A Review of Three Decades of Doctoral Studies Using the
Principal Instructional Management Rating Scale:
A Lens on Methodological Progress in Educational Leadership

Philip Hallinger
Hong Kong Institute of Education
Hallinger@gmail.com
Abstract

This report continues in the lineage of reviews of research in educational leadership and management by examining methodological approaches used by doctoral researchers in studying principal instructional leadership. The paper reviews the full set of 126 doctoral dissertations completed over the past three decades that used the Principal Instructional Management Rating Scale (PIMRS). The report analyzes trends in the research foci, conceptual models, research designs, and statistical methods employed in these studies. The study finds that interest in instructional leadership among scholars and practitioners remained strong throughout the period of the review, the PIMRS has proven a reliable and valid data collection tool, and the use of research methodology has improved in several specific areas. Nonetheless, the results also suggest that the conceptual frameworks and methodologies used by these doctoral students were, on the whole, inadequate for the task of contributing to either the theoretical or practical knowledge base in this field. This impression of weak knowledge accumulation was further reinforced by a citation analysis that found limited citations of the dissertations by other researchers in the field at large, or by the dissertation authors themselves. These conclusions applied equally to Ed.D. and Ph.D. dissertations, regardless of the level of research university from which it was produced. The review adds empirical evidence of trends in research quality gathered over a 30 year period to the renewed debate over the purpose and direction of the doctoral dissertation in this professional field.
This report continues in the lineage of published reviews of research on educational leadership and management conducted by scholars over the past 50 years. Some of these efforts have examined research questions and conceptualizations (Bridges, 1982; Hallinger & Heck, 1996a; Heck & Hallinger, 2005; Lipham, 1964; Pitner, 1988). Others have focused on methodological issues (Bridges, 1982; Erickson, 1967; Haller, 1979; Hallinger & Heck, 1996b; Heck & Hallinger, 1999; Rowan et al., 1982). Some reviews have used critical synthesis to illuminate emergent themes across a broad range of studies (Bridges, 1982). Others have used synthesis (e.g., Bell, Bolam, & Cubillo, 2003; Bossert, Dwyer, Rowan & Lee, 1982; Hallinger & Heck, 1998; Leithwood & Montgomery, 1982; Leithwood, Begley & Cousins, 1990; Leithwood, Day, Sammons, Harris, & Hopkins, 2006; Lipham, 1964) or meta-analytic techniques (Robinson, Lloyd & Rowe, 2008; Waters, Marzano & McNulty, 2003) to assess knowledge accumulation in a particular sub-domain of study.

Fifteen years ago Hallinger and Heck (1996a) observed that instructional leadership had been the most prevalent perspective adopted by researchers engaged in the study of school leadership effects between 1980 and 1995. They further noted that the Principal Instructional management Rating Scale (PIMRS; Hallinger 1982, 1990) had been the most commonly used instrument used by researchers in studies of school leadership effects during that era. An updated perusal of the school leadership literature in 2008 found that the PIMRS instrument had been employed in more than 110 doctoral studies (author, 2008). Renewed global interest in the related constructs of ‘instructional leadership’ (Gewirtz, 2003; Hallinger, 2003; Stricherz, 2001) and ‘leadership for learning’ (Leithwood et al., 2006; MacBeath & Cheng, 2008) suggest the timeliness for a review that examines the methodological approaches that scholars are using to conduct research in this important domain.

This review is limited to doctoral studies of instructional leadership conducted between 1983 and 2010 that used the PIMRS instrument. The review seeks to describe and analyze
approaches to research used in these studies rather than provide a synthesis of substantive findings. In an attempt to assess and portray methodological progress in the field over time, this report addresses four specific questions.

1. How has interest in the empirical study of instructional leadership varied over the three decades covered by the review?
2. What types of school leaders do scholars study?
3. How are doctoral student researchers using the PIMRS instrument to study instructional leadership?
4. What is the focus of research on instructional leadership?

This approach to the review of research offers several potential contributions. The examination of a large number of studies that used a common instrument over a long period of time provides a unique opportunity to investigate patterns of change in the conduct of empirical research in this field. This may yield useful recommendations concerning high impact foci for future research as well as desirable conceptual models and research methods. Since the review consists of studies conducted by doctoral students across a large number of universities our lens may also illuminate changing patterns of practice in graduate research training.

**Instructional Leadership as a Domain of Educational Leadership and Management**

Among the global trends in educational leadership and management that emerged over the past 30 years, few have been more significant, widespread or persistent than the focus on understanding linkages between leadership and learning (Bell et al., 2003; Bossert et al., 1982; MacBeath & Cheng, 2008; Hallinger & Heck, 1998; Leithwood, Anderson, Mascall & Strauss, in press; Robinson et al., 2008; Southworth, 2002; Witziers et al., 2003). While early interest in this issue can be traced to the conceptual work of Bridges (1967) and the empirical research of Gross and Herriott (1965) in the USA, it was not until the early 1980s that scholars began to
explore this relationship in a sustained fashion (Hallinger & Heck, 1996a). It was during this era when findings from research on school effectiveness and school improvement emerged to reinforce a strongly held belief among policymakers and practitioners that principal leadership ‘makes a difference’ in school performance (Edmonds, 1979; Purkey & Smith, 1983). More specifically, these bodies of research identified principal instructional leadership as a key factor in instructionally effective schools (Bossert et al., 1982; Leithwood & Montgomery, 1982).

Yet, scholars were also quick to note specific limitations in the empirical investigation of the principal’s role as an instructional leader (Bossert et al., 1982; Cuban, 1984; Leithwood & Montgomery, 1982; Rowan, Dwyer & Bossert, 1982). These included:

- Lack of clearly explicated conceptual frameworks for studying relevant constructs;
- Lack of valid and reliable instrumentation for studying the role;
- Lack of theoretical models that articulated how this role influenced student learning;
- Reliance on weak research designs, ill-equipped to test for causal effects. (Bossert et al., 1982; Bridges, 1982; Cuban, 1984)

These limitations were cause for concern, especially in light of burgeoning attempts to embed the research findings into government policies and principal training curricula (Barth, 1986; Cuban, 1984).

This emergent focus on instructional leadership subsequently led to the development of new conceptual frameworks and instruments (e.g., Bamburg & Andrews, 1990; Hallinger, 1982; van de Grift, 1990; Villanova, Gauthier, Proctor & Shoemaker, 1981). Over the ensuing decades scholars used these research tools to generate a substantial body of research that has been the subject of analytical reviews conducted by scholars in several countries (Bell et al., 2003; Hallinger, 2003; Hallinger & Heck, 1996a, 1996b, 1998; Leithwood et al., in press, 1990, 2006, in press; Robinson et al., 2008; Witziers et al., 2003). Yet, with but one exception (Hallinger &
Heck, 1996b), these reviews have focused on the synthesis of findings rather than on the analysis of methodological approaches used in this body of empirical research.

In 1980 when instructional leadership emerged as a new construct, many questioned both its relevance and viability as a guiding metaphor for school leadership. Thirty years later, ‘instructional leadership’ and ‘leadership for learning’ are widely accepted by policymakers and practitioners as an essential focus for management practice in schools. Indeed, recent reviews of research largely confirm the early assertions concerning the relationship between instructional leadership and student learning (see Leithwood et al., in press; Robinson et al., 2008). Contrary to early predictions, instructional leadership has demonstrated impressive staying power as a core concept guiding both practice in the field of educational leadership and management.

**Instructional Leadership Framework**

Thirty years ago, in their seminal review of the literature, Bossert and colleagues (1982) defined the construct of *instructional management*. They selected the term instructional management because they inferred that this role of the principal revolved around managerial functions concerned with the coordination and control of curriculum and instruction (see also Cohen & Miller, 1980). Their instructional management framework became an influential model that, to this day, continues to guide researchers in this field.

While Bossert and his colleagues coined the term *instructional management*, over time *instructional leadership* came to be accepted as the term commonly used by scholars and practitioners. In our view, the formal distinction between these conceptual terms lies in the sources of power and means proposed to achieve results. Instructional *leadership* become the preferred term due to the recognition that principals who operate from this frame of reference rely more on expertise and influence than on formal authority and power to achieve a positive and lasting impact on staff motivation and behavior and student learning.
The Principal Instructional Management rating Scale (Hallinger, 1982, 1990) is grounded in a conceptual framework that proposes three dimensions in this role: Defining the School’s Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate (Hallinger & Murphy, 1985; see Figure 1). These dimensions are delineated into 10 instructional leadership functions. Two functions, Framing the School’s Goals and Communicating the School's Goals, comprise the dimension, Defining the School’s Mission. These functions concern the principal’s role in working with staff to ensure that the school has a clear mission and that the mission is focused on academic progress of its students. While this dimension does not assume that the principal defines the school’s mission alone, it does propose that the principal is responsible for ensuring that such a mission exists and for communicating it widely to staff. This dimension is the starting point for creating a learner-centered school.

The second dimension is Managing the Instructional Program. This incorporates three leadership functions: Supervising and Evaluating Instruction, Coordinating the Curriculum, Monitoring Student Progress. This dimension focuses on the role of the principal in “managing the technical core” of the school. In larger schools, it is clear that the principal is not the only person involved in monitoring and developing the school’s instructional program. Yet this framework assumes that coordination and control of the academic program of the school is a key leadership responsibility of the principal.

The third dimension, Promoting a Positive School Learning Climate includes several functions: Protecting Instructional Time, Promoting Professional Development, Maintaining High Visibility, Providing Incentives for Teachers, and Providing Incentives for Learning. This dimension is broader in scope and intent than the second dimension and overlaps with dimensions incorporated into transformational leadership frameworks (Hallinger, 2003;
The final, definitive version of this paper has been published in Educational Administration Quarterly, 47(2), April 2011 by SAGE Publications Ltd, All rights reserved. © The University Council for Educational Administration 2011. The article is published: http://online.sagepub.com

Leithwood, 1994). It conforms to the notion that successful schools create an “academic press” through the development of high standards and expectations and a culture that fosters and rewards continuous learning and improvement.

The PIMRS Instrument

The original form of the PIMRS (Hallinger, 1982) contained 11 subscales and 72 “behaviorally anchored” items (See Hallinger (1982) and Latham and Wexley (1981) for discussions of behaviorally anchored rating scales and their development). Subsequent revision of the instrument reduced the instrument to 10 subscales and 50 items (Hallinger, 1983). For each item, the rater assesses the frequency with which the principal enacts a behavior or practice associated with that particular instructional leadership function. Each item is rated on a Likert-type scale ranging from (1) almost never to (5) almost always (see Figure 1). The instrument is scored by calculating the mean for the items that comprise each subscale. This results in a profile that yields data on perceptions of principal performance on each of the 10 instructional leadership functions.

Three parallel forms of the instrument have been developed and tested: a self-assessment form to be completed by the principal, a teacher form and a supervisor form. The items which comprise each form are identical; only the stems change to reflect the differing perspectives of the role groups. Early studies found significant differences in perceptions across role groups (Hallinger & Murphy, 1985; Krug, 1986; O’Day, 1984). Validation studies in the United States indicate that the PIMRS form that solicits teachers’ perceptions provides the most valid data of the three forms.

Insert Figure 2 about here

The original validation study found that the PIMRS met high standards of reliability (Hallinger, 1983). All ten subscales exceeded .80 using Cronbach’s test of internal consistency.
Subsequent studies have generally substituted Ebel’s (1951) test for calculating inter-rater reliability for Cronbach’s formula. This test provides a more accurate test of reliability for ratings aggregated from a set of schools where respondents within schools (e.g., teachers) are rating a feature of the school (i.e., the principal). These studies have supported the original validation study in its conclusion that the scale provides reliable data on instructional management (e.g., Dunn, 2010; Fulton, 2009; Harris, 2002; Howe, 1995; Jones, 1987; Leitner, 1990; Mercer, 2004; Moore, 2003; O’Day, 1984; Taraseina, 1993). Studies have further tested the PIMRS for face validity, content validity and discriminant validity. Initially, the instrument was judged to be a valid measurement tool for use at the elementary school level. Subsequent studies referred to in this paper expanded on the instrument’s validation (e.g., Howe, 1995; Jones, 1987; Leitner, 1990; O’Day, 1984; Taraseina, 1993).

It is not the purpose of this report to systematically assess the PIMRS as a tool for research. Given the purposes of this review of research, it suffices to conclude that the instrument appears to have provided a consistently reliable and reasonably valid means of assessing the instructional leadership of school principals. Given earlier critiques that centered on the lack of reliable instrumentation for use in the field of educational leadership and management, this finding lays an important foundation for the subsequent review.

Methodology of the Review

Following in the tradition of other published reviews of research (e.g., Bell et al., 2003; Bridges, 1982; Erickson, 1967; Haller, 1979; Hallinger & Heck, 1996a; Leithwood et al., 1990, 2006; Lipham 1964), this effort used critical synthesis to illuminate and understand patterns in the characteristics of the constituent studies. As noted, however, this report focuses on research questions and methodological approaches rather than findings. A critical assessment of methodological trends offers a necessary complement to reviews that focus on substantive
The final, definitive version of this paper has been published in Educational Administration Quarterly, 47(2), April 2011 by SAGE Publications Ltd, All rights reserved. © The University Council for Educational Administration 2011. The article is published: http://online.sagepub.com

results (Bridges, 1982; Erickson, 1967; Haller, 1979; Hallinger & Heck, 1996b; Rowan et al., 1982).

The Sample

The author searched Dissertation Abstracts for studies that used the PIMRS between 1983 and 2010 using a variety of keywords: instructional leadership, PIMRS, principal leadership, instructional management, Hallinger. The author obtained access to an additional seven dissertations that did not show up in the search from authors via a search on Google Scholar. Some had been completed in foreign countries or at North American universities that did not provide dissertations to UMI. The full dataset totaled 126 studies. The distribution of projects across degree obtained by the researchers included 88 Ed.D., and 37 Ph.D. dissertations. Most of the dissertations were downloaded in pdf format for analysis. In cases where the full dissertation file was not available, data were gleaned as far as possible from the abstract. Where this was insufficient, this is noted in the text.

Analysis of the Studies

A team of researchers worked intermittently over a period of years to read the studies, and then to generate and collate the data. The following categories were used to guide the extraction of information from the studies.

1. Job title of the role group(s) studied (e.g., principal, superintendent, vice principal);
2. Institutional setting in which the administrator was employed (e.g., public/private, elementary, middle, secondary);
3. Sample size by respondent role group (i.e., principal, teacher, supervisor);
4. Research questions and hypotheses (i.e., research focus and questions);
5. Major constructs or variables included for study (i.e., independent and dependent variables);
6. Conceptual model guiding the research (e.g., direct effects, mediated effects, reciprocal effects);
7. Research design (e.g., survey, experiment, case study);
8. Instrumentation (i.e., name of other scales used in concert with the PIMRS);
9. Data analysis approach and tests (i.e., description, single causal factor, single causal factor with controls, multiple factor);
10. Main findings and significance;
11. University sponsoring the research (i.e., name);
12. University Type (i.e., level of research activity by Carnegie categories);
13. Degree type (i.e., MA, Ed.S., Ed.D., Ph.D.);

These data were extracted, compiled, and constructed into a master table that displayed the characteristics of each of the individual studies. Data from this table were then employed for the subsequent analyses of trends across studies, within groups of studies, and over time.

It should be noted that our approach to analyzing methodological progress in the field was designed to build explicitly on the methods and findings of prior reviews. For example, we followed a ‘template’ employed in Bridges (1982) review in the selection and definition of several of the data categories mentioned above. Similarly, the criteria used to categorize the conceptual models that framed the studies were adapted from prior reviews conducted by Pitner (1988) and Hallinger and Heck (1996a). The selection of comparable categories for analysis enabled the author to more explicitly link the findings of this review to trends reported in reviews conducted at earlier points in time. We suggest that this approach ‘magnified’ the value of this particular review effort by leveraging knowledge gained from prior reviews.

Limitations of the Study

The main limitation of this report lies in acknowledging that the review is limited to doctoral dissertations and does not include published studies. An online search using Google Scholar found that only six of the dissertations resulted in journal publications (Hallinger & Murphy, 1983; Hallinger, Taraseina & Miller, 1994; Leitner, 1994; Nogay, & Beebe, 1997; O’Donnell, & White, 2005; Sheppard, 1993). This suggests that one might expect a greater
degree of variation in quality within this dataset than in a sample comprised wholly of journal articles (e.g., Robinson et al., 2008), or journal articles plus dissertations (e.g., Bridges, 1982; Erickson, 1967; Hallinger & Heck, 1996a).

The decision to focus solely upon doctoral dissertations was based on the large size of the dataset and the belief that this delimitation would enable a clearer identification of methodological patterns. That is, we would not be mixing apples (i.e., dissertations) with oranges (published research reports). Other scholars who followed a similar strategy were able to provide a useful picture of research development in prior reviews (e.g., Haller, 1979). Moreover, it is an empirical fact that doctoral dissertations represent the bulk of research conducted in the field of educational administration (Agusto, 2009; Archbald, 2008; Bridges, 1982; Haller, 1979; Hoffer, Welch, Williams, Hess, Webber & Lisek, 2005). We address this limitation by carefully considering the extent to which findings from doctoral studies might reflect research trends in the field as a whole.

A second limitation lies in the decision to focus solely on studies that used the PIMRS. While this selection criterion enabled a clearly defined set of instructional leadership studies, it resulted in an incomplete, description of the ways in which researchers have studied instructional leadership. For example, the exclusive focus on PIMRS studies will have resulted in under-reporting the incidence of qualitative studies of instructional leadership. We acknowledge this limitation and interpret the results accordingly.

Results

The results are organized around the four main questions posed at the outset of the review.
How has Researchers’ Interest in Instructional Leadership Varied Over Time?

The first issue concerns the level of scholarly interest in instructional leadership over the past three decades. The data in Table 1 indicate consistent and sustained interest in the study of instructional leadership throughout the past three decades. Despite skepticism about the sustainability of instructional leadership as a guiding construct for the school principalship (e.g., Cuban, 1984), these data portray a remarkable consistency of interest in the construct on the part of scholars. Since educational administrators form the bulk of the student population in doctoral programs in educational leadership and management, one could reasonably conclude that this trend also applies to school practitioners.

We further observe that although interest in this construct is centered in the USA, PIMRS doctoral studies were conducted at 83 universities in at least nine different countries (not tabled). These include the United States (116), Canada (2), Thailand (3), Taiwan (2), Cameroon (1), Guam (1), and the Philippines (1). This probably under-reports the number of countries since doctoral studies conducted in foreign countries often do not find their way into the UMI Dissertation Abstracts. Although they were not included in the substantive analyses conducted for this report, we note that all of the Master theses collected in this effort were conducted outside of the USA: Philippines (3), Canada (1), UK (1), Hong Kong (1).

Who is studied?

A continuing question of interest concerns the role groups that scholars choose to study in the field of educational leadership and management. Earlier reviews consistently reported a preference among scholars for studying school administrators at the elementary school level (Erickson, 1967; Bossert et al., 1982; Bridges, 1982; Hallinger & Heck, 1996a). The data in Table 1 show a continuing preference for studying principals (45% of the total studies) at the
elementary school level. When the data were sorted by both year of the study and level of the school, we found that the level of interest in studying instructional leadership at different institutional levels has not changed in the past 30 years. This is a potentially important trend since there is a need to better understand how instructional leadership is enacted in secondary schools which tend to be larger and more complex than elementary schools. We also note that two studies examined the instructional leadership of school superintendents (Edwards, 2006; Kroeze, 1992).

How is Instructional Leadership Studied?

Our approach to understanding how researchers have used the PIMRS to study instructional leadership incorporates an analysis of several methodological dimensions that combine to impact the quality of research outcomes: conceptual models, research designs, statistical methods.

Conceptual models. For the purposes of this review, the studies were classified into four basic models and associated variants: antecedent-effects, direct-effects, mediated-effects, reciprocal-effects (see Figure 3). These models describe the study’s conceptualization, whether implicit or explicit, of the relationship between instructional leadership and other variables. Nine studies were classified as NA (not applicable) for cases where this conceptual framework for our own analysis could not be applied.

Researchers who employed Model A, an antecedent-effects model, explored the effects of different variables on the exercise of instructional leadership. If a study examined the
relationship between demographic (e.g., age, experience, gender) or other personal characteristics (e.g., self-efficacy, years of teaching experience, knowledge of instruction) of the principal and perceptions of instructional leadership, it was classified as Model A-1. Studies that explored the relationship of school context factors (e.g., school level, school size, urban/rural, SES level) to the principal’s instructional leadership were classified A-2. Note that role set studies (Merton, 1957) were also included in the A-2 category. Role set studies compare the perceptions of the principal’s instructional leadership through the eyes of two or more groups of respondents holding different formal roles in the organization (e.g., principals, teacher, supervisors). Studies that incorporated a combination of both personal and context variables as independent variables studied in relation to instructional leadership were classified A-3. As displayed in Table 2, researchers used antecedent-effects models in almost 50% of the studies.

Direct-effects studies (see Model B in Figure 3) frame instructional leadership as an independent variable and explore its relationship to a dependent variable. This is usually a measure of a school-level condition (e.g., school climate, school mission, teacher collective efficacy, teacher satisfaction) or a school outcome (e.g., student achievement, school effectiveness). When antecedent variables were also included as controls, the studies were classified as B-1. When antecedent variables were not included in the model, the studies were classified as B-2. A total of 45 or about 36% of the studies were classified as Model B: 26 as B-1 and 19 as Model B-2.

The next category, Model C, consists of mediated-effects studies. In this model, the researcher seeks to understand the avenues or paths through which instructional leadership influences distal dependent variables such as student achievement, teacher satisfaction, or school effectiveness (Bossert et al., 1982; Hallinger & Heck, 1996a; Pitner, 1988). Model C suggests that the effects of leadership are mediated by other conditions in the school (e.g., school culture or climate, mission, structure, teacher involvement in decision-making, team
alignment) and has been recommended as a preferred approach to studying leadership effects (Hallinger & Heck, 1996a, 1996b). Relatively few (12) studies employed Model C as a framework for analysis; nine included antecedents (Model C-1) and three did not (Model C-2).

Model D portrays a reciprocal-effects model of leadership effects. When employing this approach, the researcher seeks to understand the mutual influence of leadership and related conditions in the school. This is a methodologically challenging approach to the study of leadership effects and the incidence of published studies that have employed a reciprocal-effects model, either inside or outside of education, is very low (Hallinger & Heck, 1996b, 2010). Therefore, it was not surprising to find that none of the studies employed Model D.

This analysis suggests that doctoral researchers have continued to rely heavily upon relatively weak two-factor conceptual models, a trend that was strongly critiqued in prior reviews (e.g., Bridges, 1982; Haller, 1979; Hallinger & Heck, 1996b). We wondered however, given the extended length of the current review, if there might have been change in the pattern of use of these conceptual models over time. Table 3 shows the frequency of use of the conceptual models across three nine-year periods. We wish to highlight two trends in the data:

- Use of Models A and B was consistently strong throughout the 27 year period of the review, thereby continuing a pattern noted in other reviews going back to the 1960s. Unfortunately, studies employing Model A have tended to do so in an atheoretical fashion. For example, with few exceptions (e.g., Cunningham, 2004), Model A-1 studies that examined the relationship between principal gender and instructional leadership did so without testing a theoretical explanation for why gender differences would impact leadership. They simply employed a bivariate test for differences between groups of male and female principals. Therefore,
after three decades and 23 PIMRS studies of principal gender, we know that female principals consistently receive higher ratings on the PIMRS than males, but are no closer to understanding why.

- Model C represents a more sophisticated approach to understanding leadership effects in the school organization and has been recommended by reviewer as a model of choice for 30 years (Bossert et al., 1982; Bridges, 1982; Hallinger & Heck, 1996b; Leithwood et al., 2004, 2006; Pitner, 1988). We noted that this model was used primarily by doctoral researchers during the final period of the review. While this trend is encouraging, the relatively small total number of Model C studies does not qualify this as evidence of major progress.

Since leadership is enacted in complex organizational settings, we suggest that researchers employ models that are able to portray systemic, multivariate rather relationships. Model A studies focused on factors that influence the exercise of leadership, but too many studies have failed to link that relationship to the impact of leadership. Model B studies focused on leadership effects, but did not employ sufficiently comprehensive perspectives. Thus, we join other reviewers in recommending that researchers incorporate antecedents into Model C studies (i.e., C-1). This offers the possibility of contextualizing the enactment of leadership and its effects on school-level conditions and school outcomes.

**Research design.** In 1979, Haller conducted an analysis of research methods used in doctoral dissertations from UCEA institutions in the United States and concluded that questionnaires used in combination with cross-sectional research designs predominated. In 1982 Bridges examined the use of research methods in both doctoral dissertations and published research in the field’s most selective research journals and concluded:

> Although researchers apparently show a greater interest in outcomes than was the case in the earlier period, they
continue their excessive reliance on survey designs, questionnaires of dubious reliability and validity, and relatively simplistic types of statistical analysis. Moreover these researchers persist in treating research problems in an ad hoc rather than a programmatic fashion. . . . Likewise the research seemed to have little or no practical utility. (pp. 24-25)

He went on to report that that the survey research design was used in over 90% of the 322 studies that he reviewed. Like Haller (1979) and Bridges (1982) we found the most commonly used research design was a cross-sectional, post-hoc survey. While the PIMRS was used to generate data in several case studies, there were no instances of longitudinal, experimental or quasi-experimental studies in the dataset.

Statistical approaches. Although the conceptual models described earlier lend themselves to a variety of different statistical analyses, past reviews have noted a preference among doctoral students for using a limited set of relatively weak data analytic techniques. For example, Haller noted: “Not only are questionnaires the major source of data, these data are typically generated within a cross-sectional research design and analyzed using simple descriptive or bivariate procedures (1979, p. 48).

Bridges (1982) conducted a more fine-grained analysis of data analysis techniques used by scholars in the field using a scheme comprised of four levels that describe statistical approaches in terms of their analytical power.

1. Description which involves the use of numbers to represent central tendencies and/or variability of scores.
2. Single causal factor/correlational which entails the examination of the relationship or association between two variables, one of which presumably covaries with or influences the other.
3. Single causal factor/correlational with controls which entails the examination of the relationship between two variables while controlling for the influence of one or more other variables.
4. Multiple factor which involves probing the differential effects of multiple sources of influence on a particular variable. (Bridges, 1982, p. 16)
In light of advances in statistical methods since the time of the Bridges review, we added a fifth level that we termed **Advanced Modeling**. We defined this level as comprised of tests that are capable of exploring relationships among multiple independent and dependent variables in a manner that allows for the examination of moderating and/or mediating effects (e.g., through the use of statistical methods such as confirmatory factor analysis, HLM, and structural equation modeling).

Using this scheme, Bridges concluded: “The bulk of the research on school administrators uses either description (60%) or a single factor/correlational without control approach (25%) in data analysis. Those approaches that enable the investigator to render rival explanations implausible are used in less than 16% of the studies” (1982, p. 16). These findings mirrored those of Haller who separately reported that doctoral students in educational administration were “much less likely to apply multivariate analyses to their questionnaire data. . . than were students from other disciplines” (1979, p. 48). We note that the importance of this issue is based on the centrality of shedding empirical light both on causal relationships between instructional leadership and learning and factors that shape the exercise of that leadership.

Have the patterns of statistical use by researchers in the field changed over the past three decades? The data in Table 4 suggest some improvement in this domain both in the overall pattern, as well as in the trend of use over time. For example, only about 10% of the PIMRS studies relied solely upon descriptive statistics and 36% used data analytic techniques ‘capable of ruling out plausible rival explanations.’ Moreover, Table 4 shows that in the most recent period (2001-2010), almost 50% of the studies used statistical techniques from levels 3 and 4 compared with only 28% during the initial period (1983-1991). In particular, the percentage of
studies using Multiple Factor tests rose in this period, climbing to 40% from 20% in the first period.

In order to place these findings in perspective, we need to recall that the study was limited to doctoral dissertations. Moreover, two-thirds of our sample was comprised of Ed.D. dissertations. One would not necessarily expect practitioners studying in Ed.D. programs to employ state-of-the-art research tools or push the outside of the conceptual envelope. Ed.D. programs are intended to be practice-oriented, and therefore contain fewer courses in research methods than are found in PhD. programs (Archbald, 2008; Agusto, 2009). With this in mind, one might expect to find a higher percentage of Ph.D. dissertations using Model C (mediated effects) frameworks in concert with Level 3, 4 and 5 statistical methods.

We examined whether the predominance of Ed.D. dissertations in the sample could have skewed the results towards the use of less comprehensive conceptual models and simpler statistical methods. Table 5 displays the breakdown of data by Type of Degree Program (i.e., Ed.D. or Ph.D.), Type of Conceptual Model (i.e., Model A, B, or C) and Level of Statistical Method (i.e., Level 1-5). Statistical analysis of these data revealed neither a significant correlation between the Degree Program and Conceptual Model used to frame the research, nor between the Degree Program and the Level of Statistical Method. This rather surprisingly suggests that the doctoral studies can be treated as a single group since the Ed.D. and Ph.D. dissertations were neither differentiated by their use of conceptual models nor by statistical methods.
Although the degree program did not provide leverage in explaining the pattern of results, it seemed plausible that the approaches to research adopted by the doctoral students could have been related to the type of university at which they were studying. With this question in mind, we categorized the universities according to a coding scheme proposed by the Carnegie Foundation (2010). This classifies universities into three categories based on the level of research activity of the university’s faculty:

- Type 1: RU/VH: Research Universities (very high research activity)
- Type 2: RU/H: Research Universities (high research activity)
- Type 3: DRU: Doctoral/Research Universities

One might expect that students studying at Type 1 universities would be exposed to more substantial funded research projects and activities. We coded the dissertations according to the Carnegie Foundation’s list and found that 58 of the studies had been completed at Type 1 universities, 34 at Type 2 universities, and 36 at Type 3 universities. We examined the Conceptual Models and Statistical Methods employed by the researchers in terms of the Type of University they had attended but found no statistically significant differences for either set of variables. Finally, we note that there was also no significant interaction effect between Degree Program and Type of University with either Conceptual Model or Level of Statistical Method. Thus, it was no more likely that students studying for a Ph.D. at a Type 1 research intensive university would conduct studies that employed more comprehensive models and powerful statistical approaches than Ph.D. or Ed.D. students at Type 2 or Type 3 universities.

In sum, while there were some demonstrable improvements in patterns of use of statistical methods, this mostly entailed a reduced reliance on the use of descriptive statistics and an increased use of bivariate tests without controls. Improvements did not extend to the widespread use of more powerful statistical methods able to shed light on important relationships concerned with leadership and learning. Finally, there was little “within-group
variation” across the studies when these patterns were further analyzed. Ed.D. and Ph.D. dissertations seemed remarkably similar in their methodological characteristics and it did not seem to matter whether they had been conducted at a research-intensive university or not.

What is the Focus of Research in Studies of Instructional Leadership?

The next issue concerned the focal questions and problems addressed in the research studies. We first group the studies in terms of antecedents, role set, and leadership effects in order to identify the variability in research issues addressed by the studies. Then two approaches to citation analysis are employed as a means of testing the extent to which this body of empirical research conducted over a thirty year period contributed to knowledge accumulation in this domain of educational leadership and management.

**Studies of antecedents effects on instructional leadership.** These studies followed Model A as discussed earlier. As noted by previous reviewers, the most frequent approach has been to study how different personal characteristics of the principals influence their instructional leadership. The most popular variables for study have been:


- **Years of administrative experience** (20%: Adkins, 1990; Brown, R., 1991; Cantu, 1994; Coltharp, 1989; Groff, 2002; Hallinger, 1983; Howell, 1989; Keith, 1989; Kincaid, 2006; Knezeck, 2001; Marshall, 2005; McCrier, 2004; McCTaggart, 1991; Meeks, 1999, Parker, 1990; Poovatanakul, 1993; Pratley, 1992; Rogers,
• **Years of teaching experience** prior to becoming a principal or other measures designed to act as proxies for knowledge of instruction (5%: Burwell, 1988; Delano, 1985; Kincaid, 2006; Lehl, 1989; McCrier, 2004; Simpson, 1992; Sterrett, 2005).

Other personal variables studied with some frequency included the principal’s preparation for the principalship, age, ethnicity, and self-efficacy.

A variety of school context variables have also been studied in terms of their effects on the principal’s instructional leadership. These include:

- **School level of the principal** (7%: Duryea, 1998; Gallon, 1998; Garcia, 1998; Hart, 2006; Howell, 1989; Kincaid, 2006; Rose, 1991; Sheppard, 1993; Singleton, 2006; );

- **School size** (13%: Anderson, 2006; Campbell, 1998; Duryea, 1998; Gallon, 1998; Garcia, 1998; Griffin, 1993; Haack, 1991; Howell, 1989; Macneil, 1992; Meeks, 1999; Pratley, 1992; Rogers, 2005; Schoch, 1992; Sheppard, 1993; Sterrett, 2005; van Pelt, 1993; Yang, 1996);

- **Private schools** (3%: Chi, 1997; Griffin, 1993; Hart, 2006; Howe, 1995).

Other context factors studied included the effects of a school-based management context, different district contexts (e.g., urban and rural), and district size.

These results offer an interesting elaboration on the earlier discussion of Model A studies by indicating the specific antecedents of instructional leadership that were of interest to scholars. The concentration of studies around a relatively limited set of personal and contextual independent variables is a positive finding; selection of topics around a concentration of antecedent variables contrasts with Bridges’ observation that topics appeared be ‘intellectual
random events’. At the same time, however, our assessment of the models used to study these topics leads us to reinforce Bridges’ assertion made thirty years ago:

Studies that merely describe the traits or attitudes should be discontinued unless they shed light on a problem of practical, social or theoretical significance. . . . [D]emographic traits such as sex also warrant investigation if they are treated in a theoretically rich fashion. . . . This research should also employ multivariate approaches to data analysis, approaches that allow the investigator to rule out plausible rival explanations and to estimate the relative influence of the traits or sentiments being studied. (1982, p. 26)

Role set studies. We classified this type of study as Model A-2 in our earlier analysis of conceptual models. Studies that compare the perceptions of different role groups (i.e., teachers, principals, supervisors) towards the instructional leadership behavior of the principal have proven to be a popular approach among this set of dissertations. Fully 16% of the studies included this type of research objective and related analyses (e.g., Brown, J., 1991; Chi, 1997; Duryea, 1988; Haack, 1991; Haasl, 1989; Hallinger, 1983; Henderson, 2007; Krug, 1986; Marshall, 2005; O’Day, 1984; Poovatanakul, 1993; Ratchaneeladdajit, 1997; Rogers, 2005; Smith, 2008; Stevens, 1996; Taraseina, 1993; Vinson, 1997; Wotany, 1999; Yang, 1996; Yogere, 1996). In some studies, role group comparison was an explicit objective of the study, while in others it was included as a means to understanding a larger issue (e.g., confirmation of the instrument’s validity).

As noted several times, it is not an objective in this review to discuss specific findings of the studies. Nonetheless, we would be remiss if we did not comment further on this particular specie of study since it lies at the intersection of research focus and methodology. As reported by the scale’s author (Hallinger & Murphy, 1985), the original validation study included role group comparison among teachers, principals and supervisors as a step in the validation process. It was essential to determine if and to what extent perceptions varied across the role groups and
to link those perceptions of behavior to alternative measures of those same behaviors (Hallinger & Murphy, 1985).

Subsequently, a number of studies have confirmed the original study’s finding of statistically significant differences across role groups in their perceptions of the principal’s instructional leadership (e.g., Brown, J., 1991; Haack, 1991; Haasl, 1989; Henderson, 2007; Krug, 1986; O’Day, 1984; Smith, 2008; Vinson, 1997). It is interesting to note that this result was also found in studies conducted in non-Western cultures such as Taiwan (Chi, 1997; Yang, 1996), Thailand (Poovatanakul, 1993; Ratchaneeladdajit, 1997; Taraseina, 1993) and the Philippines (Yogere, 1996). With few exceptions (e.g., Duryea, 1988; Krug, 1986; Marshall, 2005; Rogers, 2005), principal self-reports yield higher ratings than reports from their teachers. In each instance where additional validation procedures have been carried out (e.g., Hallinger, 1983; Taraseina, 1993), the teacher results have been found to most closely match independent sources of evidence. These findings suggest that teacher perceptions continue to constitute the preferred source of data on the principal’s instructional leadership for both research and evaluation purposes. Thus, we suggest that the analysis of role set perceptions does not, by itself, constitute a suitable focus for further research unless there is a larger question that the researcher is seeking to address.

**Leadership effects studies.** As suggested earlier, leadership effects studies could followed either Model B which posited direct effects of instructional leadership on some facet of the school, or Model C in which instructional leadership was linked to a distal dependent variable(s) through a mediating process. Topics studied with some frequency included the following.

- *Teacher stress, morale, collective efficacy, effectiveness and satisfaction* (8%: Courtney, 1987; Fancera, 2009; Keith, 1989; Kennedy, 1993; Lubbers, 1996; MacNeil, 1992; Reid, 1987; Ruzicska, 1988; Shatzer, 2009; Watkins, 1992),
• School climate (8%: Adkins, 1990; Campbell, 1999; Howe, 1989; Leitner, 1990; Lord, 2001; Reid, 1987; Sheppard, 1993; Simpson, 1990; Skiptunas, 1990; Wilson, 2005),


• Principal and school effectiveness (17%: Cantu, 1994; Coltharp, 1989; Dickerson, 1999; Fulton, 2009; Garcia, 1999; Gerrell, 2005; Gibson, 2005; Grier, 1988; Harris, 2002; Hunter, 1994; Johnson, 2005; Knezek, 2001; Lubbers, 1996; Maciel, 2005; Moore, 2003; Orange, 1990; Rose, 1991; Sinha, 2008; Singleton, 2006; Stroud, 1989; Wilson, 2005; Zeanah, 1986).

These data suggest strong interest in studying the relationship between instructional leadership and school effectiveness and student achievement. The studies that examined principal and school effectiveness actually represent a quite diverse group. As noted most used design in which principals were rated on their instructional leadership in a small sample of high and low performing schools. These profiles were then compared and sometimes tested for differences. While the choice of this research focus reflects an interest in understanding the contribution of instructional leadership to learning, this is a rather weak research design. While we recognize that this is an intuitively appealing approach, in the absence of a substantial sample size the research design offers only limited leverage on the problem of understanding the relationship between leadership and learning in these different school contexts.

Given its relevance to educational policy and its focus as a topic in quite a few high profile research reviews, we examined the trend in studies that examined the relationship of instructional leadership with student achievement more closely. Given critiques and recommendations of earlier reviewers (i.e., Bossert et al., 1982; Bridges, 1982; Hallinger & Heck, 1996a), we hoped that this dataset might offer insights into progress on this timely issue.
Our analysis, based partially on data shown in Table 6, suggests the following trends:

- Only a relatively small percentage of studies focused on studying the relationship between instructional leadership and student learning.

- At the same time, however, there has been increasing interest in studying this relationship in recent years. Thus, we noted that only six studies that examined this issue were completed between 1983 and 1990, and 15 were completed between 1998 and 2010.

- Among the 21 studies that examine this relationship explicitly, 13 used Model B to guide the analysis. This is problematic since Model B offers an overly simplistic framework for understanding the means by which school leadership impacts learning (Hallinger and Heck, 1996b).

- Similarly, the data in Table 6 indicate that Model B studies tended to use statistical tests that are less able to shed light on causal relationships among the relevant variables. In contrast, all of the Model C studies used Multiple Factor or Advanced Modeling techniques to explore relationships.

- Finally, it is interesting to note that there was no clear trend in the application of more powerful models and tests to this issue over time. Thus, for example, during the period 1998-2010, nine of these studies were guided by Model B frameworks and only six by Model C (not tabled).

We remind the reader that an implicit premise behind the development of the PIMRS was for use as a tool to examine the linkage between leadership and learning. About 35% of the studies engaged this empirical challenge in some fashion. However, when they did, the models and methods employed were generally inadequate to the task. This again supports an impression of many people ‘spinning their intellectual wheels’ without the ability to gain traction on the problem.

*Inter-relatedness and accumulation of knowledge.* Another way of shedding light on the research foci embedded in this set of studies concerns the degree to which they contributed
Knowledge accumulation represents as an important goal of empirical research, both for advancing theory-oriented and practical knowledge in an applied field. The concentration of PIMRS studies on a relatively small number of research foci (e.g., gender, teaching and administrative experience, school size, school effectiveness, student learning) suggested abundant potential for knowledge accumulation in this body of literature on instructional leadership. As noted, however, our analysis of conceptual models and research methods generated some degree of skepticism on whether that potential had been realized.

Therefore, we conducted citation analyses designed to illuminate the extent to which the studies had contributed to knowledge accumulation on this timely construct. Two features of the dataset rendered it particularly useful for this purpose. First, its strict delimitation to doctoral studies focused on instructional leadership would lend clarity to the interpretation of the results. Second, the inclusion of studies conducted over an extended period (i.e., almost 30 years) meant there would have been adequate time for the research findings to disseminate widely.

Two means of assessing knowledge accumulation were used: (1) citation frequency of the dissertations in the general literature, and (2) cross-citation of the dissertations by other authors within the sample. In phase one, the Google Scholar search tool was employed for assessing the citation impact of these studies in the literature on educational leadership and management. The author’s name and keywords taken from the title of the study were entered as search terms in Google Scholar’s advanced search engine. This generated a citation frequency number that was entered into the excel table for subsequent analysis. We note that the results from Google Scholar cast a quite broad net in terms of capturing studies as compared with other indices (e.g., ISCI). Thus the results could be interpreted as a ‘generous’ interpretation of impact.
The results of this citation analysis found the total citations of the dissertations was 366, with a mean of 3, a standard deviation of 14, and a range from 0 to 147. Eleven of the 126 studies accounted for over 75% of the citation impact; the original PIMRS dissertation (Hallinger, 1983; Hallinger & Murphy, 1985), by itself, accounted for 147 of the 366 citations. At first glance, 366 citations may suggest a reasonable degree of impact. However, when one considers the large number of studies and long duration covered by this database, this result reinforces the aforementioned impression of unfulfilled potential.

In phase two of the citation study, we analyzed the extent to which the authors built upon the work of previously completed PIMRS dissertations. As noted, the studies clustered around a relatively small number of research foci, thereby affording the researchers a prime opportunity to build upon the work of other investigators who had used the PIMRS to study instructional leadership. One would imagine, for example, that a doctoral student investigating the relationship between gender and instructional leadership in 2008 would build explicitly upon the empirical research of prior PIMRS users who had studied this issue going back to 1983.

For this analysis, we selected a subsample consisting of the 34 PIMRS dissertations completed since 2004, reasoning that these authors would have had access to the largest number of completed studies at the time of their research. We had obtained the full copy of the dissertation for 29 of the 34 studies; this was necessary in order to analyze the references. We examined the reference list in each of the 29 studies and found that the total number of cross citations of other PIMRS dissertations was 32, or a mean of just over one citation per study. Moreover, this exceedingly low cross-dissertation citation rate was skewed by one dissertation that accounted for 14 of the 29 citations. Again, neither type of degree or university was related to the pattern of citation for either type of citation analysis. This pattern of results indicates a remarkable absence of inter-connectedness among a group of studies in which the authors were
often investigating similar foci (e.g., the impact of gender or school size on instructional leadership), within the same sub-domain (i.e., instructional leadership), using the same instrument for data collection (i.e., the PIMRS).

These results provide greater clarity to our interpretation of the data that were presented on models, methods and foci as they pertain to knowledge accumulation. The citation analysis suggests that this body of studies has not even yielded an impact on knowledge accumulation among those are toiling into the same field of inquiry. This finding is startling, both for the magnitude (or lack thereof) and clarity of interpretation. That said, we also note that these results closely mirror those reported in the Bridges (1982) review, except with greater clarity and certainty. Indeed, his conclusion continues to ring true almost 30 years later: [T]hese results point to a lack of interconnectedness among the research studies on school administrators. This lack of systematic knowledge-building seems to be a pervasive characteristic of research in the field of educational administration (Bridges, 1982, p. 24).

Discussion

This review was undertaken with the stated purpose of contributing to our understanding of methodological progress in studying educational leadership and management. The report focused on the analysis of a dataset of 126 dissertation studies, all of which used the PIMRS instrument as a tool for data collection and that were completed between 1983 and 2010. In this final section, we summarize the main findings, revisit the limitations of this approach to reviewing a body of literature, and discuss the implications for future research in educational leadership and management.

Summary of Findings

This review sought to build explicitly on findings from prior reviews of research in educational leadership and management conducted (Bossert et al., 1982; Bridges, 1982; Haller,
1. Contrary to the expectations of several notable scholars (e.g., Barth, 1986; Cuban, 1984; Leithwood, 1994), interest in the construct of instructional leadership among practitioners and scholars appeared to be strong throughout the three decade period of the review. The data suggest that instructional leadership has achieved status as a core construct in educational leadership and management since early analyses conducted in the 1960s (e.g., Bridges, 1967; Gross & Herriott, 1965) and its full-scale emergence in the 1980s (e.g., Edmonds, 1979; Bossert et al., 1982; Purkey & Smith, 1983).

2. Consistent with findings reported in previous reviews (e.g., Bridges, 1982; Erickson, 1967; Haller, 1979; Hallinger & Heck, 1996b), this study found a strong preference for studying principals at the elementary school level. To our dismay, researchers continue to avoid the challenge of studying the more complex processes associated with instructional leadership in secondary schools. There were, however, exceptions to this trend that may provide useful approaches for using the PIMRS to study the distribution of instructional leadership in secondary schools (e.g., Todd, 2005).

3. The approaches to studying instructional leadership embraced in these dissertations closely mirror the broad trends reported by prior reviewers. That is, the studies tend to employ relatively simple conceptual models, cross-sectional survey designs, and bivariate statistical methods without the use of control variables. While we did note a substantial reduction in the percentage of studies that relied solely upon descriptive statistics, one could say that the distribution simply moved ‘up’ one level in our categorization scheme. Now the bulk of the studies used Level 2 methods (i.e., bivariate statistical methods without controls). Increased use of statistics that enable the researcher to gain real purchase on understanding the relationship among variables (i.e., Levels 3 and 4) was insufficient to support a claim of important progress on this score. Moreover, when we synthesized the findings across both conceptual models and research methods, the results suggested a lack of meaningful improvement in the overall conduct of research.
4. Finally, when we examined the research foci of these studies, the results were eerily reminiscent of findings from earlier reviews. The previously reported preference for studying personal antecedents of leadership in a relatively ad hoc fashion continues (Bridges, 1982). While studies of leadership impact on learning representing a substantial portion (35%) of the studies, the approaches used by the researchers were too often inadequate to this task. Our synthesis of the data on models, methods, and foci reinforced the picture painted in an earlier era (Bridges, 1982; Haller, 1979) of a lack of knowledge accumulation on important theoretical or practical problems. A subsequent citation analysis lent additional clarity to this conclusion; citation impact in the field at large was weak, and cross citation of the dissertations by other authors within the sample occurred so infrequently that it could be called a curiosity.

This last point is perhaps the most troubling since it speaks directly to the development of a deeper understanding of linkages between instructional leadership and learning. While evidence reported in this paper indicates that the PIMRS instrument is a potentially useful research tool, the findings suggest that this potential has yet to be tapped in a meaningful way, at least in doctoral research. Relatively few of the dissertation studies combined the use of comprehensive mediated-effects models with multivariate statistical methods as recommended in prominent research reviews (e.g., Bossert et al., 1982; Bridges, 1982; Haller, 1979; Hallinger & Heck, 1996b; Heck & Hallinger, 1999; Mulford, 2005) and modeled by scholars in widely-cited journal publications in educational leadership and management (e.g., Cheng, 1991, 1994; Hallinger et al., 1996; Heck et al., 1990; Kruger et al., 2007; Leithwood & Jantzi, 2000; Leitner, 1990; Marks & Printy, 2003; Mulford & Silins, 2003; Pounder, Ogawa, & Adams, 1995; Silins, 1994; Wiley, 2001). Thus, we conclude that the approaches to research embedded in these doctoral studies do not offer substantial leverage on understanding the critically important relationship of leadership and learning.
We also reported a rather unexpected finding that emerged from our examination of possible causal explanations for these methodological trends. Neither the dissertation model (i.e., Ed.D./Ph.D.) nor its brand category (i.e., level of research university) predicted the methodological approaches. In the field of educational leadership and management, it appears that, on average, a dissertation is a dissertation, regardless of where it is completed or the type of program.

Limitations

Earlier we noted that the findings from this review must be interpreted with the understanding that the data set consisted entirely of doctoral studies. While six of the 126 dissertations led to publications, this still leaves an open question as to the extent to which these trends characterize the broader empirical research on instructional leadership and educational leadership and management. While we cannot definitely answer these questions, we will attempt to place the findings in perspective.

We earlier noted that the results of this analysis of doctoral studies mirror trends reported earlier in reviews that examined doctoral dissertations alone (Haller, 1979) or doctoral dissertations and published research (Bridges, 1982; Hallinger & Heck, 1996b). Moreover, while doctoral dissertations do not represent the gold standard in research production, we noted that the majority of research in educational administration is carried out by doctoral students (Bridges, 1982, p. 13; Agusto, 2009; Archbald, 2008; Haller, 1979; Hoffer et al., 2005). Thus, while one cannot conclude that these trends generalize to published research in the field, they offer credible commentary on an important body of empirical literature.

Thirty years ago, Emil Haller (1979) offered a relevant observation based on his own assessment of methodologies employed in doctoral studies in educational leadership and management:
Evaluating the appropriateness of a research design is a complex matter involving the relative effectiveness and efficiency of competing alternatives for examining the particular question under investigation. Certainly, any thoughtful perusal of our research literature will turn up important scholarly studies based on cross-sectional questionnaire data that were subjected to a straightforward descriptive or bivariate analysis. And just as certainly, the same perusal would turn up studies of little or no practical or theoretical interest that were, nevertheless, either sophisticated multivariate statistical analyses of longitudinal data or tedious ethnographic accounts of the obvious. (p. 64)

This point seems equally applicable to the interpretation of the data presented in this review. This review sought to identify trends and draw conclusions about the quality of doctoral research by aggregating the characteristics of individual studies into a group profile. As Haller suggests, this does not mean that each study that used a particular combination of models and methods was an example of poor research. We note that it is also possible that the researchers were intent on exploring practical administrative problems that did not require sophisticated research approaches. Indeed one might even expect the Ed.D. dissertations to lean in this direction, though, as already noted we found no conceptual or methodological differences between Ph.D. and Ed.D. studies.

Another limitation of this review concerns the exclusive focus on studies that used the PIMRS instrument. This criterion naturally skewed the sample towards quantitative studies of instructional leadership. In fact, this dataset did include several mixed method case studies in which the PIMRS data were used in concert with qualitative data collected via interviews, observations, and/or school documents and data (Adams, 2002; Stroud, 1989). This is a useful trend and reinforces the contention that the broader program of research on leadership and learning should include quantitative, qualitative and mixed methods studies (Heck & Hallinger, 1999, 2005). Thus, it is not our contention that the only useful studies of instructional leadership must use multivariate methods to investigate complex conceptual models.
Implications for Research

Thirty years ago instructional leadership was identified as an approach that characterized the leadership practice of a relatively small number of ‘outlier’ principals (Bossert et al., 1982; Edmonds, 1979). Influential scholars reasonably questioned whether this could ever become a sustainable model in the broader practice of school leadership (Barth, 1986; Cuban, 1984). Indeed, just a decade after its adoption by policymakers, instructional leadership was pushed off center stage by growing interest in teacher leadership (Barth, 1990), and subsequently by transformational leadership (Leithwood, 1994), and distributed leadership (Gronn, 2002). Yet, the data offered in this report do suggest that instructional leadership has become entrenched in the firmament of professional practice (Gewirtz, 2003; Hallinger, 2003; MacBeath & Cheng, 2008; Stricherz, 2001) and remains an important focal construct in the eyes of scholars.

Evidence offered in this review indicates that the PIMRS can play a potentially useful role in empirical research on instructional leadership. Evidence was presented indicating that the instrument provides reliable and valid data on instructional leadership when the assessments come from teachers. That said, the instrument has been used for almost 30 years in a policy context that has altered quite dramatically, as well as in schools across a variety of different institutional and cultural contexts. Thus, it would seem advisable for researchers to continue to conduct analyses of reliability as a routine test in current studies rather than rely solely on the findings from prior research. Moreover, this point further implies that researchers continue to examine the validity of the PIMRS’ measurement of this construct in a changing global policy context.

With respect to desirable foci for research, we suggest that scholars who use the PIMRS more squarely accept the challenge of investigating the linkages between instructional leadership, school academic capacity, teacher effectiveness, and student learning. Numerous scholars have noted the need to shed light on the “black box” which contains the processes
through which leadership contributes to the improvement capacity of schools to create a positive impact on student learning. We noted increased interest in studying this issue over the last decade, and wish to encourage it further using comprehensive conceptual models, multivariate statistics, and also through mixed method studies.

Similarly, studies of how responsibilities for instructional leadership are shared or distributed between the principal and other staff seem timely and important (e.g., Marks & Printy, 2003; Todd, 2005), especially at the secondary school level. In our view, studies of the antecedents of instructional leadership, whether personal or contextual, are useful to the extent that they are linked to the impact of leadership. When antecedents are studied in relation to instructional leadership more substantial theorizing is required as well as methods that employ controls for other relevant variables.

As suggested above, we believe that impact should be studied in terms of student learning. However, worthy research can target other intermediate and distal variables such as teacher collective efficacy, satisfaction and commitment, school health, organizational learning, teacher change, and student engagement. Useful models for conducting empirical research on the relationship between school leadership and these variables using comprehensive models and robust statistical methods exist in the literature. Doctoral students are encouraged to draw upon these models and methods, rather than simply citing findings from these studies.

Discussion of the proper role and model of research for the Ed.D. and the Ph.D. goes beyond the brief of this report. Nonetheless, the findings from this review bear directly on the topic of research training in educational leadership and management at the doctoral level. Doctoral dissertations represent the extended effort of a large number of mid-career practitioners who are, presumably, intent on contributing to knowledge about the practice of educational leadership and management (Agusto, 2009; Archbald, 2008; Haller, 1979). Yet, the portrait presented here is of doctoral research that is neither making significant contributions to
theoretical knowledge nor to solving practical problems on a topic -- instructional leadership -- about which professionals tend to be passionate. The research reviewed in this paper does not even appear to be having an impact on the research of other scholars doing research in the same domain and who are using the same instrument.

It was noted that we were unable distinguish the methodological quality of dissertations based on the type of degree (i.e., Ed.D. or Ph.D.) or the level of research activity of the university. Citation analyses further confirmed this impression of unfulfilled potential for contributing to knowledge accumulation. If our results had indicated that both types of doctoral degree were generating productive knowledge, the field could perhaps live with this ambiguity of purpose, but they did not. Indeed, the finding that the Ph.D. dissertations suffered from the same limitations as the Ed.D. studies, raises critical questions about the nature of research training in Ph.D. programs. With this finding in mind, this review concludes that the Ph.D. in educational leadership and management is just as ill as the Ed.D. and the field must take steps to revitalize both.
References


descriptions/basic.php


Figure 1: PIMRS Conceptual Framework

PIMRS Framework

- **Defining the School Mission**
  - Frames the School’s Goals
  - Communicates the School’s Goals

- **Managing the Instructional Program**
  - Coordinates the Curriculum
  - Supervises & Evaluates Instruction
  - Monitors Student Progress

- **Developing the School Learning Climate Program**
  - Protects Instructional Time
  - Provides Incentives for Teachers
  - Provides Incentives for Learning
  - Promotes Professional Development
  - Maintains High Visibility

The final, definitive version of this paper has been published in Educational Administration Quarterly, 47(2), April 2011 by SAGE Publications Ltd, All rights reserved. © The University Council for Educational Administration 2011. The article is published: http://online.sagepub.com
<table>
<thead>
<tr>
<th></th>
<th>Almost Never</th>
<th>Almost Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. FRAME THE SCHOOL GOALS</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1. Develop a focused set of annual school-wide goals</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. Frame the school's goals in terms of staff responsibilities for meeting them</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3. Use needs assessment or other systematic methods to secure staff input on goal development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Use data on student academic performance when developing the school's academic goals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Develop goals that are easily translated into classroom objectives by teachers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From *PIMRS* - Teacher Form 1.3 (Hallinger, 1983, p. 2)
The final, definitive version of this paper has been published in Educational Administration Quarterly, 47(2), April 2011 by SAGE Publications Ltd, All rights reserved. © The University Council for Educational Administration 2011. The article is published: [http://online.sagepub.com](http://online.sagepub.com)

<table>
<thead>
<tr>
<th>Institutional Unit</th>
<th>Number of Studies by Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>School District</td>
<td>2</td>
</tr>
<tr>
<td>Multiple Levels</td>
<td>5</td>
</tr>
<tr>
<td>Elementary School</td>
<td>24</td>
</tr>
<tr>
<td>Middle School</td>
<td>2</td>
</tr>
<tr>
<td>High School</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43</strong></td>
</tr>
</tbody>
</table>

Table 1

Trend of PIMRS Studies by Institutional Level of Focal Administrator 1983-2010
Figure 3: Conceptual Frameworks for Studying Principal Leadership

**Model A-1:**
Antecedent Effects on Instructional Leadership

- Antecedent Variables (Principal Characteristics and/or School Context Factors) → Instructional Leadership

**Model B:**
Direct Effects of Instructional Leadership

- Instructional Leadership → School-Level Factors

**Model C:**
Mediated Effects of Instructional Leadership

- Instructional Leadership → Mediating School-Level Variables → School-Level Outcomes or Conditions

**Model D:**
Reciprocal Effects of Instructional Leadership

- Instructional Leadership → Mediating School-Level Variables → School-Level Outcomes or Conditions

adapted from Pitner, 1988, pp. 105-108
<table>
<thead>
<tr>
<th>Model</th>
<th>Description of Conceptual Model</th>
<th>Number of Studies</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1</td>
<td>Antecedent effects of Principal Demographics on Instructional Leadership</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>A-2</td>
<td>Antecedent effects of School Context Variables on Instructional Leadership</td>
<td>33</td>
<td>64</td>
</tr>
<tr>
<td>A-3</td>
<td>Antecedent effects of Demographics and Context on Instructional Leadership</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>B-1</td>
<td>Direct effects of Instructional Leadership on School Variables or Outcomes with Antecedent Variables included in the model.</td>
<td>26</td>
<td>45</td>
</tr>
<tr>
<td>B-2</td>
<td>Direct effects of Instructional Leadership on School Variables or Outcomes without Antecedent Variables included in the model.</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>C-1</td>
<td>Multivariate effects of Instructional Leadership on School Variables and Outcomes with Antecedent Variables included in the model.</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>C-2</td>
<td>Multivariate effects of Instructional Leadership on School Variables and Outcomes without Antecedent Variables included in the model.</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Reciprocal effects of Instructional Leadership on School Variables and/or Outcomes</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>Not applicable; used another design such as case study or profile of principals</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

*Note that the total is higher than the sample of studies due to numerous studies using more than one model to address multiple research questions. Some studies could not be categorized by this coding scheme.

Table 2
Summary of Theoretical Models Used to Study Instructional Leadership
<table>
<thead>
<tr>
<th>Period</th>
<th>Model A</th>
<th>Model B</th>
<th>Model C</th>
<th>Model D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983-91</td>
<td>22</td>
<td>17</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1992-2000</td>
<td>18</td>
<td>13</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2001-2010</td>
<td>24</td>
<td>15</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td><strong>Raw Total</strong></td>
<td><strong>64</strong></td>
<td><strong>45</strong></td>
<td><strong>12</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td><strong>% Total</strong></td>
<td><strong>50%</strong></td>
<td><strong>36%</strong></td>
<td><strong>9%</strong></td>
<td><strong>0%</strong></td>
</tr>
</tbody>
</table>

Table 3
Distribution of Model Frequency of Use Over Time
<table>
<thead>
<tr>
<th>Statistical Approach/Period</th>
<th>Descriptive</th>
<th>Single Causal Factor-Correlational</th>
<th>Single Causal Factor-Correlational with Controls</th>
<th>Multiple Factor</th>
<th>Advanced Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983-91</td>
<td>5</td>
<td>26</td>
<td>3</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>1992-2000</td>
<td>5</td>
<td>21</td>
<td>7</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>2001-2010</td>
<td>2</td>
<td>21</td>
<td>5</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Raw Total</td>
<td>12</td>
<td>68</td>
<td>15</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>% of Total</td>
<td>10%</td>
<td>54%</td>
<td>12%</td>
<td>24%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 4
Breakdown of Studies Over Time by Statistical Approach"
<table>
<thead>
<tr>
<th>Level</th>
<th>Statistics</th>
<th>Ed.D.</th>
<th>Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>2</td>
<td>Single factor w/out control</td>
<td>56%</td>
<td>46%</td>
</tr>
<tr>
<td>3</td>
<td>Single factor with control</td>
<td>13%</td>
<td>11%</td>
</tr>
<tr>
<td>4</td>
<td>Multiple Factor</td>
<td>23%</td>
<td>31%</td>
</tr>
<tr>
<td>5</td>
<td>Advanced Modeling</td>
<td>0%</td>
<td>3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Conceptualization</th>
<th>Ed.D.</th>
<th>Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Antecedent effects</td>
<td>47%</td>
<td>49%</td>
</tr>
<tr>
<td>B</td>
<td>Direct effects</td>
<td>37%</td>
<td>34%</td>
</tr>
<tr>
<td>C</td>
<td>Mediated effects</td>
<td>9%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Table 5
Comparison of Ed.D. and Ph.D. Dissertations by Statistical Level And Conceptual Models
Table 6
Classification of Studies of Instructional Leadership and Student Achievement by Conceptual Model and Level of Statistical Method

<table>
<thead>
<tr>
<th>Model</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>B-2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>C-1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>C-2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

The author would like to acknowledge the useful feedback offered on this manuscript by Edwin Bridges whose own review of the literature in 1982 offered important baseline data against which the findings from this study could be compared. This study was supported in part by funding support of the Research Grant Council (RGC) of Hong Kong for its support through the General Research Fund project # 840509.

Analysis of the substantive findings of the studies will be reported in a separate paper.

With respect to research interest in the construct, a Google Scholar search in March 2010 found that the most frequently cited articles in Educational Administration Quarterly since its inception were the Hallinger and Heck (1996) and Bossert et al (1982) reviews of research, both of which focused on instructional leadership and learning outcomes in schools.

The author notes that the total of 126 doctoral studies included 1 Ed.S project as well. Since it was indistinguishable from the doctoral studies in most scope and approach it was included in the sample as an Ed.D. study. Six Master degree studies were also identified, but these were not included in the analyses for this report. By way of information, the Master degree studies used Models A (1) and B (2) with two NA. In terms of statistical methods, 2 studies used Level 1 and 4 used Level 2 methods.

Most of the studies classified as NA were either case studies or studies that employed a single variable (i.e., instructional leadership). For example, some studies simply generated an instructional leadership profile for a group of principals without conducting an analysis in relation to other variables (e.g., Corkill, 1994).

It should be noted that this type of study is more properly described as a moderated effects design (Baron & Kenny, 1986). However, for the purposes of this report, the author simply wished to denote the fact that the researcher was examining the effects of another variable on the principal’s instructional leadership.

The Mantel-Haenszel test (nominal by ordinal chi-square test) showed there’s no significant difference between EdD and PhD in using 4 types of methods (p=.338). A second run of the same test comparing only two groups (Methods 1, 2, and 3) and (Method 4) also showed no significant difference between Ed.D. and Ph.D. in using Method 4, Multiple Factors, (p=.354), although Ph.D. dissertations tended to use Method 4 slightly more than Ed.D. dissertations (Ph.D. 32.4% vs. Ed.D. 22.7%).

Note that for the purposes of this analysis, publications based on the dissertations were counted as well.

Whereas this review focused solely on PIMRS studies of instructional leadership, Bridges’ review examined studies across a broad range of topics in educational administration. Thus, the interpretation of findings from his citation analysis was clouded somewhat by the selection criterion. For example, one would not necessarily expect studies in school finance to be cited in studies of principal work activities or school climate. In contrast, the citation analysis in this review had a built-in ‘control’ since all of the studies were conducted in the same sub-domain of the field.

Note that the total number of studies listed here does not match the total number of studies included in the report because some studies used more than one method.