent success in turning around schools is quite superficial and illusory" (2005, p. 174). While he cites evidence of programs bringing change within 3–4 years, these efforts most often include substantial leadership and organization of resources from the district level, not single principals

acting on their own in schools (2005). Taken together, our student achievement and principal log data underscore how varied principals' jobs are in these complex environments and how slowly change can come to these organizations.

These conditions further complicate researchers' efforts to identify specific leadership activities that have positive relationships with student achievement in every school. It may be more sensible to expect that each school has different characteristics and capacities, and that the most effective principals are able to evaluate these factors and leverage the ones that are most likely to impact instruction and learning in their schools. Leithwood et al. (2004) and others have stressed the contingent nature of leadership:

> There is a rich body of evidence about the relevance to leaders of such features of the organizational context as geographic location (urban, suburban, rural), level of schooling (elementary, secondary) and both school and district size. Each of these features has important implications for what it means to offer successful leadership. (p. 8)

These results also suggest that contextual factors not only have strong influences on student achievement but also exert strong influences on what actions principals need to take to successfully improve teaching and learning in their schools.

One might be inclined to conclude from these results that principals' activities have little influence on student performance over a 3-year period; however, our statistical models detect only systemic relationships that appear consistently across the full sample of students and schools. For example, our models are designed to detect whether a greater emphasis on instructional leadership is consistently associated with greater value added to student achievement across the entire sample of schools. If the mechanisms by which principals influence student achievement are specific to each school, then our models will fail to find relationships. In other words, if the success of a principal requires a unique approach to leadership given a school's specific context, then simple comparisons of time spent on activities will not reveal leadership effects on student performance. To capture those effects, we would need to establish not just what the principal was doing but whether the principal's activities were focused on the most important things for improving his/her school. Such relationships would suggest that leadership effects are moderated by school context, implying that the size and direction of effects of specific leadership practices depends on one or more other factors.

Other explanations may rest with neither the mediated nor the contextual explanations of the relationships between principal leadership and student outcomes, but the need to explore the combined set or configuration of leadership activities as well as the sustainability of a set of activities over time, neither of which could we study with the current design. We suggest, as does Yukl (1999), in a critique of transformational leadership theory, that a seminal question for school leadership research is how do principals influence groups or organizational processes? We suggest that a set of task-oriented behaviors, such as organizing personnel and resources to align with goals and objectives, the focus on planning and implementing agreed-upon goals, and engaging with parents are the part of the core set of task of school leadership. Much research has focused on individual tasks (see Robinson et al., 2008) but not on the order or collective efficacy of a set of activities. This is an important line of research that can expand our understanding of school leaders' effects on student outcomes.

Acknowledgements

This research was funded by the Institute of Education Sciences, U.S. Department of Education (Grant R305E040085).

Notes

1. Value added is a term that refers to estimates of a school's or teacher's unique contribution to student performance in a given year as estimated by one of several longitudinal statistical models (see McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Sanders, Sexton, & Horn, 1997).
2. This is analogous to group-mean-centering the time-varying predictors around the school average.
3. The student-level demographic indicator variables were also group-mean-centered around the school average so that changes in performance over time reflected that of the average student from each school.
4. The auto regressive structure fit the data significantly better (as determined by a likelihood ratio chi-square test) than simpler variance components and compound symmetry structures; however, model fit could not be compared to more heterogeneous covariance structures due to lack of convergence for these more complicated models.
5. This is the Missing at Random (MAR) assumption as defined by Rubin (1987).

References

Alig-Mielcarek, J.M. (2003). *A model of school success: Instructional leadership, academic press, and student achievement* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.

Braughton, R., & Riley, J. (1991). *The relationship between principals' knowledge of reading processes and elementary school reading achievement* (ERIC Document Reproduction Service No. ED341952).

Briggs, D.C., & Weeks, J.P. (2011). The persistence of school-level value-added. *Journal of Educational and Behavioral Statistics*, *36*, 616–637.

Bryk, A.S., & Schneider, B. (2003). Trust in schools: A course resource for model reform. *Educational Leadership*, *60*(6), 40–44.

Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., & Easton, J.Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Camburn, E., Spillane, J.P., & Sebastian, J. (2006, April). *Measuring principal practice: Results from two promising measurement strategies*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Cantu, M. (1994). *A study of principal instructional leadership behaviors manifested in successful and nonsuccessful urban elementary schools* (Unpublished doctoral dissertation). University of Texas, Austin, TX.

Cheng, Y.C. (1994). Principal's leadership as a critical factor for school performance: Evidence from multi-levels of primary schools. *School Effectiveness and School Improvement*, *5*, 299–317.

D'Agostino, J.V. (2000). Instructional and school effects on students' longitudinal reading and mathematics achievements. *School Effectiveness and School Improvement*, *11*, 197–235.

Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., … Kington, A. (2009). *The impact of school leadership on pupil outcomes*. Nottingham, UK: Department for Children, Schools, and Families.

Drake, T.L., & Roe, W.H. (2003). *The principalship*. New York, NY: Macmillan College.

Fullan, M. (2005). Turnaround leadership. *The Educational Forum*, *69*, 174–181.

Fullan, M. (2006). *Turnaround leadership*. San Francisco, CA: Jossey-Bass.

Glasman, N. (1984). Student achievement and the school principal. *Educational Evaluation and Policy Analysis*, *6*, 283–296.

Goddard, R.D., Hoy, W.K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Education Research Journal*, *37*, 479–507.

Goldring, E.B., Huff, J.T., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration*, *46*, 332–352.

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, *4*, 221–239.

Hallinger, P., Bickman, L., & Davis, K. (1996). School context, principal leadership and student achievement. *Elementary School Journal*, *96*, 498–518.

Hallinger, P., & Heck, R. (1996a). The principal's role in school effectiveness: A review of methodological issues, 1980-95. In K. Leithwood, J. Chapman, D. Corson, P. Hallinger, & A. Weaver-Hart (Eds.), *The international handbook of educational leadership and administration* (pp. 723–784). New York, NY: Kluwer.

Hallinger, P., & Heck, R. (1996b). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, *32*, 5–44.

Hallinger, P., & Heck, R. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement*, *9*, 157–191.

Hallinger, P., Leithwood, K., & Murphy, J. (1993). *Cognitive perspectives on educational leadership*. New York, NY: Teachers College Press.

Hallinger, P., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, *86*, 217–247.

Hallinger, P., & Murphy, J. (1986). The social context of effective schools. *American Journal of Education*, *94*, 328–355.

Harris, D.N., Rutledge, S.A., Ingle, W.K., & Thompson, C.C. (2010). Mix and match: What principals really look for when hiring teachers. *Education Finance and Policy*, *5*, 228–246.

Heck, R., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal*, *46*, 659–689.

Heck, R., & Hallinger, P. (2010). Testing a longitudinal model of distributed leadership effects on school improvement. *The Leadership Quarterly*, *21*, 867–885.

Heck, R.H., & Marcoulides, G.A. (1992). Principal assessment: Conceptual problem, methodological, or both? *Peabody Journal of Education*, *68*(1), 124–144.

Horng, E.L., Klasik, D., & Loeb, S. (2010). Principals' time use and school effectiveness. *American Journal of Education*, *116*, 491–523.

Hoy, W.K., & Hannum, J.W. (1997). Middle school climate: An empirical assessment of organizational health and student achievement. *Educational Administration Quarterly*, *33*, 290–311.

Hoy, W.K., & Hannum, J.W., & Tschannen-Moran, M. (1998). Organizational climate and student achievement: A parsimonious and longitudinal view. *Journal of School Leadership*, *8*, 336–359.

Jacobson, S. (2011). Leadership effects on student achievement and sustained school success. *International Journal of Educational Management*, *25*, 33–44.

Knapp, M.S., Copland, M.A., Plecki, M.L, & Portin, B S. (2006). *Leading, learning, and leadership support*. Retrieved from http://www.wallacefoundation.org/SiteCollection Documents/WF/Knowledge%20Center/Attachments/PDF/OverviewLeadingLearningand Leadership.pdf

Krüger, M.L., Witziers, B., & Sleegers, P. (2007). The impact of school leadership on school level factors: Validation of a causal model. *School Effectiveness and School Improvement*, *18*, 1–20.

Kruse, S., Louis, K.S., & Bryk, A.S. (1994). Building professional community in schools. In *Issues in restructuring schools, No. 6* (pp. 3–6). Madison, WI: University of Wisconsin, Center on Organization and Restructuring of Schools.

Larsen, T.J., & Hartry, A. (1987, April). *Principal/teacher perceptual discrepancy: Instructional leadership in high and low-achieving California schools*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Leithwood, K. (1994). Leadership for school restructuring. *Educational Administration Quarterly*, *30*, 498–518.

Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement*, *17*, 201–227.

Leithwood, K., & Riehl, C. (2003). *What do we already know about successful school leadership?* Paper prepared for the AERA Division A Task Force for the Development of an Agenda for Future Research on Educational Leadership.

Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004, September). *How leadership influences student learning*. Retrieved from http://cehd.umn.edu/CAREI/leadership/ExecutiveSummary.pdf

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., & Schabenberger, O. (2006), *SAS for mixed models* (2nd ed.). Cary, NC: SAS Institute Inc.

Lockwood, J.R., & McCaffrey, D.F. (2007). Controlling for individual heterogeneity in longitudinal models, with applications to student achievement. *Electronic Journal of Statistics*, *1*, 223–252.

Lockwood, J.R., McCaffrey, D.F., Mariano, L.T., & Setodji, C. (2007). Bayesian methods for scalable multivariate value-added assessment. *Journal of Educational and Behavioral Statistics*, *32*, 125–150.

Louis, K.S., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, *21*, 315–336.

Louis, K.S., Leithwood, K., Wahlstrom, K., & Anderson, S. (2010). *Investigating the links to improved student learning: Final report of research findings*. Retrieved from http://www.cehd.umn.edu/carei/Leadership/Learning-from-Leadership_Final-Research-Report_July-2010.pdf

Martin, W.J., & Willower, D.J. (1981). The managerial behavior of high school principals. *Education Administration Quarterly*, *17*, 69–90.

Marzano, R.J., Waters, T., & McNulty, B. (2005). *School leadership that works: From research to results*. Aurora, CO: ASCD and McREL.

McCaffrey, D., Lockwood, J.R., Koretz, D., & Hamilton, L. (2003). *Evaluating value-added models for teacher accountability*. Washington, DC: RAND.

McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T.A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*, 67–101.

Miskel, C. (1982). *An analysis of principal effects*. Unpublished speech to the National Graduate Student Seminar in Educational Administration, Princeton, NJ.

O'Day, K. (1983). *The relationship between principal and teacher perceptions of principal instructional management behavior and student achievement* (Unpublished doctoral dissertation). Northern Illinois University, Normal, IL.

Peterson, K.D. (1977). The principal's tasks. *Administrator's Notebook*, *26*(8), 2–5.

Pitner, N. (1982, March). *The Mintzberg method: What have we really learned?* Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Pitner, N.J., & Hocevar, D. (1987). An empirical comparison of two-factor versus multifactor theories of principal leadership: Implications for the evaluation of school principals. *Journal of Personnel Evaluation in Education*, *1*, 93–109.

Porter, A.C. (2005). Prospects for school reform and closing the achievement gap. In C.A. Dwyer (Ed.), *Measurement and research in the accountability era* (pp. 59–95). Mahwah, NJ: Lawrence Erlbaum Associates.

Robinson, V.M.J., Lloyd, C.A., & Rowe, K.J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, *44*, 635–674.

Rowan, B., Dwyer, D., & Bossert, S. (1982, March). *Methodological considerations in the study of effective principals*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Rubin, D. (1987). *Multiple imputation for non-response in surveys*. New York, NY: Wiley.

Sammons, P., Gu, Q., Day, C., & Ko, J. (2011). Exploring the impact of school leadership on pupil outcomes: Results from a study of academically improved and effective schools in England. *International Journal of Educational Management*, *25*, 83–101.

Sanders, W.L. (2006, October). *Comparisons among various educational assessment value-added models*. Paper presented at The Power of Two – National Value-Added Conference, Columbus OH.

Sanders, W.L., Sexton, A.M., & Horn, S.P. (1997). The Tennessee Value-Added Assessment System: A quantitative outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid education measure?* (pp. 137–162). Thousand Oaks, CA: Corwin Press.

Shouse, R.C. (1996). Academic press and sense of community: Conflict, congruence, and implications for student achievement. *Social Psychology of Education*, *1*, 47–68.

Silins, H.C., Mulford, W.R., & Zarins, S. (2002). Organizational learning and school change. *Educational Administration Quarterly*, *38*, 613–642.

Spillane, J.P., & Zuberi, A. (2009). Designing and piloting a leadership daily practice log: Using logs to study the practice of leadership. *Educational Administration Quarterly*, *45*, 375–423.

Stein, M., & Spillane, J. (2003). *Research on teaching and research on educational administration: Building a bridge*. Paper prepared for the AERA Division A Task Force for the Development of an Agenda for Future Research on Educational Leadership.

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, *46*, 31–56.

Tekwe, C.D., Carter, R.L., Ma, C.-X., Algina, J., Lucas, M.E., Roth, J., … Resnick, M.B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, *29*, 11–36.

Van de Grift, W. (1990). Educational leadership and academic achievement in elementary education. *School Effectiveness and School Improvement*, *1*, 26–40.

Waters, T., Marzano, R.J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement. A working paper*. Aurora, CO: Mid-Continent Research for Education and Learning.

Witziers, B., Bosker, R.J., & Krüger, M.L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly*, *39*, 398–425.

Yukl, G. (1999). An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *Leadership Quarterly*, *10*, 285–306.