Methodologies for Studying School Leadership:
A Review of 25 years of Research Using the
Principal Instructional Management Rating Scale

Dr. Philip Hallinger

Paper prepared for presentation at the annual meeting of the
American Educational Research Association
New York, March 2008

1 Dr. Philip Hallinger specializes in the study of leadership, leadership development organizational change, and school improvement. He can be reached at Hallinger@gmail.com The author wishes to acknowledge the contributions of numerous graduate students who worked on this project over the years including Melissa Hall Barron, Patchanee Taraseina, Ellen Singer, David Ng, Charlie Hausman, Parinya Showanasai and Apichai Somboonpong. He would also like to acknowledge the comments of Ronald Heck on parts of the paper.
Methodologies for Studying School Leadership:  
A Review of 25 years of Research Using the PIMRS

Abstract
Influential reviews of research on school leadership conducted over the past 25 years (e.g., Erikson, Bridges, Bossert and colleagues, Leithwood, Hallinger and Heck) have examined the methodologies used by scholars to understand school leadership and its effects. This paper examines the evolution of scholarship in school leadership by focusing on methods used by researchers over the past 25 years in studying instructional leadership. The paper reviews 119 doctoral studies conducted between 1983 and 2008, all of which used the Principal Instructional Management Rating Scale (Hallinger, 1983) as the instrument for data collection. The paper offers conclusions concerning trends in the types of research questions being asked, analytical methods in use, application of theory, as well as general findings concerning how instructional leadership is employed by school principals.
The early 1980’s witnessed the advent of a period of educational reform in the United States that has demonstrated surprisingly long staying power. Among the educational trends that emerged during that era, few have been more significant or widespread than the continuing focus on principal effectiveness. During the 1980s, American educational policymakers, eager to change practice in schools, came to view school principals as key agents in the reform of schools and classrooms. Research on school effectiveness and school improvement reinforced this perspective on the importance of principals in policy implementation (Edmonds, 1982; Purkey & Smith, 1983). These bodies of research identified principal instructional leadership as a key factor in successful schools and refocused the attention of scholars on school principals (Bossert, Dwyer, Rowan & Lee, 1982; Leithwood & Montgomery, 1982; Murphy, Hallinger & Mitman, 1983).

During the ensuing 25 year period, the focus on principal leadership has continued unabated with an ever-expanding global interest, though the choice of lens for viewing school leadership has shifted periodically. Between 1992 and 2002, interest in the instructional leadership lens appeared to wane somewhat as interest in transformational leadership, teacher leadership, and distributed leadership increased. Interest in these models reflected the deepening concerns of education professionals and policymakers for teacher empowerment, teacher career development, and the recognition that “principals cannot do it alone” (Barth, 1991; Donaldson, 2001; Leithwood, Jantzi, & Steinbach, 1999).

Since the turn of the millennium, however, the leadership pendulum has swung back again. The North American infatuation with educational performance standards has turned into a global love affair with accountability (Leithwood, 2003; Murphy, 2002; Murphy & Shipman, 2003). Principals again find themselves at the nexus of accountability and school improvement with an increasingly explicit expectation that they will function as instructional leaders (Gewirtz, 2003; Leithwood, Louis, Anderson, & Wahlsttom, 2004; Stricherz, 2001a, 2001b).
Given the passage of formal government standards for education through the world, principals who ignore their role in monitoring and improving school performance as instructional leaders do so at their own risk (Bolam, 2003; Jackson, 2000; Lam, 2003; Leithwood, 2003; Southworth, 2002; Tomlinson, 2003).

In a review of literature on principal leadership effects conducted a decade ago, Hallinger and Heck (1996a) noted that instructional leadership had been the most prevalent perspective adopted by researchers studying principal leadership between 1982 and 1995 and identified the Principal Instructional management Rating Scale (PIMRS; Hallinger 1982, 1990) as the most commonly used instrument. The Hallinger and Heck (1996a, 1996b, 1998) reviews, however, focused on studies that linked principal leadership and student achievement outcomes and did not specifically on instructional leadership.

These limitations of the prior reviews form the rationale for this paper. This paper examines studies that explicitly employed an instructional leadership perspective to study principal leadership. With the reemergence of interest in instructional leadership, it is timely to examine what has been learned from systematic empirical investigation of this construct over the past 25 years. The PIMRS has been used in over 119 studies since its development in 1982, thereby offering a useful opportunity to examine patterns in both results and methodologies. The use of a single instrument in such a large number of studies over an extended period also provides a unique opportunity to investigate the appropriateness of conceptual models and methods used by researchers in our field (Heck 2006, personal communication).

This report presents a review of studies that used the PIMRS to study principal instructional leadership between 1982 and 2006. In a broad sense, the review seeks to examine patterns in the methodologies used to study instructional leadership with the PIMRS instrument over the past 25 years. The specific objectives of the review include:
1. To identify the most frequently studied research questions studied by users of the PIMRS instrument;
2. To describe and assess the methodological approaches employed by users of the PIMRS;
3. To assess the appropriate role of the PIMRS in future research on instructional leadership.

Theoretical Perspective

In the early 1980s, despite ardent advocacy among policymakers for stronger principal instructional leadership, scholars were quick to note the paucity of systematic investigation of this role. Scholars identified several specific factors that impeded development of a clear understanding of how principals contributed to school effectiveness and improvement:

• Lack of theoretical models that articulated how this role influenced student learning;
• Lack of clearly explicated conceptual frameworks for studying relevant constructs;
• Lack of valid and reliable instrumentation for exploring the role empirically;
• Reliance on studies poorly designed to test for causal effects. (Bossert et al., 1982; Bridges, 1982; Cuban, 1988; Leithwood & Montgomery, 1982; Miskel, 1982; Murphy, Hallinger & Mitman, 1983)

These limitations were cause for concern, especially in light of the burgeoning attempts to put the findings from these studies into practice. For example, based upon his assessment of the literature on the school principal, Bridges (1982) 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)"
These conflicting trends between policy and scholarship posed a challenge for scholars. In response, several international efforts were undertaken to develop stronger methodologies for studying principal leadership. This included the development of new conceptual frameworks as well as instrumentation. Notably these early efforts focused primarily on the role of the principal as an instructional leader.

Important contributions were made during the subsequent decade by Andrews and colleagues at the U. of Washington (e.g., Andrews & Soder, 1987; Andrews, Soder, & Jacoby, 1986; Bamburg & Andrews, 1990), Glasman at the U. of California, Santa Barbara (e.g., Glasman, 1983, 1984), Hallinger and Murphy at Vanderbilt U., (e.g., Hallinger, 1982, 1990; Hallinger, Bickman & Davis, 1996; Hallinger & Murphy, 1985; Murphy, 1988; Murphy et al., 1983), Heck and colleagues at the U. of Hawaii (e.g., Heck, 1992, 1993; Heck, Larsen, & Marcoulides, 1990), Cheng and colleagues in Hong Kong (e.g., Cheng, 1991, 1994), Leithwood and colleagues at OISE in Canada (e.g., Leithwood, Begley & Cousins, 1990; Leithwood & Montgomery, 1986), Southworth and others in the United Kingdom (Day, Harris, & Hadfield, 2001; Jackson, 2000; Southworth, 2002), Villanova and colleagues in the State Education Dept. in Connecticut (Villanova, Gauthier, Proctor, & Shoemaker, 1981), and a variety of researchers in the Netherlands (van de Grift, 1987, 1989, 1990; Witziers, Bosker, & Kruger, 2003).

This emergent scholarly focus on instructional leadership included the development of several instruments designed to measure principal instructional leadership (e.g., Andrews, Soder, & Jacoby, 1986; Hallinger, 1982, 1984a; Leithwood & Montgomery, 1986; van de Grift, 1990; Villanova, Gauthier, Proctor, & Shoemaker, 1981). Given the development of these research tools, investigators subsequently generated a substantial body of research on principal instructional leadership (Hallinger & Heck, 1996a, 1996b, 1998). As noted earlier, the most commonly used instrument has been the Principal Instructional Management rating Scale or PIMRS (Hallinger, 1982). Studies using this instrument represent the focus of this paper.
Instructional Management Framework

The theoretical framework for this review is based on the conceptual framework underlying the PIMRS instrument. The PIMRS assesses three dimensions instructional leadership: Defining the School’s Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate (Hallinger & Murphy, 1985; see Figure 1).

PIMRS Framework

These dimensions are further delineated into 10 instructional leadership functions, each of which is measured by behaviorally anchored items. 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 the dimension does not assume that the principal defines the school’s mission alone, it does assume that the principal’s responsibility to ensure that such a mission exists and to
communicate 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 developing the school’s instructional program. Yet this framework assumes that development 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. It conforms to the notion that successful schools create an “academic press” through the development of high standards and expectations and a culture of continuous improvement.

**Instructional Leadership vs. Instructional Management**

During the early 1980’s the term instructional leadership became increasingly popular. In their seminal review of the literature, Bossert and his colleagues (1982) defined a similar construct as instructional management. They selected the term instructional management because they inferred that this role of the principal – at least as discussed in the studies of effective schools – revolved around traditional centralized *managerial* functions of directing, coordinating and controlling. The difference was that these principals focused their efforts at coordination and control more explicitly upon curriculum and instruction.

Over the course of the 1980’s, scholars increasingly came to conceive of *instructional management* as a highly directive role centered on the principal. This contrasted, for example,
with conceptions that emphasized *shared or distributed leadership* in the school (e.g., Gronn, 2000, 2002; Lambert 2002). While this distinction held weight among scholars, practitioners came to use instructional leadership and instructional management synonymously. The distinction was further muted during the 1990’s. With the subsequent conceptualization of the transformational leadership role of principals, instructional leadership and instructional management came to be treated as one and the same, even among researchers. The aforementioned description of the PIMRS items bears out the blended focus on management of the instructional program through leadership via the school’s vision and culture.

**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, 1990). 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. Studies have 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.

Figure 2
Sample PIMRS Rating Subscale: Teacher Form

To what extent does your principal...?

I. FRAME THE SCHOOL GOALS

<table>
<thead>
<tr>
<th></th>
<th>Almost Never</th>
<th>Almost Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Develop a focused set of annual school-wide goals</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2. Frame the school's goals in terms of staff responsibilities for meeting them</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3. Use needs assessment or other systematic methods to secure staff input on goal development</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4. Use data on student academic performance when developing the school's academic goals</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5. Develop goals that are easily translated into classroom objectives by teachers</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

From PIMRS - Teacher Form1.3 (Hallinger, 1983, p. 2)

It should be emphasized here that a high score on a particular leadership does not necessarily indicate effective performance, only active leadership in that area. Principals who obtain a high rating on a given leadership function are perceived as engaging more frequently in instructional leadership behaviors and practices associated with principals in effective schools. The PIMRS ratings do not, however, measure the quality of principal instructional leadership. Nor does the scale address the thinking that underlies the exercise of the specified leadership behaviors, potentially important information for understanding how and why behaviors are enacted in context (see Leithwood et al., 1990).

Thus, the data generated by the PIMRS is used most effectively to highlight patterns in the instructional leadership of principals. For individual principals these data can be used to identify relative emphasis given to different facets of the instructional leadership role. When
data from a principal’s self-assessment is compared with data obtained from other role groups (e.g., teachers’ and the principal’s perceptions), this can be useful for problem-solving or goal-setting. When data for a number of principals are aggregated across an administrative unit (e.g., school district or cluster), the resulting profile will again highlight areas of relative activity in different aspects of the principals’ instructional leadership role. This can be used for needs assessment, program evaluation, policy analysis, or research (Hallinger & Murphy, 1985).

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 initial validation study in its conclusion that the scale provides reliable data on instructional management (e.g., Howe, 1995; Jones, 1987; Leitner, 1989; O’Day, 1984; Taraseina, 1993).

The original validation study further tested the instrument 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, 1989; O’Day, 1984; Taraseina, 1993).

Methodology of the Review

This research review employs a critical synthesis approach to the analysis of the studies. Although the author would have preferred to adopt a meta-analytic approach, this proved impractical due to various limitations of the studies that are described in the paper.
The Sample

The author searched Dissertation Abstracts for studies that used the *PIMRS* between 1983 and 2001 using a variety of keywords: instructional leadership, *PIMRS*, principal leadership, instructional management. The resulting sample for the review included 118 completed studies that used the *PIMRS* as a means of collecting data on principal instructional leadership. Of these six were Master theses and the others were doctoral dissertations.

Most of the dissertations were downloaded in *pdf* format and printed out for review and analysis. In selected cases, the full dissertation file was not accessible and data was instead gleaned from the abstract. After reviewing these studies, it was determined that there was insufficient data in the abstract for seven studies (Coltharp, 1989; Haasl, 1989; Kroeze, 1992; Rose, 1991; Roudebusch, 1996; Ryans, 1989).

The author of the review obtained access to an additional seven dissertations that did not show up in the search. Some had been completed in foreign countries or at North American universities that did not routinely provide dissertations to UMI. The full set of studies is listed in Table 1. They totaled 119 studies.

Organization of the Review

A team of researchers worked intermittently over a period of several years to read the studies and to generate and collate the necessary information. Information was extracted from the studies according to the following categories:

- Research questions and hypotheses;
- Variables included for study;
- Theoretical framework;
- Research Design;
- Instrumentation;
- Sample and procedures;
- Data analysis approach and tests;
• Findings.

These data were compiled within studies and gradually constructed into tables to display the characteristics of the studies as a group. Notations were made as well as to strengths and weaknesses of individual studies. These were used to help understand the patterns that emerged both respect to methods and results.

Limitations of the Study

This study has several salient limitations that it is best to identify in advance. First, this paper will not report on the results of this body of studies. Its purpose is limited to understanding how scholars have used the PIMRS in studying principal instructional leadership. Second, the study is limited to master and doctoral degree dissertations. Only a few of these dissertations were subsequently published in refereed articles. Consequently, there is wide variation in the quality of the sampled studies. Studies ranged in quality from poor to excellent.

Despite this limitation, there are also substantial grounds for the validity of this sample.

• In the field of educational administration, doctoral dissertations traditionally represent the largest source of empirical data.

• The PIMRS instrument has been validated. Therefore, despite other limitations of the dissertation studies, they used a common instrument. This may yield interesting and useful information not only about instructional leadership, but also about how researchers are studying this popular topic.

• Finally, as noted above, the study reviewed all 119 studies that had used the PIMRS to study instructional leadership. This will provide a potentially useful picture of the broad trends in doctoral studies of principal leadership as conducted over the past 20 years.

Results

The results of the review are presented in this section.
Time-based Trends in Studies of Instructional Leadership

The first question to be considered is how the use of the PIMRS has varied over time. If we break the time-span of the review into three-year periods, the frequency of studies is as follows:

- 1983-1988: 19 studies
- 1989-1994: 43 studies
- 1995-2000: 28 studies
- 2001-2006: 28 studies
- 2007-2008: 3 studies

These figures suggest a waxing and waning of interest in the topic of instructional, at least as measured by use of the PIMRS. The first half of the period of this review, 1983-1991, shows the initial spurt of interest in instructional leadership following the emergence of the effective schools movement. As the PIMRS instrument became more widely known, the timeliness of the topic of instructional leadership generated many additional studies.

During the mid-1990’s, however, attention shifted somewhat away from effective schools and instructional leadership. Interest in these topics was displaced by concepts such as school restructuring and transformational leadership. Indeed, during the early 1990’s, a noted scholar, Ken Leithwood, explicitly questioned the viability of the instructional leadership construct in light of the changing context and needs of schools during an era of rapid reform. He questioned whether the instructional leadership construct would last in the face of these changes. Nonetheless, if we look at the subsequent 12 year period from 1995 to 2007, it is clear that interest in the topic remains consistent and strong.

Who Conducted the Studies?

As noted earlier, six of the studies were Master degree Theses. The remaining 113 studies were doctoral theses completed by students completing degrees in Doctor of Education or Doctor of Philosophy.
The studies were conducted in 8 countries. These included the United States (103), the Philippines (4), Canada (3), Thailand (3), Taiwan (2), England (1), Hong Kong (1), Cameroon (1), Guam (1). This may somewhat under-report the number of countries. Not all of the studies conducted in foreign countries were reported in the dissertation abstract searches and the authors in foreign nations did not always send their completed reports to the author.

The 119 studies were conducted at 74 different universities. This suggests a wide knowledge of the instrument. The most frequent users were found at the following universities: U. of South Carolina (7), Texas A & M (6), Stanford University (4), U. of Texas (5), U. of Virginia (4). There are numerous instances of the instrument being used three times at other universities.

Who Was Studied?

A continuing question of interest concerns whom scholars choose to study in the field of educational administration. In earlier reviews of research on school administrators, Bridges (1982) and Erickson (1967) found a preference among scholars for studying administrators at the elementary level. This review analyzed the subsequent trend among those studying instructional leadership with the PIMRS. The breakdown among this set of studies was as follows:

- District level (i.e., superintendents): 3 studies
- Cross Levels: 15 studies
- Elementary School: 59 studies
- Middle School: 9 studies
- High Schools: 33 studies

These data show a clear, continuing preference for studying principals at the elementary school level. Reasons for this persisting preference for studying elementary school principals are several.
First, availability and access are key issues for doctoral students. Since there is often a ratio of anywhere from 3 to 1 to 6 to 1 of elementary to secondary schools, availability for samples makes elementary schools an easier choice for studies that require multiple schools.

Second, it is worth reminding the reader of the linkage between the concept of instructional leadership and that of effective elementary schools. The effective schools movement was borne through a set of studies conducted primarily at the elementary school level. These studies found evidence of strong instructional leadership among principals of urban elementary schools serving poor children. Subsequently, there was much interest in validating these findings. Indeed, the PIMRS was itself only validated initially at the elementary school level.

It was only after the initial burst of enthusiasm for this construct that scholars and practitioners began to question its applicability to secondary schools that tend to be larger and more highly differentiated in their staffing. This pattern is reflected in the results. When the data were sorted by both year and level, we found that only four out of the total of 24 secondary studies were published between 1983 and 1990. The largest portion was published during the latter period.

It is also notable who was not studied. Neither assistant principals, department chairpersons nor other senior teachers or teacher-leaders were studied. While it is true that the PIMRS was only validated for principals, it does still seem curious that nobody undertook its adaptation for these role groups during an era of teacher empowerment. This omission is even more salient given the widely accepted need to create a broader capacity for instructional leadership in schools.

Who were the respondents?
In his 1982 review, Bridges also commented on the inadequate size of samples used for the study of school administrators. In assessing the samples for this review, size was only one of several relevant factors.

The first factor considered here is the source of the data. The initial validation study reported that the validity of the PIMRS results varied by respondent role group (Hallinger, 1983; Hallinger & Murphy, 1985). The PIMRS yielded valid results when data were collected from teachers. However, data collected from principal self-reports and supervisors were skewed when compared with other data sources. That is, the teacher results matched most closely with results obtained from interviews and documentary evidence. This led to the conclusion that for the purposes of research, it was recommended that users of the PIMRS rely on teacher reports more than on self-report or supervisor’s assessments.

Most PIMRS users complied with this recommendation. Twenty-seven studies were completed using only principal self-reports. This is a positive sign. Use of the teacher form of the instrument offers greater confidence in the results.

This factor is especially relevant since only 10 of the 119 studies explicitly retested their samples for reliability and or validity of the PIMRS. The others accepted the initial validation of the instrument as sufficient and simply carried out the analyses related to their research questions. Although reliability and validity are discussed later in the paper, here it is worth noting that the instrument was generally used in a manner consistent with its conception and purpose.

Methods of Studying Instructional Leadership

The presentation of results thus far may be further illuminated by additional examination of the models used by the researchers as they employed the PIMRS. The theoretical models used in the research refer to the frameworks used by the authors to select the variables for study and then to organize relationships among the variables.
For the purposes of this review, the studies were classified into five general models: Antecedent Effects Studies, Direct Leadership Effects Studies, Meditated Effects Studies, Reciprocal Effects Studies, Moderated Effects Studies (see Figure 2). This classification scheme was adapted from Pitner (1988) and used in previous research reviews conducted by Hallinger and Heck (1996a, 1996b). Briefly, these models describe the manner in which the study authors conceptualized the relationship between instructional leadership and the other variables in their studies.

Figure Two: Conceptual Frameworks for Studying Principal Leadership

After reviewing the 119 studies, they were classified into these models.
Antecedent effects models. Studies using an Antecedent Effects Model studied the effects of either principal demographics or school context factors on the instructional leadership of the sample of principals. Studies that explored the relationship of principal demographics to instructional leadership typically examined variables such as the principal’s age, experience, gender, self-efficacy, teaching experience or knowledge of instruction. These studies were classed as Model A-1.

Studies that explored the relationship of the school context to instructional leadership typically included variables such as the school size, school level, district size or type (see Table One). These studies were classed as Model A-2. When studies incorporated both types of Antecedents; they were classified as Model A-3.

As displayed in Table One, Antecedent Models accounted for 30 of the 77 studies. There was an almost fifty-fifty split among the two basic types of this model: demographic and context factors.

Direct leadership effects model. The next classification of studies was the Direct Leadership Effects Model. As shown in Figure Two, this model studies the relationship between instructional leadership and a second variable, usually an in-school variable (school climate, school missions) or school outcomes (e.g., teacher satisfaction, student achievement, school effectiveness). Other studies examined the relationship between instructional leadership and another measure of leadership, such as transformational leadership.

This model included a sub-group of studies that examined the relationship between instructional leadership and student achievement and school effectiveness. When conceptualized as a Model B study, the authors sought to answer the question: Is there a relationship between principal instructional leadership and student achievement? As these
studies all sought to understand the direct relationship between instructional leadership and an outcome variable, they can be thought of as bivariate studies.

In some cases, antecedent variables were also studied in terms of their direct effect on instructional leadership. These studies were classified as B-1 Model. Notably, however, even when antecedent variables were included in these studies, the investigators still tended to study relationships in terms of direct effects as displayed in Figure Two.

When antecedent variables were not included in the model, the studies were classified as B-2. A total of 39 studies were classified as direct leadership effects studies: Twenty-three included antecedents (B-1) and sixteen did not (B-2). Therefore, studies of the

Mediated effects models. Model C studies represent efforts to investigate the mediated effects of instructional leadership on school outcomes, with or without the inclusion of antecedent variables. The difference between Model C and Model B is that the mediated effects studies seek to understand the avenues through which instructional leadership influences school outcomes such as achievement. They answer the question: How does leadership influence student achievement and school effectiveness?

It is interesting to note that none of the studies included in this sample used this approach. This is notable because in our earlier reviews Ronald Heck and I concluded that this model was among the most powerful approaches to studying school leadership and its effects. It represents a significant advance over direct effect models in its ability to illuminate relationships. Yet none of the doctoral or masters studies utilized this Model.

Reciprocal effects model. Model D was Reciprocal Effects model. In this approach the researcher seeks to understand the interactive effects of variables without assuming the direction of effects in advance. As noted by Hallinger and Heck (1996a, 1996b) this is a highly sophisticated approach to studying leadership effects. It does, however, require longitudinal data that may be beyond the reach of most doctoral students. No studies employed this model.
Table One

Summary of Theoretical Models Used to Study Instructional Leadership

<table>
<thead>
<tr>
<th>Model Name</th>
<th># of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-1: Antecedent Effects of Principal Demographics on Instructional Leadership</td>
<td>22</td>
</tr>
<tr>
<td>A-2: Antecedent Effects of School Context Variables on Instructional Leadership</td>
<td>17</td>
</tr>
<tr>
<td>A-3: Antecedent Effects of Demographics and Context on Instructional Leadership</td>
<td>9</td>
</tr>
<tr>
<td>B-1: Direct Effects of Instructional Leadership on School Variables or Outcomes with Antecedent Variables included in the model.</td>
<td>25</td>
</tr>
<tr>
<td>B-2: Direct Effects of Instructional Leadership on School Variables or Outcomes without Antecedent Variables included in the model.</td>
<td>16</td>
</tr>
<tr>
<td>C-1: Multivariate Effects of Instructional Leadership on School Variables and Outcomes with Antecedent Variables included in the model.</td>
<td>5</td>
</tr>
<tr>
<td>C-2: Multivariate Effects of Instructional Leadership on School Variables and Outcomes without Antecedent Variables included in the model.</td>
<td>3</td>
</tr>
<tr>
<td>D: Reciprocal Effects of Instructional Leadership on School Variables and/or Outcomes</td>
<td>0</td>
</tr>
<tr>
<td>Non-classified</td>
<td>18</td>
</tr>
</tbody>
</table>

Research Designs and Statistical Approaches

The theoretical models discussed above lend themselves to various research designs and approaches to data analysis. Given that the PIMRS was designed as an instrument for survey research, it is not surprising that most of the studies employed statistical rather than qualitative analysis. Nonetheless, as noted below, there were several case studies in which the author employed the PIMRS as one tool for data collection.

*Research designs.* Researchers used a variety of research designs to study instructional leadership using the PIMRS. These included case studies, usually employing multiple methods and descriptive statistics, as well as a variety of post hoc or ex post facto designs. No
longitudinal, experimental or quasi-experimental designs were found among the sampled studies.

The most commonly used research design was a cross-sectional, post-hoc, design. In this type of design the researcher sampled a population of schools and studied principal leadership in relation to other variables. As suggested above, this type of design could take the form of Model A (Cunningham, 2004; Moore, 2003; Sawyer, 1997; Sterrett, 2005; Trout, 1985), B (Campbell, 1998; Cantu, 1994; Geiselman, 2004; Howe, 1995; Maciel, 2005; MacNeil, 1992; Meek, 1999) or C (Balsamo, 2004; Jones, 1987; Knezek, 2001; Leitner, 1989; Mallory, 2002; O’Donnell, 2002; Prater, 2004; Singleton, 2006; Van Pelt, 1993) depending upon the variables selected for study. This type of design can range from simple bi-variate models (e.g., Model A or B) to more complex models (Model C). Regardless of the complexity of the model, limitations inherent in this type of cross-sectional survey study, make it difficult to tease out questions of causal relationships (e.g., between leadership and student learning). This is especially true in the absence of a theoretical model, a substantial sample, and sophisticated analytical tools (Heck & Hallinger, 1999).

A second type of cross-sectional survey research design sought to compare perceptions of the principals’ instructional leadership across respondent role groups (e.g., principals with teachers). This type of study did not seek causal explanation but instead simply sought to determine if differences existed in perceptions across role groups. The purpose of this type of study derived from an interest in understanding the validity of perceptions from teachers, principals, and supervisors (e.g., Chi, 1997; Hallinger, 1983; Haack, 1991; Haasl, 1989; Marshall, 2005; Saavedra, 1987; Smith, 2007; Stevens, 1996; Taraseina, 1993; Vinson, 1997).

A third type of research design observed in this body of dissertations was the contrasting groups design. In quite a few instances researchers employed the school as the unit of analysis and selected schools based on specific characteristics such as effectiveness, poverty level, test
performance, or some combination of such criteria. The school was then considered as the independent variable and the researcher sought to examine differences in the instructional leadership of the principals in the contrasting groups (e.g., Bravo., 1991; Cantu, 1994; Dickerson, 1999; Garcia, 1999; Gerrell, 2006; Gibson, 2005; Hunter, 1994; Johnson, 2005; Johnson, 2006; Krug, 1986; Lubbers, 1996; Orange, 1990; Palmer, 2000; Rose, 1991; Stroud, 1989; Winger, 1992; Zeanah, 1986).

Statistical approaches. The theoretical models employed by these researchers lend themselves to a variety of different types of statistical analysis. The breakdown along these approaches is displayed in Table Three.

<table>
<thead>
<tr>
<th>Statistical Approach</th>
<th>Case Study</th>
<th>Descriptive Only</th>
<th>Bivariate without controls</th>
<th>Inferential with controls</th>
<th>Advanced Inferential</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Studies</td>
<td>4</td>
<td>14</td>
<td>46</td>
<td>53</td>
<td>4</td>
</tr>
</tbody>
</table>

In the simplest type of study, researchers developed profiles of the principals and described the pattern of instructional leadership across their subjects using descriptive statistics (e.g., Augustine, 1998; Yogere, 1996). They employed measures of central tendency and standard deviations to describe these patterns on the different dimensions and sub-scales of instructional leadership measured by the PIMRS. These were the simplest set of studies though sometimes the use of descriptive statistics was complemented by qualitative data obtained from interviews or observation.

The next approach to statistical analysis involved the use of bivariate statistics, usually without controls for variables beyond the variables in questions such as gender and instructional leadership. These tests included t-tests, U-tests, correlation procedures such as Pearson’s or
Wilcoxen’s tests, Anova, and bivariate regression analysis. This approach to analysis was employed frequently in both Model A and Model B studies.

For Model A studies, the most common variables tested for their relationship to instructional leadership using these statistical tests were school level and size, principal gender, years of experience, and teaching experience. In Model B studies, these tests were used to examine the relationship between principal instructional leadership and a number of school-level variables assumed to be influenced by instructional leadership including teacher effectiveness, school effectiveness, teacher morale, teacher self-efficacy, and school climate. Tests such as t-tests were also used to test for differences in the perceptions across role groups, another common research question.

The key limitation of these statistical tests is that they, at best, provide insight into associations between variables. They do not, however, offer deeper insight in the nature of the relationship or the causal basis for relationships. This is particularly troubling in social science research such as this where assumptions of causality are often made with insufficient basis. For example, a number of studies sought to examine the relationship between instructional leadership and student achievement through the use of t-tests (Buzek, 2004; Neuwenswander, 1986). Neither the test nor the design would allow meet assumptions of causality regardless of the level of significance of the result.

The third set of statistical approaches employed to analyze these models involved the used of inferential tests with controls. This includes tests such as multiple regression (e.g., Geiselman, 2004; Jones, 1987; Leitner, 1989; Lubbers, 1996; Maciel, 2005; Parker, 1990; Prater, 2004; Sheppard, 1993; Van Pelt, 1993), factor analysis (Edwards, 2006; Garcia, 1999; Lord, 2001; Mallory 2002) and discriminant analysis (Delano, 1985; Griffin, 1993; Werner, 1991). These tests allow researchers to test for relationships among several variables. Thus, they
are better able to illuminate relationships between instructional leadership and other relevant school factors.

Inferential statistics have become the minimum acceptable level of testing for many of the most important research questions of interest in this domain of research in school administration (Hallinger & Heck, 1996a, 1996b; Heck & Hallinger, 1999). For the most part, they were selected correctly for these studies because neither Model A nor Model B makes great demands on the statistical tests. The problem with these studies, as a group, is not with their application of statistics, but with the limitations of the theoretical models that they employed. As Hallinger and Heck noted in earlier reviews of research on principal leadership (1996a, 1996b; Heck & Hallinger, 1999), advances in this domain of research will depend on the concurrent application of more powerful theoretical models (e.g., Models C and D) with robust statistical methods.

This brings us to the final set of statistical approaches that allow for the modeling of direct and indirect effects among variables. Tests such as hierarchical linear modeling, path modeling, and structural equation modeling allow the researcher to accomplish this important goal. Only a few studies used such tests (i.e., hierarchical linear regression). This creates severe limitations on the extent to which these studies will advance the field today and in the future. Moreover, it should be noted that quite a few studies could have employed such techniques. This characteristic of the dissertations will be remarked upon later as it may indicate the need for new standards in doctoral dissertation research.

Findings on the reliability and validity of the PIMRS. Earlier it was noted that the PIMRS was validated in an initial study (Hallinger, 1982, 1983). While relatively few researchers using the instrument sough to replicate the initial findings, several did. The replication studies of reliability and validity included Howe (1995), Jones, (187), Nogay (1995), Sawyer (1997), Taraseina (1993), Wotany (1999).
Reliability coefficients seldom dropped below .80 in any of these studies. Although some questions remain concerning validation of all subscales at the secondary level (e.g., Jones, 1987), the findings across studies are surprisingly consistent. The PIMIRS is yielding valid and reliable data for the purposes of research.

**Topics and Findings**

Although this review focuses on methods rather than results, some mention of the trend of results is warranted as well. Three types of studies define the most popular topics among these studies of instructional leadership: role group perception studies, antecedent studies, and effects studies.

*Role group studies.* As noted above, there are three forms of the *PIMRS*. The initial validation study suggested differences in the perceptions of teachers, principals and supervisors concerning the principal’s instructional leadership. The original validation study conducted for the *PIMRS* concluded that teacher perceptions yielded the highest degree of validity (Hallinger, 1983).

Numerous role group perception studies have been undertaken with the *PIMRS* in order to determine the degree of agreement among teachers, principals and supervisors concerning the principal’s instructional leadership. Subsequent research continues to find, statistically significant disagreement across role groups in their perceptions of the principal’s instructional leadership (e.g., Brown, J., 1991; Haack, 1991; Haasl, 1989; Krug, 1986; O’day, 1984; Vinson, 1997). It is interesting to note that this finding was confirmed 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 some exceptions (e.g., Duryea, 1988; Krug, 1986; Marshall, 2005; Rogers, 2005), principals self-reports yield higher ratings than reports from their teachers and the findings are statistically significant.
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.*

**Antecedents of leadership: Principal demographic variables.** The second type of model is referred to in Figure Two as Antecedent Effects Studies, of which there are two categories of antecedents: demographic characteristics of the principal (and sometimes the teacher respondents) and the school context. The single 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 principal gender, years of experience as a principal, years of teaching experience prior to becoming a principal, age, and ethnicity. Although not strictly a *demographic* variable, several investigators also studied the relationship between the principal’s knowledge of instruction and their instructional leadership.

It is not the purpose of this review to review the substantive findings of the studies in depth. Nonetheless, it is worth noting general trends in the findings. The first finding of interest among the demographic variables concerns the relationship between gender and instructional leadership. The initial PIMRS study (Hallinger, 1983) found that gender (female) appeared to be associated with higher levels of instructional leadership activity among the sample of principals. This finding supported earlier work by Gross and Herriott (1976) and led to the recommendation that future researchers focus more attention on this issue. That call was answered and 16 completed studies within this set used the PIMRS to study the relationship between gender and principal instructional leadership.

The finding of more active instructional leadership among female principals is unusual in its consistency. For example, this finding held across school levels (Babcock, 1991;
Cunningham, 2004; Duryea, 1988; Gallon, 1998; Hallinger, 1983; Howell, 1989; Lehl, 1989; McCrier, 2004; McTaggart, 1991; Meek, 1999; Nogay, 1995; Parker, 1990; Poovatanakul, 1993; Pratley, 1992; Prater, 2004; Ryans, 1989; Schoch, 1992; Sterret, 2005; Wells, 1993) and even in some non-Western cultures such as Thailand and Taiwan. Although some studies reported no effects of gender on instructional leadership (e.g., Geiselman, 2004; O’Donnell, 2002; Rogers, 2005; Saavedra, 1987; Trout, 1985; Yang, 1996), these were exceptions to the more general trend.

Several other demographic variables were also studied fairly extensively including age, experience as an administrator, prior teaching experience and ethnicity. Among these variables only years of experience as an administrator and prior teaching experience yielded any significant findings.

Several researchers who studied the impact of years of experience as an administrator found positive findings (Adkins, 1990; Benoit, 1990; Brown, R. 1991; McTaggart, 1991; Parker 1990; Rogers, 2005; Singleton, 2006; Sterrett, 2005). Meek (1999) reported mixed results. It is interesting to note that several studies, including four conducted in Asia, all reported either no relationship or a negative correlation between administrative experience and instructional leadership (Brown, R., 1991; Chan, 1992; Howell, 1989; Poovatanakul, 1993; Saavedra, 1987; Yang, 1996). Burwell reported an interaction between gender and experience which suggested that more experienced female principals exercised the most active instructional leadership.

A number of researchers also studied the relationship between variables that might act as a proxy for instructional expertise. These included assignment at the level at which the principal had taught, prior years of teaching experience, and knowledge of instruction (measured by an independent test of knowledge). Although several found positive associations (Delano, 1985; Orange 1990; Pratley 1992), other studies did not confirm these findings (Lehl, 1989; Simpson, 1988).
Despite the number of studies conducted, these findings concerning the personal antecedents of instructional leadership do little to illuminate the characteristics of instructional leaders. Even the most clear-cut conclusion that female principals, on average, exercise more active instructional leadership than males is of only marginal use. Research in this area must uncover the avenues through which female principals exercise this leadership. In fact, these studies begin to point the direction. A number of them concluded that it was through the mission setting and instructional program domains that the largest differences were found. We will return to this issue again later.

**Antecedents of leadership: School context variables.** Among school context variables, several stand out as particularly frequent targets for study:

- **School level of the principal,** (Duryea, 1998; Gallon, 1998; Garcia, 1998; Hart, 2006)
- **School performance rating** (e.g., Gerrell, 2006; Gibson, 2005; Johnson, 2005; Moore, 2003; Waters, 2005)
- **Private schools** (Chi, 1997; Griffin, 1993; Hart, 2006; Howe, 1995).

Other context factors include the effects of a school-based management context, comparison of instructional leaders in public and private schools, comparison of instructional leaders in public and Catholic schools, different district contexts, and district size. These did not yield significant results in any consistent fashion across studies.

**Leadership effects studies.** Among the leadership effects studies, there are two broad types: studies of intervening or mediating variables within the school and studies that linked
instructional leadership to school outcomes. Intervening variables of interest among these researchers included:

- Teacher morale and satisfaction (Apolonia, 1998; MacNeil, 1992; ),
- Teacher self-efficacy (Chan, 1992; Keith, 1989; Lubbers, 1996; Ruzicska, 1988),
- Teacher stress (Courtney, 1994),
- Principal locus of control (e.g., Duryea, 1988),
- School and organizational culture (Howe, 1989; Leitner, 1989; Reid, 1987),
- Teacher effectiveness and time on task (e.g., Kennedy, 1993; Reid, 1987; Watkins, 1992),
- Organizational climate or health (Adkins, 1990; Lord, 2001; Salvador, 1999; Sheppard, 1993; Simpson, 1990; Skiptunas, 1990; Wilson, 2005),
- Teacher participation in decision-making (Sicina, 1996).

Others have studied the relationship between instructional leadership and a variety of school outcomes. These outcomes have focused mostly on student achievement, though measured in different ways. The first subset of leadership effects studies examined student achievement by linking the PIMRS measurement of principals directly to student test scores (Balsamo, 2004; Campbell, 1998; Corkill, 1994; Dilworth, 1997; Geiselman, 2004; Jones, 1987; Krug, 1986; Leitner, 1989; Maciel, 2005; Meek, 1999; Neuwenswander, 1986; O’Day, 1984; Ruzicska, 1987; Singleton, 2006; Stroud, 1989). The results of this approach were generally disappointing with researchers reporting infrequent positive results and generally inconsistent results across studies.

A second approach to studying “whether principals make a difference” was through studying the association between instructional leadership and school effectiveness. In most cases, researchers using this approach linked measurements on the PIMRS of principals in two or more groups of schools that contrasted on achievement (Cantu, 1994; Dickerson, 1999; Garcia, 1999; Hunter, 1994; Lubbers, 1996; Moore, 2003; Orange, 1990; Rose, 1991; Stroud, 1989; Zeanah, 1986). Consistent with the school effectiveness movement, most of these
researchers sought to control for student background. While there were a few findings of positive effects of instructional leadership on school effectiveness, the general trend of results did not support this conclusion.

The third approach to studying the link between instructional leadership and school success was done by again through a comparative groups design. In this case, measurements on the PIMRS of principals were linked to two or more groups of schools that contrasted on a common standard of success such as national or state recognition (Brown, R., 1991; Werner, 1991). These studies found small differences on certain subscales of the PIMRS in favor of the principals in the successful schools. However, the differences are not conclusive.

To place these findings in perspective, it is necessary to revisit the methodological sections of this review. There it was noted that almost all of the studies here fell into the Model A and B categories. Both the theoretical models and the statistical tests employed in most of these studies were inadequate to the task of explaining causal relationships such as the relationship between instructional leadership and school effectiveness or school improvement. To make progress on this front will require studies built on Models C and D and which use Advanced Inferential statistics.

This point was made in the Hallinger and Heck (1996a, 1996b) reviews of principal leadership effects. In that review it was clearly shown that the effects of principal instructional leadership are indirect, not direct. It requires a more sophisticated model of the paths through which principals create more effective schools in order to reveal these types of effects. Therefore, the results of the studies are hardly surprising.

Summary and Conclusion

In brief the results of this review suggest the following:

1. The topic of instructional leadership remains an active topic of investigation among doctoral researchers.
2. The PIMRS has proven a reliable means of collecting data on principal instructional leadership for both the elementary and secondary levels; Differences in the role of the principal at the secondary level, however, lead to less robust results with respect to validity than are found at the elementary level;

3. Teacher perceptions continue to constitute the preferred source of data on the principal’s instructional leadership for both research and evaluation purposes when using the PIMRS.

4. Research on instructional leadership using the PIMRS provides surprisingly consistent evidence that female principals at all levels exercise more active instructional leadership than their male counterparts. The strength and consistency of this finding are unusual in the social sciences and are, therefore, noteworthy. Nonetheless, researchers studying this issue will need to use more creative and robust designs in order to further our understanding of this finding.

5. With respect to school achievement outcomes, the findings from these studies are very clearly inconclusive. Although 36 studies examined the relationship between instructional leadership and either school effectiveness or student achievement, the methodological limitations of most of these studies would have hidden the indirect effects even if they existed. The sophistication of most of the studies was simply inadequate to the task they undertook. Researchers are recommended to examine high quality studies that have examined this relationship in the past in order to construct higher quality studies in the future (e.g., Cheng, 1991, 1994; Hallinger, Bickman & Davis, 1996; Heck, Larsen, & Marcoulides, 1990; Leithwood, 1994; Leithwood & Jantzi, 2000; Marks & Printy, 2003; Mulford & Silins, 2003; Silins, 1994; Wiley 2001).
6. Although the *PIMRS* appears to provide reliable and valid data on principal instructional leadership, criticisms lodged by Bridges in 1982 continue to impede broader progress in understanding this approach to principal leadership.

- The choices of topics for study focus too much on personal traits and their relationship to principal leadership;
- Too few researchers embed the conceptual framework implied in the *PIMRS* within a broader theoretical framework of antecedent or outcome effects.
- With the exception of a relatively small percentage of studies the theoretical models and statistical tests used to explore instructional leadership do not reflect more general advances in the field. Not coincidentally, the exceptions were studies that also employed more sophisticated theoretical frameworks.
- Finally, it is a clear trend in the field that the principal’s role as instructional is increasingly viewed as one to be shared with others (Barth, 1990; Blasé & Blasé, 1998; Kleine-Kracht, 1993; Lambert, 1998, 2002; Gronn, 2000, 2002)

During the 1980’s Larry Cuban published a series of critiques that called into question the validity of an all encompassing focus on the principal’s role as an instructional leader (e.g., Cuban, 1988). He questioned how long the fascination with this role would last given the difficulty of enacting it within real school contexts. Despite its noted limitations, the body of research reviewed in this study is substantial in scope, depth and longevity. While it provides few answers to the questions that surround the principal’s instructional leadership role, it does provide a useful base from which to build.


Bio Note

Philip Hallinger is Professor and Executive Director in the College of Management at Mahidol University, Thailand. He specializes in the study of leadership, leadership development, organizational change, and school improvement. He can be reached at Philip@leadingware.com.