Leadership for Learning: Does Collaborative Leadership Make a Difference in School Improvement?

Philip Hallinger and Ronald H. Heck

Educational Management Administration & Leadership 2010 38: 654
DOI: 10.1177/1741143210379060

Article

emal

Educational Management
Administration & Leadership
38(6) 654–678

DOI: 10.1177/1741143210379060
emal.sagepub.com
SAGE

Leadership for Learning: Does Collaborative Leadership Make a Difference in School Improvement?

Philip Hallinger and Ronald H. Heck

Abstract

Although there has been a sizable growth spurt in empirical studies of shared leadership over the past decade, the bulk of this research has been descriptive. Relatively few published studies have investigated the impact of shared leadership on school improvement, and even fewer have studied effects on student learning. This longitudinal study examines the effects of collaborative leadership on school improvement and student reading achievement in 192 elementary schools in one state in the USA over a 4-year period. Using latent change analysis, the research found significant direct effects of collaborative leadership on change in the schools' academic capacity and indirect effects on rates of growth in student reading achievement. In addition, the study identified three different growth trajectories among schools, each characterized by variations in associated school improvement processes. The study supports a perspective on leadership for learning that aims at building the academic capacity of schools as a means of improving student learning outcomes.

Keywords

collaborative leadership, educational change school accountability, school improvement, student learning

Introduction

This article focuses on understanding whether and how collaborative leadership makes a difference in elementary school improvement and student learning. A panel of internationally recognized scholars recently offered their perspectives on the state-of-the-art on shared school leadership in a volume entitled, *Distributed Leadership According to the Evidence* (Leithwood et al., 2009). In reflecting on this body of work in the final chapter, the editors offered the following conclusion:

Corresponding author:
Philip Hallinger, The Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change, Hong Kong Institute of Education, 10 Lo-Ping Road, Tai Po, NT, Hong Kong
Email: Hallinger@gmail.com

> By this point in the text, most of our 'hard nosed' readers will be experiencing profound disappointment at the lack of serious effort in the text to assess the contribution of greater leadership distribution to the long list of desirable outcomes typically invoked by advocates—greater student learning, more democratic practices, greater commitment by staffs to the mission of the organization, increased professional development for a wider range or organizational members, better use of intelligence distributed throughout the organization outside those in formal roles, and the like. We have considerable sympathy for such disappointment but have come to the grudging conclusion that research focused on outcomes would have been premature, at least until quite recently. (Leithwood et al., 2009: 280)

The challenge to examine the impact of shared leadership on important school processes and learning outcomes represents the problem space addressed in this article. More specifically, we ask three related research questions. First, does collaborative leadership contribute to school improvement and student achievement? Second, is there an identifiable set of patterns in how schools improve (or decline) in their academic performance over time and, if so, how are these patterns related to change in the schools' leadership and school organizational processes?

We address these questions in the context of a longitudinal, time-series study of school improvement in 192 elementary schools in one state in the USA. The study employed annual surveys of teachers and parents as a means of understanding patterns of change in the strength of collaborative leadership and academic capacity in their schools over a 4-year period. These perceptions were then compared with growth in reading achievement of a cohort of 12,480 elementary school students as they moved from Grade 3 through to Grade 5. This approach allowed us to assess how changes in leadership were associated with patterns of change in the capacity of schools to improve, and subsequent rates of growth in reading achievement.

Three features of the study deserve explicit mention at the outset. First, the study used student reading achievement as the focal measure of school performance. There are many other useful indicators of school performance, and some scholars have justifiably questioned why one would even choose to study the impact of leadership on a distal variable such as the academic achievement of students (D.M. Mcinerney, pers. Comm.., 7 January 2009). Indeed, prior research supports the view that principal leadership effects on school learning are largely indirect, mediated by a variety of school level factors (Bell et al., 2003; Bush and Glover, 2003; Hallinger and Heck, 1996a, 1996b; Leithwood et al., 2004, in press; Robinson et al., 2008). Yet, the accountability focus that dominates international discussions of educational reform has raised student achievement to a preeminent position in the goal hierarchy of many school leaders. Moreover, given a growing global interest in 'leadership for learning', we suggest that it should be a priority to explore the effects of leadership both on organizational variables and student learning outcomes.

Second, we assert that a comprehensive model of leadership and school improvement must include features of the school's external environment and internal organizational processes (Bossert et al., 1982; Hallinger and Heck, 1996a; Ogawa and Bossert, 1995). This study employs a conceptual model in which the effects of collaborative leadership on reading achievement are subject to the influence of selected environmental factors and mediated by the school's academic capacity (Bossert et al., 1982; Hallinger and Heck, 1996a, Pitner, 1988).

Finally, our interest in school performance lies explicitly in monitoring how leadership contributes to changes in school performance improvement over time. This longitudinal approach

contrasts with the predominant use of cross-sectional research designs in prior studies of school leadership effects (Hallinger et al., 1996; Heck et al., 1990; Heck and Hallinger, 2005; Leithwood and Jantzi, 1999; Luyten et al., 2005; Pounder et al., 1995; Wiley, 2001). We suggest that a longitudinal time series research design provides incrementally greater leverage over several methodological and theoretical problems that have persisted in this field of inquiry (Hallinger and Heck, 1996b).

Conceptual Framework

In this section, we discuss the conceptual model that frames the study. We begin by presenting the conceptualization of 'leadership for learning' that was employed in this study. Then we discuss how the relationship between school leadership, school academic capacity and student achievement was modeled in the research.

Conceptualizing 'Leadership for Learning'

The literature in organizational theory often defines leadership as an influence process that shapes the behavior of individuals and groups towards the attainment of goals (Yukl, 2006). Who exercises influence, how goals are defined and the means by which individuals, teams and organizational units move towards goals have been the subject of extensive theoretical and empirical examination (Bass and Bass, 2008; Yukl, 2006). While this study's perspective on leadership falls within this broad tradition of leadership research, it is also more specifically oriented towards the practice of leadership in 21st century educational organizations (Leithwood et al., 2009; MacBeath et al., 2008). We highlight two dimensions of the conceptualization of leadership used in this study.

First, this study focused explicitly school leadership that is team-oriented or collaborative (Marks and Printy, 2003; Mulford and Silins, 2003). This refers to school-wide leadership exercised by those in management roles (for example, the principal, assistant principals, department heads) as well as others (for example, teachers, parents, staff, students). Shared leadership, or what we will refer to in this article as collaborative leadership, thus encompasses both formal and informal sources of leadership, and conceptualizes leadership as an organizational property aimed at school improvement (Ogawa and Bossert, 1995; Pounder et al., 1995).

As noted at the outset of the article, empirical research that examines the impact of shared forms of leadership on school performance is scarce (Marks and Printy, 2003; Pounder et al., 1995). The data in this study came from one state in the USA in which new educational policies were explicitly designed to support the implementation of collaborative school leadership. School councils were mandated and both teachers and parents were expected to be involved in the development and implementation of school improvement plans.

The second dimension of leadership that we wish to highlight concerns the means by which leadership impacts school performance. A large body of international research supports the view that school leadership can have a significant indirect impact on student learning outcomes (Bell et al., 2003; Bossert et al., 1982; Hallinger and Heck, 1998; Leithwood et al., 2006; Robinson et al., 2008; Witziers et al., 2003). Notably, in recent years scholars increasingly assert that school leadership achieves this impact by shaping conditions that build school capacity for change and foster effective teaching and learning (Hallinger and Heck, 1996a; Leithwood et al., 2004, 2006, in press; Robinson et al., 2008; Southworth, 2002). The current research agenda in this field is, therefore, geared towards deepening our understanding of the means or paths through which

leadership achieves improvement in teaching and learning (Hallinger and Heck, 1996a; Leithwood et al., in press; Robinson et al., 2008).

In recent years, the phrase 'leadership for learning' has gained international currency (MacBeath et al., 2008; Robinson et al., 2008). In our view, this approach to school leadership represents a blend of two earlier leadership conceptualizations: instructional leadership and transformational leadership (Hallinger, 2003). Drawing inspiration from a body of work on instructional leadership, leadership for learning signals, among other things, the critical role that leadership plays in creating and sustaining a school-wide focus on learning (Hallinger and Murphy, 1985; Hallinger et al., 1996; Heck et al., 1990; Marks and Printy, 2003). It further highlights the importance of learning, not only for students but also for teachers and staff (Barth, 1990; Fullan, 2001; Leithwood et al., in press; Robinson et al., 2008). This capacity-building perspective is especially supported by findings from studies of transformational school leadership (Leithwood and Jantzi, 1999; Marks and Printy, 2003; Mulford and Silins, 2003).

Three areas of focus were incorporated into the *means of leadership:*

- Vision—making decisions to facilitate actions that focus the energy of the school on improving student outcomes and fostering commitment.
- Governance—empowering staff and encouraging participation.
- Resource allocation—obtaining and allocating resources to support teaching and learning.

Proponents of shared leadership also suggest that collaborative leadership has the potential to account for the broader range of naturally occurring leadership processes that exist in schools beyond the formal leadership exercised by principals (Barth, 1990, 2001; Harris, 2003; Lambert, 2002; Ogawa and Bossert, 1995). This perspective is reflected in the growing interest expressed in 'distributed school leadership' (Gronn, 2002; Leithwood et al., 2009). Scholars have suggested that acknowledging and developing the broader leadership capacity in schools may hold the key to unlocking the store of leadership potential grounded in instructional expertise that principals are often unable to provide (Barth, 1990, 2001; Crowther et al., 2008; Donaldson, 2001; Fullan, 2001; Gronn, 2002; Grubb and Flessa, 2009; Hall and Hord, 2001; Leithwood et al., 2009; Marks and Printy, 2003).

With this in mind, this study examines collaborative leadership drawn from a variety of sources, including but not limited to the principal. This provides the conceptual rationale behind the study's operational construct of collaborative school leadership. As described later in the article, survey items were selected to reflect the *means* by which collaborative leadership achieves results as well as the extent to which varied *sources* of leadership were accessed in the school.

Modeling Collaborative Leadership in the School Improvement Process

The question of whether a particular school and its teachers have a substantial effect on student achievement is central to systemic efforts to increase accountability for student learning. Students typically attend a school over a period of several years, during which they receive instruction from multiple teachers. Student learning, therefore, depends at least in part on the quality of teaching across classrooms. From this perspective, school improvement represents a *dynamic* process in which schools seek to develop the *breadth and density* of instructional expertise among their teachers.

Yet each school starts out at a different point on its school improvement journey (Jackson, 2000). At any point in time, the school's efforts to improve are shaped by multiple factors, including its intake of students, quality of faculty, academic organization, school culture and social organization (Bossert et al., 1982). Although leadership is often viewed as a catalyst for change, we suggest that the 'particular location' of each school in its journey of school improvement creates the need for and shapes the behavior of school leadership. Leadership, therefore, not only impacts school improvement, but is also shaped by the context in which it is exercised (Bossert et al., 1982; Hallinger and Murphy, 1986; Jackson, 2000; Leithwood et al., 2006; Luyten et al., 2005; Southworth, 2003).

Increasingly, scholars have begun to view school improvement within a framework of organizational learning (Mulford, 2007). For example, Mulford and Silins (2003) highlight how the complex interaction of organizational systems affects the capacity of schools to learn and change. This supports our own view that the empirical study of school improvement requires *dynamic* models that take into account changing relationships among relevant organizational processes over time (Blalock, 1989; Kelly and McGrath, 1988; Langlois, and Robertson, 1993). Given this perspective, we suggest that empirical studies that employ cross-sectional designs are ill-equipped to shed light on issues of *school improvement*. After all, school improvement, by definition, entails a change in the state of the organization over some period of time (Luyten et al., 2005).

Figure 1 presents our proposed model of how *changes* in school context, collaborative leadership, and school academic capacity are related to *changes* in student learning. The model highlights three features of data that must be incorporated into data analysis. First, data structures must reflect the multilevel (or nested) structure of school organizations (Hill and Rowe, 1996; Ogawa and Bossert, 1995). Second, in longitudinal studies, the repeated observations that describe changes in individual students or changes in schools processes also represent nested data structures (that is, within individuals or within organizational units). Third, longitudinal models require specification of a temporal sequence of relationships among the measured variables.

In this study, we employ multilevel latent change analysis (LCA), a variant of structural equation modeling (SEM), to examine changes in leadership, school academic capacity and student reading outcomes over a four-year period. In the LCA approach, repeated observations of student outcomes can each be represented by two correlated latent (or underlying) factors (Raykov and Marcoulides, 2006). These are represented as ovals in Figure 1. The *level factor* represents the level of a particular variable (for example, leadership, academic capacity or student achievement) at a chosen point in time. The *shape factor* represents change or growth in the variable over a particular interval.

For example, in Figure 1 we represent student reading achievement at two organizational levels (that is, the within-school [or student level] and the organizational [or school level]). We propose that initial student achievement levels as well as rates of growth are parameters that vary randomly within the population of schools. The subsequent focus of the research is to explain this variability in initial achievement and growth by monitoring sets of static and dynamic contextual and organizational variables.

At the school level in Figure 1, we define similar latent *level* (initial status) and *shape* (growth) factors describing collaborative leadership and school academic capacity. These factors serve as mediating organizational processes between schools' contexts (for example, student composition and enrollment size, staffing resources) and their outcomes (that is, reading

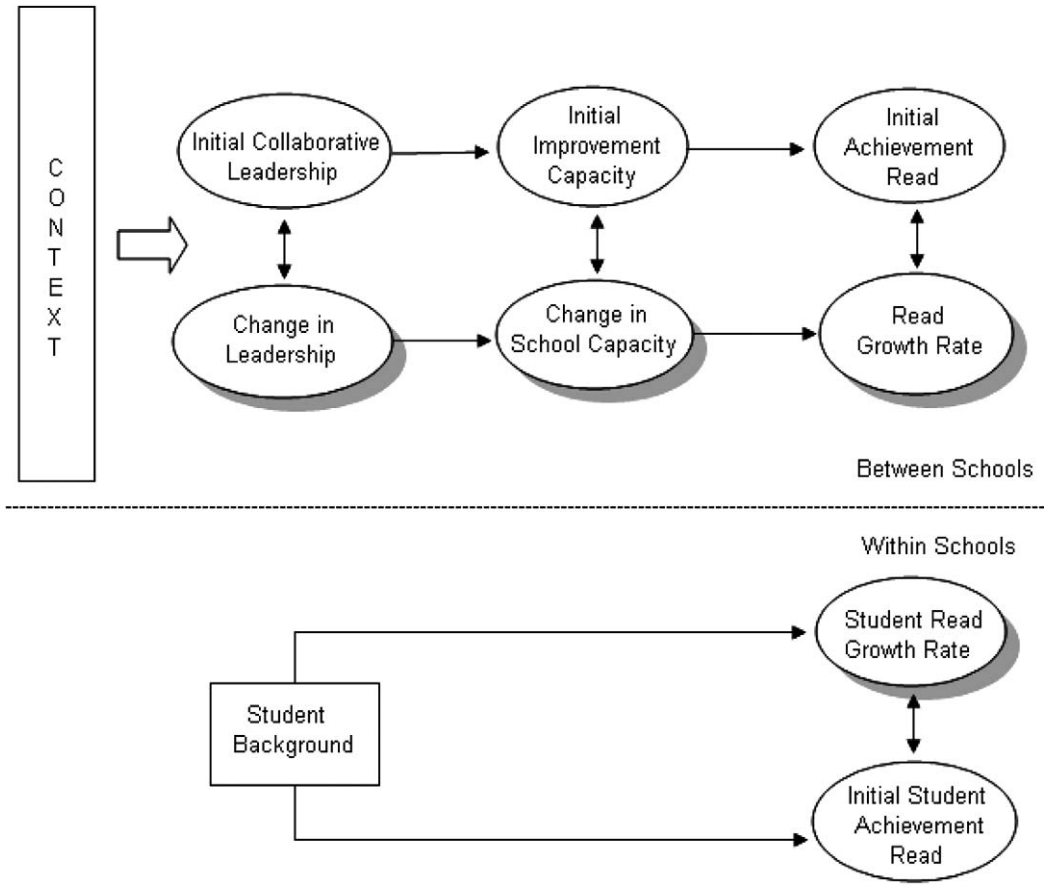

Figure 1. Conceptual model of school improvement leadership and student learning

achievement). The model proposes that changes in levels of collaborative leadership in a given school will impact the school's academic capacity. As the school builds its academic capacity over time, we would expect to see subsequent changes in teacher practices, student behavior and learning outcomes. The initial status leadership and academic capacity factors are assumed to be causally prior (defined in year 1) to initial achievement (defined in year 2). They are proposed to affect initial reading achievement indirectly and directly, respectively. The component variables in the dynamic model are proposed to affect student growth rates in a similar fashion.

Two-headed arrows between the level and shape factors in Figure 1 indicate expected negative correlations between the initial level of a particular variable and its subsequent change. For example, we expect that a school with a high initial level of collaborative leadership may change *less* over time than a school that starts out with low initial level of collaborative leadership due to ceiling effects. Our model implies, therefore, that the leadership and capacity-building trajectories of individual schools (and the reading trajectories of individual students) have common algebraic forms, but that not every school has the same trajectory. We

subsequently identify three emergent subsets of schools experiencing similar patterns of growth in reading and explore differences among the three subsets related to variables in our theoretical model.

Research Questions

We propose three research questions in this study. The questions are framed within the conceptual model proposed earlier in the article and portrayed in Figure 1.

1. Does Collaborative Leadership Impact School Performance?

This question is aligned with past research on school leadership effects. It seeks to understand how leadership impacts school organizational processes (that is, academic capacity) as well as performance outcomes, specifically student achievement. We address this question by analyzing the relationship between leadership and school performance at a single point in time (that is, year two of the study). Conceptually, this research question is located within the *static* portion of the model in Figure 1.

Within this research question, we suggest three specific propositions. We propose that the initial level of collaborative leadership in schools will be *directly* and significantly related to the school's initial academic capacity. In turn, initial academic capacity will be *directly* and significantly related to initial student achievement in reading. We also propose the initial level of collaborative leadership will be *indirectly* and significantly associated with initial student learning levels (mediated by initial academic capacity).

2. Does Collaborative Leadership Impact the Improvement of School Performance Over Time?

This research question shifts the focus towards the *dynamic* portion of the model. It seeks to understand how *changes* in levels of collaborative leadership in schools carry over into *changes* in academic capacity and reading achievement. The analyses related to this question take advantage of the longitudinal data set by exploring patterns of change across schools over a period of several years.

We tested several propositions in relation to this question. We proposed that change in collaborative leadership will be *directly* and significantly related to change in schools' academic capacity. In turn, changes in academic capacity will be *directly* and significantly related to improvements in student reading achievement. We also proposed that change in collaborative leadership will be *indirectly* and significantly related to change in reading achievement. Finally, we proposed that the size of leadership and academic capacity effects on outcomes will be significantly larger on student growth rates than on initial achievement levels.

3. How Do Schools Differ in Their Improvement Over Time and How Are Those Differences Related to Changes in School Leadership and Capacity?

This last research question derives from our earlier discussion of school improvement as a journey. From this perspective, at the beginning of the study, each school was at a particular point in a process of growth (or decline). This research question sought first to explore whether there were

identifiable growth trajectories (or patterns) in improvement among the 192 schools. Second, we were then interested to see if the organizational processes associated with school improvement (that is, leadership and academic capacity) vary in any systematic fashion for schools with different growth trajectories (for example, high-growth or low-growth schools).

Research Design and Method

This study employed a longitudinal panel time-series design (Cook, 2002) covering a period of 4 years. Our goal in this research was to examine how changes in leadership and academic capacity over time might be related to patterns of growth in student achievement. Although time-series studies are superior to cross-sectional studies in their ability to shed light on possible causal relationships, they still lack the power of experimental designs (Cook and Campbell, 1979). Despite this limitation, we note that longitudinal panel studies are recommended in domains where experimental manipulation of key variables is difficult or impractical (Marsh and Cravens, 2006). As Podsakoff, a prominent methodologist studying general leadership effects has noted, this makes longitudinal panel studies a 'method of choice' when conducting large-scale studies of leadership effects (Podsakoff, 1994; Williams and Podsakoff, 1989).

Data were collected from students and teachers in 192 elementary schools over a 4-year period. We captured changes in school processes through surveys given to each school's teachers on three occasions (years 1, 3 and 4). Where surveys are repeated over time with a high level of consistency between items, sequential measures may be used to estimate changes in a population (Davies, 1994). Achievement data from the student cohort were collected in years 2, 3 and 4. Unequal spacing of observations and nonlinearity can be incorporated into a LCA model without compromising quality of data analysis (Raykov and Marcoulides, 2006).

Data

Data were collected from a random sample ($N = 12{,}480$) of students drawn from a Grade 3 student cohort that was subsequently observed over a three-year period (that is, Grade 3 through Grade 5). We highlight the fact that data on the same students were incorporated into the data analysis for the duration of the study. This requirement is an essential condition for the types of analyses used for this longitudinal study and that we describe later in this section.

The students were enrolled in 192 public elementary schools. Background data were as follows: female, 49 percent; participation in federal free/reduced lunch program, 45 percent; receiving English-language services, 7 percent; receiving special education services, 11 percent; minority, 50 percent and changed schools, 16 percent. One advantage of growth modeling is that missing data (that is, less than 5 percent) and student mobility can be incorporated directly into the analysis, which reduces parameter bias that would result from eliminating these students (Peugh and Enders, 2004).

Operationalized Model of School Improvement Effects

The conceptual model described earlier was operationalized through explicit measurement of the control, explanatory and outcome variables included in Figure 1.

Control Variables. Student background variables included female (coded 1, male coded 0), low socioeconomic status (SES) (that is, participation in the federal free or reduced lunch program

coded 1, else coded 0), special education services (coded 1, else coded 0), minority (coded 1 = minority by race/ethnicity, else coded 0), English-language learning (ELL) services (coded 1, else coded 0) and changed schools (coded 1, else coded 0).

Context indicators describe initial school contexts during the first year of the study (2002/3), unless otherwise noted. *School size* was defined as the number of students enrolled for the school year. *Student composition* was defined as a composite variable by combining several relevant student demographics to create a weighted school indicator (using principal components analysis). The variables included the percentage of children receiving free or reduced lunch, percentage of students receiving ELL services, and the percentage of racial/ethnic minority students. Larger positive values represent school settings where these percentages of students were higher. *Teacher experience* was defined as the average number of years teaching of teachers at each school. *Teacher quality* was defined as the percentage of teachers at each school who were fully qualified according to No Child Left Behind and state certification requirements. *Teaching staff stability* was defined as the percentage of teachers in each school who had been at the school for five years (that is, assessed in year 4). *Principal stability* was defined as whether the same principal (coded 1, else = 0) was at the school during the four years of the study.

Explanatory Variables: Collaborative Leadership and School Academic Capacity. The collaborative leadership and school academic capacity constructs were defined from survey items administered initially on two-year cycles and now annually within the state. The survey is given at each school to all certificated staff, all grade five students, and a random sample of parents (that is, approximately 20 percent across grade levels within each school). Information from three successive teacher surveys was used to measure the academic capacity variable. The collaborative leadership variable was defined using additional information from the three successive teacher and parent surveys.[1] Items defining the constructs were measured on five-point, Likert-type scale. Indicators were expressed as the percentage of positive agreement with each statement. Higher percentages reflect more favorable perceptions. Cronbach's alpha (α) was used to assess the reliability of each subscale.

Collaborative leadership was measured by a subscale of five items describing teacher perceptions of leadership exercised within the school ($\alpha = 0.82$) and a corresponding subscale of five items describing parent perceptions of school leadership and their own personal involvement in improving education at the school ($\alpha = 0.88$). The stem used for these items was 'To what extent does school leadership ...' The state survey items were designed to reflect three specific aspects of collaborative leadership within each school (with items paraphrased in parentheses):

- Make collaborative decisions focusing on educational improvement (that is, support efforts and decisions of members of the school community that focus on student learning; ensure teachers have a major role in decisions about curriculum development in the school; enable administrators, teachers, and staff work together effectively to achieve our school's goals).
- Emphasize school governance that empowers staff and students, encourage commitment, broad participation, and shared accountability for student learning (that is, provide opportunities for parents to participate in important decisions about their children's education through a variety of venues; encourage parent involvement in the school improvement process; ensure teachers can freely express input and concerns to the administrators; ensure parents can freely express input and concerns to the administrators; provide opportunities for teachers to plan and make school decisions).

- Emphasize participation in efforts to evaluate the school's academic development (for example, ensure adequate resources are available to the school to develop its educational programs; provide regular opportunities for all stakeholders to review the school's vision and purpose).

The 10 teacher and parent items all loaded substantively on the collaborative leadership factor.[2] Factor scores describing the measurement of the leadership factor on each of the three occasions (summarized in the results section) were then saved and used to define the LCA model of collaborative school leadership. While this scale fails to measure certain potentially important facets of instructional leadership (Hallinger and Murphy, 1985; Leithwood et al., in press; Robinson et al., 2008), we believe that it still represents a valid construct of collaborative leadership oriented towards school improvement.

School academic capacity ($\alpha = 0.94$) was formed by combining four subscales measured from the teacher surveys. Factor scores were used in equating the measurement of academic capacity across the three occasions. The scores were saved and used to define the latent change model of academic capacity. The subscale alphas and items of the subscales were as follows.

- *Standards emphasis and implementation* (learn, $\alpha = 0.91$; school's educational programs are aligned to the state content and performance standards; teaching and learning activities are focused on helping students meet the state content and performance standards; school prepares students well for the next school; students and parents are informed about what students are expected to learn; school has high academic and performance standards for students; classroom instruction includes active participation of students; curriculum and instructional strategies emphasize higher-level thinking and problem solving; instructional time is flexible and organized to support learning; teachers provide a variety of ways for students to show what they have learned; students learn to assess their own progress and set their own learning goals; students are provided with multiple ways to show how well they have learned; homework assignments are appropriate, productive, and reflective of adopted learning standards; assessment results are used to plan and adjust instruction).[3]
- *Focused and sustained action on improvement* (improve, $\alpha = 0.83$; school clearly communicates goals to staff, parents and students; vision and purpose are translated into appropriate educational programs for children; school seeks ways to improve its programs and activities that promote student achievement; teachers know what the school learner outcomes are; teachers expect high quality work; school's vision is regularly reviewed with involvement of all stakeholder groups; changes in curriculum materials and instructional practices are coordinated school-wide and I [teacher respondent] am involved in the school improvement process).
- *Quality of student support* (support, $\alpha = 0.85$; standards exist for student behavior; discipline problems are handled quickly and fairly; school environment supports learning; open communication exists among administrators, teachers, staff and parents; teachers feel safe at school; teachers and staff care about students; administrators, teachers, and staff treat each other with respect; I provide students with extra help when they need it; programs meet special needs of students; school reviews support services offered to students).
- *Professional capacity of the school* (capacity, $\alpha = 0.80$; teachers are well qualified for assignments and responsibilities; leadership and staff are committed to school's purpose; staff development is systematic, coordinated and focused on standards-based education; systematic evaluation is in place).

Literacy/Reading Achievement. The standardized test used in the study was constructed to measure state-developed reading/literacy content standards. The test consisted of constructed-response items and standardized test items from the Stanford Achievement Test (Edition 9). For literacy/reading, there were three curricular strands consisting of 47 items (that is, comprehension process, conventions and skills, and literary response and analysis). Student scores (re-scaled to range from 100 to 500) considered patterns of right, wrong and omitted responses over successive years and were equated across the three years to enable the measurement of academic growth.

Data Analysis

Data analysis proceeded in two steps. First, we investigated the relationships among variables implied by our proposed model in Figure 1. In the LCA approach repeated observations on individuals over time (yt) can be expressed as a type of confirmatory factor analysis (or measurement model), where the level and shape of latent factors are measured by multiple indicators of y. We provide further details on the specification and testing of the model in the endnotes.[4] Student-level estimates were centered on their grand means. This results in school means that are adjusted for within-school differences in student composition. School-level estimates were also centered on their grand means, except for the dichotomous indicator of principal stability.

Second, we investigated whether variables in our proposed model might be useful in categorizing schools with different growth patterns in reading over time. We used discriminant analysis to explore the extent to which schools could be classified into separate groups according to their growth trajectories in reading achievement. We identified three groups of schools: (1) low-growth schools, or schools that experienced no student growth or negative growth over time (that is, 20 schools); (2) high-growth schools, or schools that experienced student growth considerably above the norm for schools in the data set (that is, 25 schools with growth of 31 or more scale score points per interval); and (3) average schools, or schools that were clustered primarily within one standard deviation above or below the mean for growth for schools in the data set.

Results

Descriptive statistics for the variables are provided in Table 1. These indicate that on average about 12 percent of the variability in students' reading achievement (and 12 percent of growth) lies *between schools*. This implies that a multilevel theoretical model could be useful in explaining differences in achievement levels and growth rates between schools. Table 1 also suggests that schools made considerable growth in reading over time (averaging about 23.7 scale score points over the first growth interval). Although factor means do not provide an indication of the amount of change that takes place over time, *t*-tests suggested that schools changed in average academic capacity (factor means = 0.00, 0.02, 0.13, respectively, $p < 0.05$) and average collaborative leadership (that is, factor means = 0.00, 0.04, 0.05, respectively, $p < 0.05$) over time.

Tests of the proposed model were conducted with Mplus 5.2 (Muthén and Muthén, 1998–2006). Our first concern is with the fit of the proposed model to the data. Adequacy of the fit of the proposed model can be determined by a number of different model fit indices (Marcoulides and Hershberger, 1997). The root mean square error of approximation (RMSEA) describes the amount of model discrepancy per degree of freedom. Values near 0.05 or lower generally indicate an adequate fit of the model to the data. The comparative fit index (CFI) compares the fit of the proposed model against a baseline (non-fitting) model, with values near 0.95 providing evidence of an

Table 1. Descriptive statistics for between-school ($n = 192$) and within-school ($n = 12{,}480$) variables in the model

Variable name	Mean	SD	Minimum	Maximum
Context				
Enrollment	495.88	243.67	42.00	1428.00
ELL (%)	8.45	9.02	0.00	61.00
Low SES (%)	50.49	22.63	0.00	97.00
Minority mean (%)	51.16	23.97	3.00	97.00
School composition	0.01	1.00	−1.92	2.24
Staffing				
Same principal	0.31	na	0.00	1.00
Staff stability (%)	57.28	14.13	10.00	93.00
Met licensing criteria (%)	84.05	16.29	12.10	100.00
School achievement				
Read 2004 (ICC = 12%)[a]	247.33	37.72	100.01	323.33
Read 2005 (ICC = 11%)	271.92	30.37	123.87	345.69
Read 2006 (ICC = 12%)	281.37	24.90	161.26	346.96
Read growth rate (ICC = 12%)	23.65	17.74	−25.15	92.87
Initial collaborative leadership				
Leadership (%)	65.50	10.30	29.00	86.20
Leadership factor scores[b]				
Year 1	0.00	0.14	−0.43	0.28
Year 3	0.04	0.13	−0.40	0.23
Year 4	0.05	0.14	−0.45	0.31
Initial school academic capacity				
Standards-based learning(%)	87.10	6.32	69.14	98.67
Student support system (%)	78.48	10.93	37.63	98.78
Professional capacity (%)	74.53	11.82	40.00	99.11
Focused school improvement (%)	78.41	11.25	47.22	97.35
Academic capacity factor scores[2]				
Year 1	0.00	0.24	−0.70	0.23
Year 3	0.02	0.22	−0.68	0.24
Year 4	0.13	0.10	−0.16	0.27
Student background				
Low SES	0.45	na	0.00	1.00
English services	0.07	na	0.00	1.00
Special education	0.11	na	0.00	1.00
Female	0.49	na	0.00	1.00
Minority	0.50	na	0.00	1.00
Changed schools	0.16	na	0.00	1.00

[a]Intraclass correlation (ICC) refers to the variance in outcomes between schools.
[b]Successive factor scores are significantly different from initial score.

adequate model fit. Model fit indices suggested the proposed model was a plausible representation of the data (for example, CFI = 0.972, RMSEA = 0.020).

Because the model provided an adequate fit to the data, we can turn our attention to the specific parameter estimates. Table 2 summarizes the results concerning within-school and between-school variables that explained differences in initial school achievement levels and annual growth rates.

Table 2. Standardized estimates, 95% Confidence Intervals (CI), and power of variables explaining student achievement and growth

	Standardized coefficients					
	Initial	CI	Power	Growth	CI	Power
School context						
Enrollment	0.02	(−.05, .23)	0.26	−0.08*	(−.15, −.01)	0.70
Student composition	−0.05	(−.20, .14)	0.30	−0.14*	(−.32, .04)	0.79
Teacher quality				0.12*	(.00, .26)	0.73
Staff stability				−0.01	(−.12, .10)	0.12
Principal stability				−0.03	(−.10, .05)	0.20
Hypothesis 1						
Academic capacity	0.12*	(.05, .20)	1.00			
Change in capacity				0.20*	(.15, .25)	1.00
Hypothesis 2						
Leadership (indirect effect)	0.02*	(.00, .05)	0.60			
Change in leadership (indirect effect)				0.10*	(.05, .15)	1.00
Student background						
Female	0.05*	(.03, .07)	0.31	0.02*	(.00, .04)	0.73
Low socioeconomic status (SES)	−0.04*	(−.06, −.02)	1.00	−0.05*	(−.07, −.03)	1.00
English language Learner (ELL)	−0.24*	(−.26, −.22)	1.00	0.10*	(.08, .12)	1.00
Special education	−0.20*	(−.22, −.18)	1.00	−0.07*	(−.09, −.05)	1.00
Minority	0.03*	(.00, .07)	1.00	−0.11*	(−.14, −.09)	1.00
Changed	0.05*	(.03, .07)	0.92	−0.08*	(−.10, −.06)	1.00

Note: *$p < .05$.

Figure 2 provides further information about between-school relationships in the model. The coefficients are standardized, which indicates the relative size of each variable's effect, with significance level set at $p = 0.05$. As Hedges (2008) notes, when reporting effect sizes, it is desirable to include estimates of uncertainties (for example, standard error or confidence intervals; see Table 2). When interpreting effect sizes, the level of analysis matters in multilevel populations. For example, a standardized effect that is small in accounting for existing variation at the student level (for example, 0.1 or 0.2) may be large in accounting for between-school variation (Hedges, 2008). It is therefore best to consider specific effects in relation to other effects at each level of the data hierarchy.

Notably, in Table 2, contextual indicators (that is, student composition, enrollment) were not related to initial school achievement levels; however, both student composition (standardized $\gamma = -0.14$, $p < 0.05$) and enrollment size (standardized $\gamma = -0.08$, $p < 0.05$) were related to student growth. Additionally, teacher quality was significantly related to growth in reading (standardized $\gamma = 0.12$, $p < 0.05$).

Does Collaborative Leadership Impact Initial School Performance?

The results in Figure 2 provide support for the first proposition that initial level of leadership would be related to initial levels of school academic capacity (standardized $\gamma = 0.19$, $p < 0.05$) and, in

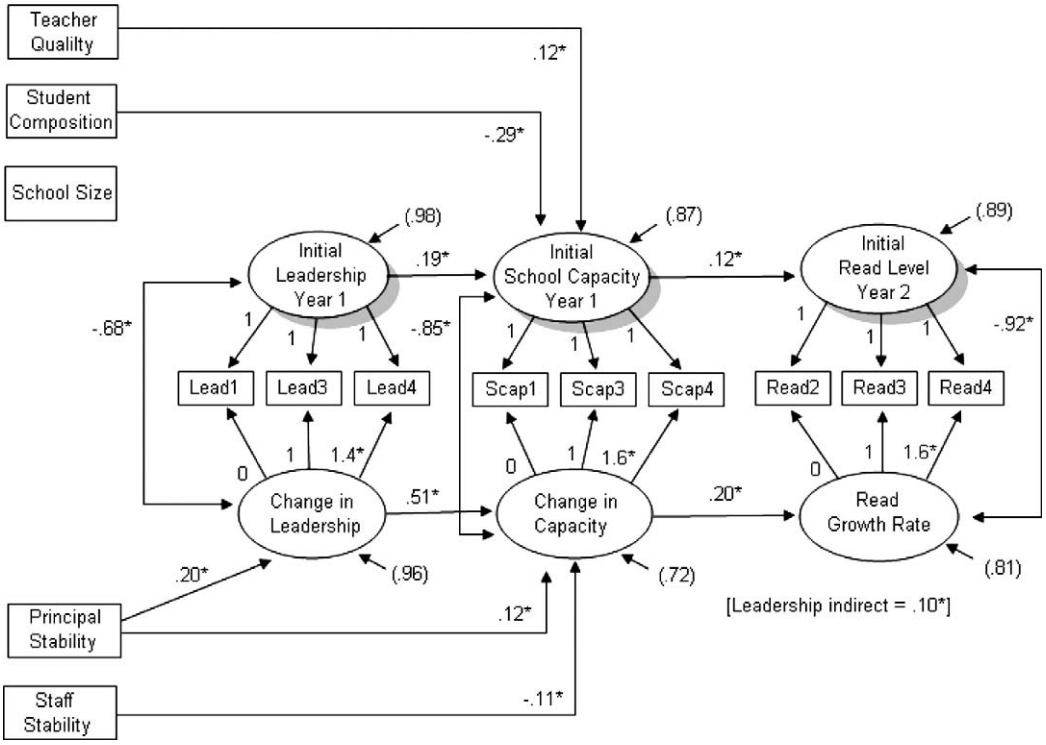

Figure 2. Between-school significant ($p < .05$) standardized effects

turn, that initial capacity would be significantly related to initial reading achievement (standardized $\gamma = 0.12$, $p < 0.05$). The second proposition indicated that initial collaborative leadership would be *indirectly* and significantly associated with initial student learning levels through initial academic capacity. We found evidence of a small indirect (but substantively unimportant) effect of initial collaborative leadership on initial reading outcomes (standardized $\gamma = 0.02$, $p < 0.05$). We note in passing that these results are quite consistent with the trend of results from prior cross-sectional studies of school leadership effects (Bell et al., 2003; Hallinger and Heck, 1998; Leithwood et al., 2004; Robinson et al., 2008).

Does Collaborative Leadership Impact Improvement of School Performance Over Time?

The second question focused on the *patterns of change* in the relationship among the variables over the 4-year period of the study. That is, would changes in leadership be positively associated with changes in school capacity and student growth in achievement? Longitudinal modeling would seek to identify whether patterns of association among the variables as they changed from year to year were significant.

Our first proposition was that change in collaborative leadership would have a significant direct effect on change in school academic capacity and that change in academic capacity would directly

affect growth in student learning. We proposed that relationships between initial processes would be weaker than subsequent relationships where schools have been successful in their efforts to improve. Since leadership is often portrayed as a catalyst for change, it follows that stronger perceptions of leadership would be positively associated with growth in school academic capacity. As proposed, we found positive change in collaborative leadership was significantly related to growth in academic capacity (standardized $\gamma = 0.51$, $p < 0.05$).

The relationship between change in academic capacity and student growth rates in reading was significant and substantial (standardized $\gamma = 0.20$, $p < 0.05$). Bloom and colleagues (2008) noted that student year-to-year gains on standardized tests can be summarized to provide a benchmark against which we can compare the effects of various school-level effects. On average, across a series of standardized reading tests, they found that students gained roughly 0.4 of a standard deviation per year during Grades 3 through 5 (Bloom et al., 2008). Using this as a rough benchmark, an increase in yearly student growth rates due to the effect of increasing academic capacity (0.20), could be viewed as a substantial enhancement to student learning (that is, approximately 50 percent).

As noted in our earlier discussion of the conceptual framework, the lack of a measure of classroom teaching represents a limitation in explicating the causal chain between changes in leadership, academic capacity and student learning. We, therefore, sought to check the validity of teachers' collective perceptions of changes in their schools' educational practices over time (as defined by the survey items). To do this, we compared their perceptions in year four of the study against students' collective perceptions of educational practices in their school for the same year. We noted substantial agreement in their perceptions ($r = 0.68$, $p < 0.05$, not tabled). We do, however, acknowledge that even this step does not fully satisfy the conditions for addressing the causality problem since we are unable to link students to particular classrooms from year to year in the dataset.

The second proposition stated that the combined effects of change in collaborative leadership on student growth rates in reading should be *indirect* rather than direct. The indirect effects of change in collaborative leadership (operating through the mediating change in academic capacity factor) on growth rates in reading achievement was significant (standardized $\gamma = 0.10$, $p < 0.05$).

Although the size of the indirect effect of collaborative leadership on student growth in reading may appear small, it is on a par with direct effects of teacher quality and initial school academic capacity on initial school achievement levels. These results also support the view that the effects of both collaborative leadership and academic capacity are stronger when reading growth is the outcome. Together these constructs accounted for about three-fourths of the 19 percent of growth variance in reading accounted for by the model (with 81 percent from other sources, shown in parentheses in Figure 1).[5] In contrast, initial leadership and academic capacity accounted for less than half of the 11 percent of variation in initial achievement accounted for by the model (with 89 percent due to other sources, also shown in parentheses in Figure 1). This highlights the added value of the growth model over a cross-sectional model of leadership effects on school improvement.

Is the Model Useful in Classifying Schools According to Their Academic Improvement?

The second part of our analysis focused on classifying schools according to their different patterns of student growth. The discriminant analysis revealed two significant functions ($p < 0.05$) that

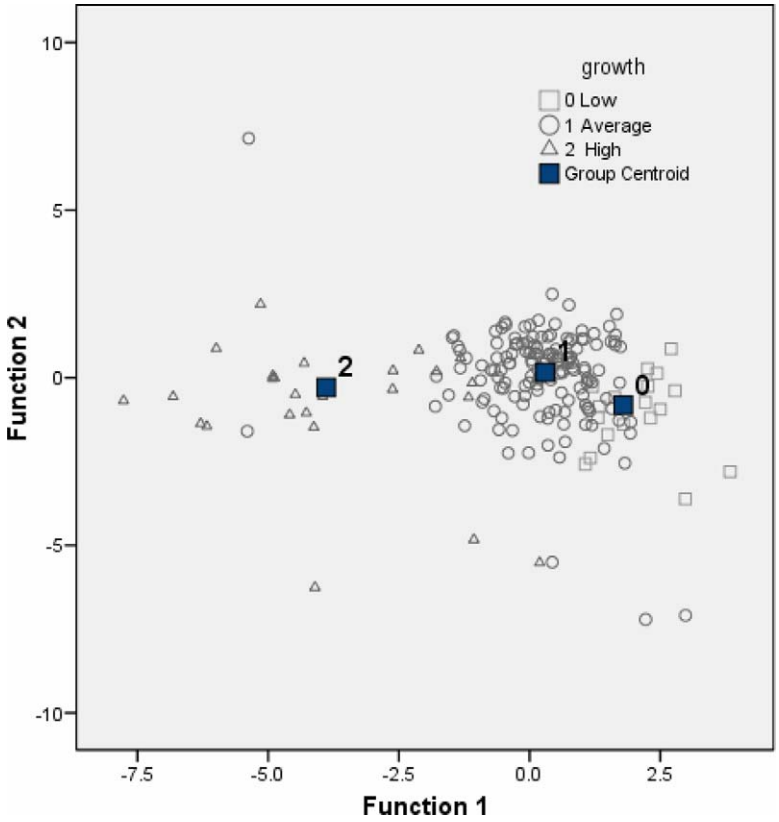

Figure 3. Classification of schools according to growth trajectories

helped separate schools according to their growth trajectories. The first function primarily separated high-growth schools (coded 2) from average (coded 1) and low-growth (coded 0) schools. This function accounted for about 80 percent of the observed variability among the groups. The second function, which accounted for the remaining 20 percent of the separation between groups, primarily separated low-growth schools from average-growth schools. A plot of the schools in multidimensional space is shown in Figure 3.

Structure coefficients (ranging between 0 and 1), which describe the degree of correspondence of each observed predictor to the underlying dimensions that separate the groups, are presented in Table 3. The size of the coefficient indicates its relative contribution in classifying schools into their appropriate groups. The positive or negative sign associated with the coefficient indicates which group of schools had the highest or lowest mean on the predictor. In this case, positive coefficients on the first function generally suggest lower means for high-growth schools compared to the other two groups of schools. More specifically, the coefficients suggest that high-growth schools were characterized by lower initial reading achievement levels (0.75), fewer students (0.12), and slightly less change in collaborative leadership (0.10) than low-growth or average-growth schools. In contrast, higher-growth schools were also differentiated by somewhat higher initial means for collaborative leadership (−0.12) and slightly greater change in capacity over time (−0.08) than the other groups of schools. Other variables contributed little to separating the schools with respect to the first function.

Table 3. Structure coefficients for classifying schools by growth

	Structure coefficients	
	Function 1	Function 2
Beginning read level	0.75	−0.01
Enrollment	0.12	−0.10
Initial leadership	−0.12	−0.30
Student composition	0.08	0.63
Change academic capacity	−0.08	0.42
Initial academic capacity	0.05	−0.47
Teacher quality	−0.01	0.26
Change in leadership	0.10	0.11

For the second function, positive coefficients suggest higher means for average-growth schools compared with low-growth schools. These coefficients indicate that average-growth schools tended to have greater percentages of low SES students, students receiving English services, and special education students (0.63); experienced greater change in academic capacity (0.42); had higher teacher quality (0.26); and experienced greater change in collaborative leadership (0.11) than low-growth schools. Negative coefficients indicate average-growth schools had corresponding lower initial means for academic capacity (−0.47) and collaborative leadership (−0.30) than low-growth schools.

Overall, the variables in the model were very useful in classifying 86 percent of the schools correctly (not tabled) according to their growth in reading (that is, including 70 percent of high-growth schools, 54 percent of low-growth schools, and nearly all of the average-growth schools). This supports the proposed model's validity in explaining differences in school reading growth. Moreover, this finding suggests that it could be fruitful to focus in future research on identifying patterns of leadership practice and school improvement strategies that are linked to the different groups of schools based on where they lie in their growth trajectories.

Conclusion

This research explored prominent issues concerning the impact of collaborative leadership on school performance. Building upon findings from prior research on school leadership effects (Hallinger and Heck, 1996a; Leithwood and Jantzi, 1999; Marks and Printy, 2003; Pitner, 1988; Robinson et al., 2008; Wiley, 2001; Witziers et al., 2003), the study employed a mediated-effects model that examined the effects of collaborative leadership on school-level academic capacity and student reading achievement. Our analysis of longitudinal data supports the view that collaborative leadership positively impacted growth in student learning indirectly through building the academic capacity in schools. The results also provide initial insight into patterns of growth that characterize different schools in their school improvement 'journeys'. We have suggested that although these findings are consistent with a substantial body of cross-sectional survey research on principal leadership effects, they also extend this knowledge base by focusing on collaborative leadership and employing longitudinal modeling. In this concluding section, we discuss limitations and implications of the study.

Limitations of the Research

Although the results generally confirmed the main research questions, they must be interpreted in light of several limitations. First, in time-series analysis, it still remains a challenge to separate growth in performance outcomes (that is, student achievement) produced by intended changes from those produced by natural trends and extraneous or unmeasured variables (Cook, 2002). Rival explanations could include random fluctuations at each time interval as well as other evolving conditions (for example, state or district policies, school and classroom conditions) that may simultaneously influence learning. Although we took steps to mitigate or test for the unanticipated effects of a variety of other factors, the limitations of non-experimental research remain relevant to the interpretation of results and most specifically to claims of causality within the model.

Thus, our time-series analysis would have been strengthened if it had been possible to find a set of 'control' schools that had not undergone any type of intervention and compare them with a set of schools that had implemented a specific school-improvement intervention (for example, change in collaborative leadership structures). In experimental and quasi-experimental research designs, planned interventions occur in selected organizational units in a series of 'before' and 'after' measurements and, hence, employ stronger controls for extraneous variables (Cook and Campbell, 1979). Nonetheless, large-scale experimentation is often impractical when conducting organizational studies in field settings (Cook, 2002; K. Leithwood, pers. comm., 12 May 2007).

In such cases, longitudinal panel studies represent a strongly preferred alternative to cross-sectional studies (Marsh and Cravens, 2006). Notably, our recent search of the empirical literature on school leadership effects found that the field remains dominated by cross-sectional studies (Luyten et al., 2005). Moreover, our direct comparison of *static* (that is, cross-sectional) and *dynamic* (longitudinal) models confirmed the superior ability of the latter approach to reveal underlying relationships among variables in the model.

Another key limitation of this study is the 'black box' that lies between academic capacity and student achievement. Although the study collected data from specific teachers about what they do in their own classrooms (that is, in the academic capacity subscale), due to limitations in the data, we were unable to link specific students to their teachers. Because the data from teachers were aggregated to the school level, we can only speculate about the extent to which school-level changes in leadership and collective academic capacity led to specific changes in teacher behavior that resulted in improved student achievement. As suggested above, it is also possible that unmeasured variables could account for the observed pattern of change in student learning.

Nonetheless, there is evidence that changes in school-level processes (for example, focus on academic improvement, increased professional capacity, alignment of curriculum and instruction) can have 'trickle down' effects on teachers' instructional practices. For example, in other research conducted with this same database, Heck (2009) determined that differences in the effectiveness of successive classroom teachers (accounting for about 10–11 percent of the total variability in outcomes) and the school's collective teaching effectiveness over time contributed meaningfully to reducing gaps in student learning in reading and math between schools. So, although the inclusion of specific observational data on classroom teaching practices would have provided greater confidence in the conclusions, we suggest that the results be interpreted in light of other research conducted in this field where the findings are largely in the same direction (Bell et al., 2003; Leithwood et al., 2006, in press; Robinson et al., 2008; Southworth, 2002).

A final limitation that we wish to highlight concerns the measurement of collaborative leadership. Although our operational measure, which sought information from both teachers and parents at each

school appeared reliable, it is possible, even likely, that the *nature of the distribution* of collaborative leadership varied across different schools. This *between-school* variability in the *sources and nature* of collaborative leadership makes interpretation of the collaborative leadership variable somewhat more difficult to interpret than in traditional studies of principal leadership effects. That is, we were unable to describe the ways in which the sources of collaborative leadership varied across schools and examine the resulting differences in effects on school improvement. Thus, our study does not offer in-depth insight into the range and variation of collaborative leadership practices across schools in the sample.

With this limitation in mind, we wish to refer briefly to a set of qualitative case studies that were conducted in 21 schools that demonstrated high levels of growth in reading. Due to space limitations, we did not include these data in the main body of the report. Nonetheless, we would note that the case studies support the view that 'high growth' schools were employing varied forms of collaborative leadership. In none of the 21 case studies schools was the principal alone providing the leadership for school improvement. By way of example, we offer one short vignette:

> Throughout all the school's efforts in making student achievement gains in reading, the faculty, staff and community members have come together as a school-wide professional learning community ... The school leadership group ... works together as a collaborative team to facilitate the school's curricular/instructional programs, student support and school operations ... At the teacher meetings, the principal meets with the teachers to plan and develop grade level plans, discuss English Language Arts and Reading and Science curriculum/instruction, analyze student work, review professional literature, and to hold student case reviews ... Family literacy has been the focus of parent workshops and classes. Workshops provide parents with training in supporting the growth and development of their child's love of reading and writing. (Interview from School 6, pp. 16–17)

These and similar statements about the linkage between collaborative leadership and improved instruction appear throughout the case studies of the high-growth schools. This perception is further corroborated by noting that almost 50 percent of these 21 schools increased 6 percent or more in distributed leadership over the 4-year period (against only 27 percent of the other schools in the full dataset).

In sum, we acknowledge that due to these limitations this study does not illuminate how differences in forms of leadership collaboration and distribution impact school improvement and student learning. The accumulation of knowledge in this emergent field of research on the impact of shared forms of leadership will require a variety of methods and an abundance of empirical investigations.

Implications

With these limitations in mind, we suggest that the methods used in this study represent an *incremental* advance in the state-of-the-art of research on school leadership effects. While the longitudinal time-series approach employed in the research is unable to provide a conclusive answer to all questions of interest regarding school leadership effects, it does offer more confidence concerning causal relationships than cross-sectional surveys or case studies. Indeed, we suggest that the ability to employ longitudinal data to model relationships as they change over time is a necessary condition in empirical studies of *school improvement* (Luyten et al., 2005). While this conclusion may appear self-evident, we note that many studies commonly referenced in the school improvement literature have employed cross-sectional designs. In some instances, even case studies of

school improvement have consisted of the retrospective analysis of data collected at a single period of time.

The design of the current study allows us to extend the substantive findings from prior research on school leadership effects in several ways. First, the finding of *indirect leadership effects* on academic growth over time reinforces an important conclusion drawn in a series of influential reviews of research on school leadership effects (Bell et al., 2003; Bossert et al., 1982; Bush and Glover, 2003; Hallinger and Heck, 1998; Leithwood and Montgomery, 1982; Leithwood et al., 2004, 2006, in press; Robinson et al., 2008; Witziers et al., 2003). Prior research in this domain has been limited to drawing conclusions about the effects of school leadership on learning from studies that described the relationship among variables at a single point in time. This is the first study of which we are aware that has identified statistically significant, indirect effects of school leadership on student growth in learning within a *dynamic* model of school improvement encompassing a period of several years.

Second, the results offer additional insight into the indirect paths through which leadership impacts student learning (Hallinger and Heck, 1998; Leithwood et al., in press). More specifically, this study provides empirical support for the proposition that collaborative leadership contributes to school improvement through building the school's academic capacity in quite specific areas (Fullan, 2001; Robinson, et al., 2008). By academic capacity, we refer to a set of organizational conditions that impact what teachers do in classrooms to influence student learning. More specifically, we suggest that leadership acts as a driver in identifying needs and devising strategies to foster school-wide academic changes over time (that is, developing and sustaining a school-wide focus on learning, upgrading the curriculum, providing individualized support for teachers and students, improving the monitoring of student progress). The longitudinal analysis found that changes in collaborative leadership over time were directly associated with these types of changes in academic capacity and indirectly related to student growth in reading achievement.

Third, the results contribute to a nascent literature on the different patterns of growth that may characterize schools during 'the school improvement journey' (Jackson, 2000). The findings suggest that there are indeed several different patterns in the 'growth trajectories' of schools. The initial evidence from this study further indicates that different strategies may be in order for schools that are at different points in their journeys.

Although this finding may also appear self-evident, we are unaware of research that has inquired into the effects of matching interventions, policies or improvement strategies to the differentiated needs of schools. The ability to match improvement strategies to a limited set of 'school improvement states or conditions' (for example, turnaround situations, plateaued growth and so on) would appear to be a worthy research challenge in an era of evidence-based practice. We urge others to build on this finding.

Finally, these findings represent an early contribution to the emerging *empirical knowledge base* on the effects of shared forms of school leadership (Leithwood et al., 2009; Marks and Printy, 2003; Mulford and Silins, 2003; Pounder et al., 1995; Timperly, 2009). The study takes note of sources of school leadership beyond the principal, and explicitly links a more team-oriented and collaborative approach to school leadership with capacity building strategies designed to impact teaching and learning.

Development of the knowledge base in a given field of professional practice takes place incrementally through series of studies that examine the phenomenon from different perspectives and using a variety of methods. Having engaged in research on school leadership effects for over 25 years, we are encouraged by evidence of the dual trend of stronger methodologies being applied

in empirical studies (Leithwood and Jantzi, 1999; Marks and Printy, 2003; Mulford and Silins, 2003; Wiley, 2001) and a concerted effort to synthesize knowledge in more systematic fashion (Bell et al., 2003; Leithwood et al., 2004, 2006, in press; Luyten et al., 2005; Robinson et al., 2008; Southworth, 2002; Witziers et al., 2003). In conclusion, although the current study does not provide a conclusive answer to the question posed in the title of this article, 'Does collaborative leadership make a difference in school improvement?', we hope that its incremental contribution offers useful direction to others engaged in this field of study.

Notes

1. Teacher return rates were 73.4, 76.4 and 75.6 percent, respectively; parent return rates were lower (that is, 22, 24 and 22, respectively) but represented a mean of approximately 40 parents per school for each survey year.
2. The teacher items (1–5) loaded on the leadership factor as follows: 0.81, 0.39, 0.60, 0.90 and 0.46, respectively. The parent item (1–5) loadings were 0.62, 0.56, 0.62, 0.30 and 0.81, respectively.
3. We defined a similar measure ($\alpha = 0.95$) consisting of student perceptions of standards implementation and learning in their classrooms during the last year of the study. We used this measure as a means of triangulating teachers' perceptions of changes in collective classroom practices at this same point in time. The scale consisted of 16 items (for example, classroom learning helps me reach content and performance standards; teachers help me learn how to think and solve problems; I have learned how to evaluate my own work and monitor my progress; my teacher tells me how I am doing and how I can improve; we learn by doing, not just listening; my teachers expect quality work).
4. Matrices and vectors facilitate the specification of models that can be relatively complex (for example, containing both observed and latent variables and relationships between variables that indicate both direct and indirect effects). The latent change model to represent individual i at time t can be written as

$$y_{it} = v_t + \Lambda_t \eta_i + K x_i + \varepsilon_{it}, \quad (1)$$

where y_{it} is a vector of reading outcomes for individual i at time t $(y_{i1}, y_{i2}, \ldots, y_{iT})'$, v_t is a vector of measurement intercepts, Λ_t is a p x m design matrix representing the change process, η_i is an n-dimensional vector of latent variables, $(\eta_{0i}, \eta_{1i} \ldots, \eta_{pi})'$, K is a p x q parameter matrix of regression slopes relating x_i covariates $(x_{1i}, x_{2i} \ldots, x_{pi})'$ to the latent factors, and ε_{it} represents time-specific errors which are contained in a covariance matrix (Θ). The factor loadings for the latent level and shape achievement factors are defined in the $\Lambda_t$factor loading matrix. The level factor loadings are fixed to 1.0. We incorporated the possibility that students' growth trajectories were curvilinear through the coding scheme for the shape factor (that is, 0, 1, *). The asterisk indicates a free parameter which is then estimated in fitting the model. The interval 0 to 1 represents the linear portion in the model describing the yearly change in reading achievement between year 2 and year 3 of the study. The last growth interval represents any nonlinear change that might be present over the latter part of the study. This coding strategy is also appropriate for handling the unequal spacing of measurement occasions (that is, for measures of distributed leadership and academic capacity). Coding the first interval 0.0 ensures that the level factor represents students' achievement levels at year 2 of the study (Raykov and Marcoulides, 2006, for further discussion). The leadership and academic capacity level and shape factors were defined in the same manner to incorporate possible nonlinear change in the latter portion of the study.
5. The structural part of the SEM analysis can then be used to investigate the effects of covariates or other latent variables on the latent change factors. We can model variability in initial reading level (η_{0i}) and shape (η_{1i}) latent variables as a function of one or more covariates (x_i) plus error:

$$\eta_{0i} = \alpha_0 + \gamma_0 x_i + \zeta_{0i}, \quad (2)$$

$$\eta_{1i} = \alpha_1 + \gamma_1 x_i + \zeta_{1i}, \quad (3)$$

where α_0 and α_1 are measurement intercepts and γ_0 and γ_1 are structural parameters describing the regressions of latent variables on a covariate. Each latent factor has its own residual (ζ_{0i}, ζ_{1i}) that permits the quality of measurement associated with each individual's growth trajectory to differ from those of other individuals. Once the overall achievement model has been defined through relating the observed variables to the latent factors that represent the change process (as in Eq. 1–3), it can be further divided into its respective individual-level measurement and structural models and its organizational-level measurement and structural models. In contrast to reading achievement and growth, the leadership and school capacity change processes are defined only at the school level.

6. Note that most of the within- and between-school variability in reading growth is due to initial achievement levels. For example, if initial school reading achievement were included in the model as a predictor of reading growth, the variance accounted for between schools would be nearly 88 percent.

References

Barth R (1990) *Improving Schools from Within*. San Francisco, CA: Jossey-Bass.

Barth R (2001) Teacher leader. *Phi Delta Kappan* 82(6): 443–449.

Bass B and Bass R (2008) *The Bass Handbook of Leadership*. New York: Free Press.

Bell L, Bolam R and Cubillo L (2003) *A Systematic Review of the Impact of School Headteachers and Principals on Student Outcomes*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Blalock H (1989) The real and unrealized contributions of quantitative sociology. *American Sociological Review* 54: 447–460.

Bloom HS, Hill CJ, Black AR and Lipsey MW (2008) Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness* 1: 289–328.

Bossert S, Dwyer D, Rowan B and Lee G (1982) The instructional management role of the principal. *Educational Administration Quarterly* 18(3): 34–64.

Bush T and Glover D (2003) *School Leadership: Concepts and Evidence*. Commissioned report. Nottingham: National College for School Leadership.

Cook T (2002) Randomized experiments in education. Why are they so rare? *Educational Evaluation and Policy Analysis* 24(3): 175–200.

Cook T and Campbell DT (1979) *Quasi-experimentation. Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin.

Crowther F, Ferguson M and Hann L (2008) *Developing Teacher Leaders*. Thousand Oaks, CA: Corwin Press.

Davies R (1994) From cross-sectional to longitudinal analysis. In: Davies R and Dale A (eds) *Analysing Social and Political Change: A Casebook of Methods*. Newbury Park, CA: SAGE, 20–40.

Donaldson GA (2001) *Cultivating Leadership in Schools: Connecting People, Purpose, and Practice*. New York: Teachers College Press.

Fullan M (2001) *Leading in a Culture of Change*. San Francisco, CA: Jossey-Bass.

Gronn P (2002) Distributed leadership as a unit of analysis. *Leadership Quarterly* 13: 423–451.

Grubb WN and Flessa J (209) 'A job too big for one': multiple principals and other non-traditional approaches to school leadership. In: Leithwood K, Mascall B and Strauss T (eds) *Distributed Leadership According to the Evidence*. London: Routledge, 137–164.

Hall G and Hord S (2001) *Implementing Change: Patterns, Principles, and Potholes*. Boston, MA: Allyn & Bacon.

Hallinger P (2003) Leading educational change: reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education* 33(3): 329–351.

Hallinger P, Bickman L and Davis K (1996) School context, principal leadership and student achievement. *Elementary School Journal* 96(5): 498–518.

Hallinger P and Heck RH (1996a) Reassessing the principal's role in school effectiveness: a review of the empirical research, 1980–1995. *Educational Administration Quarterly* 32(1): 5–44.

Hallinger P and Heck RH (1996b) The principal's role in school effectiveness: an assessment of methodological progress, 1980–1995. In: Leithwood K, Chapman J, Corson D, Hallinger P and Hart A (eds) *International H0andbook of Educational Leadership and Administration*. Dordrecht: Klewer Academic, 723–783.

Hallinger P and Heck RH (1998) Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement* 9(2): 157–191.

Hallinger P and Murphy J (1985) Assessing the instructional management behavior of principals. *The Elementary School Journal* 86(2): 217–248.

Hallinger P and Murphy J (1986) The social context of effective schools. *American Journal of Education* 94(3): 328–355.

Harris A (2003) Teacher leadership as distributed leadership: heresy, fantasy or possibility? *School Leadership and Management* 23(3): 313–324.

Heck RH (2009) Examining the impact of successive teachers on student learning: testing a multilevel cross-classified model. Paper presented at the annual meeting of the American Educational Research Association, San Diego, April.

Heck RH and Hallinger P (2005) The study of educational leadership and management: where does the field stand today? *Educational Management, Administration & Leadership* 33(2): 229–244.

Heck RH, Larson T and Marcoulides G (1990) Principal instructional leadership and school achievement: validation of a causal model. *Educational Administration Quarterly* 26: 94–125.

Hedges L (2008) What are effect sizes and why do we need them? US Department of Health and Human Services, Administration for Children & Families, Paper No.1. Available at: http://www.acf.hhs.gov/programs/opre/other_resrch/effect_sizes/pres_papers/hedges.html (accessed 6 December 2008).

Hill P and Rowe K (1996) Multilevel modeling in school effectiveness research. *School Effectiveness and School Improvement* 7: 1–34.

Jackson D (2000) The school improvement journey: perspectives on leadership. *School Leadership & Management* 20(1): 61–78.

Kelly J and McGrath J (1988) *On Time and Method*. Newbury Park, CA: SAGE.

Lambert L (2002) A framework for shared leadership. *Educational Leadership* 59(8): 37–40.

Langlois R and Robertson PL (1993) Business organisation as a coordination problem: toward a dynamic theory of the boundaries of the firm. *Business and Economic History* 22(1): 31–41.

Leithwood K and Jantzi D (1999) The relative effects of principal and teacher leadership on student engagement in school. *Educational Administration Quarterly* 35(Supplemental): 679–706.

Leithwood K and Montgomery D (1982) The role of the elementary principal in program improvement. *Review of Educational Research* 52(3): 309–339.

Leithwood K, Anderson S, Mascall B and Strauss T (in press) School leaders' influences on student learning: the four paths. In: Bush T, Bell L and Middlewood D (eds) *The Principles of Educational Leadership and Management*. London: SAGE.

Leithwood K, Day C, Sammons P, Harris A and Hopkins D (2006) *Seven Strong Claims About Successful School Leadership*. Nottingham: National College of School Leadership.

Leithwood K, Louis KS, Anderson S and Wahlsttom K (2004) Review of research: how leadership influences student learning. Wallace Foundation. Available aT: http://www.wallacefoundation.org/NR/rdonlyres/E3BCCFA5-A88B-45D3–8E27-B973732283C9/0/ReviewofResearchLearningFromLeadership.pdf (accessed 19 December 2007).

Leithwood K, Mascall B and Strauss T (2009) What we have learned where we go from here. In: Leithwood K, Mascall B and Strauss T (eds) *Distributed Leadership According to the Evidence*. New York: Routledge, 269–282.

Luyten H, Visscher A and Witziers B (2005) School effectiveness research: from a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement* 16(3): 249–279.

MacBeath J, Frost D and Swaffield S (2008) Editorial. *School Leadership & Management* 28(4): 301–306.

Marcoulides G and Hershberger S (1997) *Multivariate Statistical Methods: A First Course*. New York: Lawrence Erlbaum.

Marks H and Printy S (2003) Principal leadership and school performance: an integration of transformational and instructional leadership. *Educational Administration Quarterly* 39(3): 370–397.

Marsh HW and Cravens RG (2006) Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science* 1(2): 133–163.

Mulford W (2007) Building social capital in professional learning communities: importance, challenges and a way forward. In: Stoll L and Seashore Louis K (eds) *Professional Learning Communities: Divergence, Depth and Dilemmas*. London: Open University Press and McGraw Hill, 166–180.

Mulford B and Silins H (2003) Leadership for organizational learning and improved student outcomes—what do we know? *Cambridge Journal of Education* 33(2): 175–195.

Ogawa, R and Bossert S (1995) Leadership as an organizational quality. *Educational Administration Quarterly* 31(2): 224–243.

Peugh JL and Enders C K (2004) Using an EM covariance matrix to estimate structural equation models with missing data: choosing an adjusted sample size to improve the accuracy of inferences. *Structural Equation Modeling: A Multidisciplinary Journal* 11: 1–19.

Pitner N (1988) The study of administrator effects and effectiveness. In: Boyan N (ed.) *Handbook of Research in Educational Administration*. New York: Longman, 99–122.

Podsakoff P (1994) Quantitative methods in leadership research. *Leadership Quarterly* 5: 1–2.

Pounder DG, Ogawa RT and Adams EA (1995) Leadership as an organization-wide phenomena: its impact on school performance. *Educational Administration Quarterly* 31(4): 564–588.

Raykov T and Marcoulides G. A (2006) *A First Course in Structural Equation Modeling*, 2nd edn. Mahwah, NJ: Lawrence Erlbaum.

Robinson V, Lloyd C and Rowe K (2008) The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educational Administration Quarterly* 44(5): 564–588.

Southworth G (2002) Instructional leadership in schools: reflections and empirical evidence. School *Leadership and Management* 22(1): 73–91.

Southworth G (2003) *Primary School Leadership in Context: Leading Small Medium and Large Sized Schools*. London: Taylor and Francis.

Timperly H (2009) Distributing leadership to improve outcomes for students. In: Leithwood K, Mascall B and Strauss T (eds) *Distributed Leadership According to the Evidence*. London: Routledge, 197–222.

Wiley S (2001) Contextual effects on student achievement: school leadership and professional community. *Journal of Educational Change* 2(1): 1–33.

Williams L and Podsakoff P (1989) Longitudinal field methods for studying reciprocal relationships in organizational behavior research: toward improved causal analysis. *Research in Organizational Behavior* 11: 247–292.

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

Yukl G (2006) *Leadership in Organizations*. Paramus: Prentice Hall.

Biographical Notes

Philip Hallinger, EdD, is Chair Professor and Director of the *Asia Pacific Centre for Leadership and Change* at the Hong Kong Institute of Education. His research focuses on school leadership effects, leadership development and problem-based learning. He received his EdD in Administration and Policy analysis from Stanford University. His research publications can be accessed at philiphallinger.com

Ronald Heck, PhD, holds the Dai Ho Chun Endowed Chair in Education at the University of Hawaii-Manoa. He received his PhD in Education from the University of California at Santa Barbara. His research focuses on the relationship between school leadership and school improvement. He has also written on research methodology.