

Chapter appearing in K. Leithwood et al. (1996; eds.),
*The International Handbook of Research on
Educational Leadership and Administration*,
New York: Kluwer Press.

The Principal's Role in School Effectiveness:
An Assessment of Methodological Progress, 1980-1995

Philip Hallinger

Vanderbilt University

Chiang Mai University

and

Ronald H. Heck

University of Hawaii, Manoa

There is relatively little disagreement in either lay or professional circles concerning the belief that principals play a critical role in the lives of teachers, students and schools. This belief has led to considerable research into the nature of principals' work, attitudes, values, thought processes, and behavior (Leithwood, Begley & Cousins, 1990, 1992). When consulting the empirical literature, however, both the nature and degree of principal impact continue to be subject to debate (e.g., Pitner, 1988; Rowan, Dwyer & Bossert, 1982; van de Grift, 1990).

Over the past 15 years several substantive reviews of this research literature have been conducted with the aim of consolidating our understanding of the principal's role in schooling (e.g., Boyan, 1988; Bossert, Dwyer, Rowan, & Lee, 1982; Bridges, 1982; Leithwood & Montgomery, 1982; Leithwood et al., 1990; Murphy, 1988; Murphy, Hallinger & Mitman, 1983; Pitner, 1988). These reviews consistently paint a picture whose broad strokes seem quite clear from afar, but which becomes much fuzzier when viewed up close. While most of these scholars agree on the importance of the principal's leadership, we still lack many details concerning how principals respond to their schools' environmental contexts as they seek to shape organizational processes and student outcomes. Reviewers have concluded that the tradition of research on the impact of principals has not generally done justice to the complexity of the topic in terms of either theoretical or methodological sophistication (Bossert et al., 1982; Bridges, 1982; Murphy, 1988; Rowan, et al., 1982).

For example, following a review that focused focusing primarily on methodological features of research of educational administrators, Bridges (1982) asserted:

Research on school administrators for the period 1967-1980 reminds one of the dictum: "The more things change, the more they remain the same" . . . 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. . . . Despite the rather loose definition of theory that was used in classifying the sample of research. . . , most of it proved to be atheoretical. Likewise the research seemed to have little or no practical utility. (pp. 24-25)

While it was not the specific focus of his review, Bridges (1982) further noted that research on administrator impact was both consistent with the above characterization and plagued by additional problems as well.

The lack of integration of theory with procedures of scientific inquiry in studying the consequences of school administration is especially disheartening. We must have confidence that researchers have accurately described and explained the phenomenon under consideration if this research is to be useful in informing policy efforts to improve schooling. Methodology concerns the process by which we construct knowledge. As such analyses of methodology must address how data are collected, analyzed, and interpreted, as well as the theoretical and technical justification for these procedures (Everhart, 1988; Kaplan, 1964). Therefore, we must attend to the underlying assumptions of any scientific approach, as well as to its strengths and weaknesses (see Everhart (1988) for a thorough discussion of the relationship between methodology, method, and techniques of scientific inquiry in educational administration).

Over the past 15 years researchers in educational administration have increased their attention to the study of the principal's leadership role and its impact on a variety of school processes and outcomes. This was stimulated in part by parallel research into processes of school improvement and effectiveness. These efforts have been coupled with analytical innovations such as structural equation modeling. In combination, they have yielded a new generation of research on principal effectiveness.

Findings from this research that have found their way into scholarly and professional publications tend to give the impression that principals make a difference in student learning (e.g., Andrews & Soder, 1987; Bamberg & Andrews, 1990; Cheng, 1994; Eberts & Stone, 1988; Heck et al., 1990). In our view, however, the interpretation of substantive findings from a body

of literature must be considered in light of conceptual and methodological underpinnings. Both are crucial to determining the extent to which findings from research can be accepted as valid.

This chapter reviews research that explores the relationship between the leadership of the school principal and school outcomes concerned with student learning. The period of review extends from 1980 up to 1995. Our lens focuses more on the interplay between methodological and conceptual features of this body of research than upon the substance of the actual findings. Our aim is assess the contribution to knowledge made by these studies by understanding more clearly both the theoretical frameworks and processes of scientific inquiry used to generate the results. The purpose of the chapter, therefore, is threefold:

- 1) to analyze the theoretical frameworks that have been employed by researchers in the
- 2) to examine the methodological features of this literature;
- 3) to propose a theoretical framework and appropriate methodological approaches that

We begin by discussing the perspective for this review and issues that concern policymakers and researchers with respect to the impact of administrative leadership in schools. We briefly examine how this emerging concern was addressed by researchers during the 1980's and the resulting influence on our thinking about this review. Next, we consider conceptual and methodological issues as they emerged in our analysis of the studies. This analysis rests on the assumption that the way the research problem is conceptualized and the means of studying it both lead to what is observed and how those observations are interpreted. The chapter concludes with an attempt to frame an agenda for research on the principal's role in school effectiveness for the next generation of studies.

The Perspective and Selection of Studies for this Review

Any attempt to integrate a body of research into a coherent framework that summarizes both conceptual and methodological issues must begin with an acknowledgement of its limitations. First, the field's conceptualization of organizational processes, including the school leadership construct, is constantly evolving (Glasman & Heck, 1992; Hallinger,1992, Leithwood &

Hallinger, 1993; Leithwood et al., 1992). Hence, we assert that there is no universal paradigm or theory for examining organizational behavior that is valid in all contexts. This point is especially salient for the current review since we include studies conducted internationally.

Second, proposed theories often become problematic when they seek to model the actual detail and richness of life in organizations (Bossert et al., 1982; Hallinger & Murphy, 1986b). The complexity of extra- and intra-organizational processes represents a challenge for researchers who seek to study causal relationships (Boyan, 1988; Marcoulides & Heck, 1993; Pitner, 1988). Social scientists who seek to develop valid theoretical models and apply appropriate analytic techniques to assess how those models work in the empirical world confront a formidable set of tasks.

We began this review with the assumption that the number of studies to be included would not be so large as to require a sampling strategy (Bridges, 1982). At the same time, we sought to conduct a highly inclusive review. We first searched the ERIC (Resources in Education) and Current Journals in Education (CJIE) databases and used the resulting sources as well as our personal knowledge of published and presented research to identify additional studies.

Three criteria guided our selection of studies for review. First, we were interested in studies that had been designed explicitly to examine the school principal's beliefs and leadership behavior. The research must have clearly conceptualized and measured principal leadership as one of the independent variables. While our assumption was that most studies would use some measure of instructionally-oriented leadership, we observed that the definition of principal leadership has changed considerably over the decade and half of our review.

Second, the studies also had to include an explicit measure of school performance as a dependent variable. Most often performance was measured in terms of student achievement data, but occasionally other definitions such as effectiveness were also used. It was our desire, though not a necessary condition for inclusion, to also identify studies that examined the principal's impact on teacher and school level variables as mediating factors. The dual focus reflects the priority that we assign to student outcomes as the goal for school improvement, since we assume

that an understanding of principal impact on student outcomes must also account for the operation of classroom and school-level variables. Notably, however, we did not include studies that examined principal impact on intervening variables if they did not also incorporate a measure of school outcomes. This criterion shifted the focus of the review towards quantitative studies of impact, as opposed to studies about the nature of the principal's work.

Third, given both the focus of the Handbook and the growing interest in international perspectives on school improvement, we made an extra effort to seek out studies that examined the impact of principals conducted in a variety of countries. We were reasonably successful in attaining this goal. Although we do not undertake comparative analysis in this chapter, we have included studies conducted in a diverse set of cultural contexts including the United States, Canada, Singapore, England, Netherlands, Marshall Islands, Israel, and Hong Kong. Eleven of the studies reviewed were conducted outside of the United States.

Consequently, the review includes published journal articles, dissertation studies, and papers presented at peer-reviewed conferences. We are reasonably confident that the chapter has captured most empirical studies of principal impact on school effectiveness disseminated internationally between 1980 and 1995. We owe particular debts to the earlier efforts of Bossert and colleagues (1982), Boyan (1988), Leithwood and colleagues (1990), and Pitner (1988) for laying the groundwork for this review.

The Principal's Role in School Effectiveness: Methodological and Conceptual Issues

Using these criteria, we identified 40 studies that explored the relationship between principal leadership and school outcomes or effectiveness conducted during this time period. Twenty-two of the studies were published in blind-refereed journals. Eleven were presented as papers at peer reviewed conferences (primarily the annual meeting of the AERA). Five were doctoral dissertations. One was a book chapter and one was a synthesis of several studies conducted by the author. Of the studies identified, we were unable to obtain two papers presented at

professional meetings (Edington & Benedetto, 1984; Teddlie, Falkowski, Stringfield, Desselle, & Garvue, 1983).

With this overview in mind, we assess the conceptual and methodological trends that emerged from these 40 studies as a group. We content analyzed the studies using a classification scheme suggested by Pitner (1988) (described in further detail later in the chapter). Working independently, we classified each study as one of six model types (see Figure 1). After comparing our completed schemes, we resolved the few discrepancies and triangulated our results with previous reviews (e.g., Bridges, 1982; Pitner, 1988).

In this section of the chapter we discuss the major conceptual and methodological features of these studies. First, we briefly overview the philosophical and methodological perspectives of these studies. Then we look more specifically at the range of theoretical models that have been proposed for the study of the principal's role in school effectiveness and the corresponding operationalization of the variables proposed in the theoretical models. Next, we look more closely at a variety of design issues: the nature of the samples chosen, units of analysis, data collection methods, the analytical techniques employed to test the various models. Finally, we address the construct validity of the leadership models presented and generalizability of the knowledge generated from this set of studies.

Frame of Reference

In 1982 Bridges concluded that the frame of reference for studies in educational administration tended to be neither theoretical nor practical in nature. He found meager evidence of any systematic attack on problems of practice. Similarly, there appeared to be little accumulation of knowledge aimed at building theory. This was a bleak conclusion indeed for a field of professional practice.

This next generation of research is but subset of the literature reviewed by Bridges. Yet, it suggests considerable improvement. Perhaps because of the present studies' common lineage from the effective schools literature (Hallinger & Heck, in press), they delineate similar themes concerning the principal's role in promoting school improvement. Early studies of effective

schools (e.g., Brookover et al., 1978; Edmonds, 1979) tended to view organizations primarily from a technical/rational perspective (Bolman & Deal, 1992; Ogawa, 1992). Scholars implied that changes could be made in schools by implementing effective schools correlates such as high expectations and strong school leadership. Prolonged positive effects of such change interventions, however, have been difficult to find or to generalize across educational settings. Nonetheless, the studies target an important area of research in educational administration in a positive fashion.

As Everhart (1988) argues, however, research approaches (e.g., purposes, questions, methods) must be seen within the historical and social context in which they are formulated; the answers created are equally contextualized. Thus, it is important to note that as a group the studies included in this review are decidedly functionalist and positivistic in their philosophical tradition. This undoubtedly resulted in part from our very framing of the question that underlies the review -- how do principals make a difference in the learning of students?

What we observe in organizations is socially constructed and often reinforces, or at least reflects, dominant social and political ideologies (e.g., Anderson, 1990; Benham & Heck, 1994). While it is true that scholars have paid greater attention to alternative philosophical frameworks over the past decade (e.g., critical theorist, feminist, Marxist, phenomenological, post-modernist perspectives), relatively little of this discourse seems to have worked its way into empirical research on administrative effectiveness [see *Educational Administration Quarterly*, 27(3) for a discussion of several different frameworks for research]. The lack of diverse philosophical and methodological views manifested in these studies is notable. The implications of the rather limited philosophical perspective and methodological approach taken within "effective schools" studies, including those reviewed here are, therefore, considerable.

The examination of methodology itself forces one to admit that any particular approach is unlikely to yield universal understanding (Eisner, 1993; Everhart, 1988). From a methodological perspective, we chose to frame our review in terms of administrator effects on school effectiveness (as opposed to, for example, descriptions of principal work). Consequently, our

selection criteria weighed heavily towards identifying finding quantitative studies. In contrast, a different approach to conceptualizing effectiveness might have pointed us towards naturalistic inquiry (e.g., fieldwork, ethnography) which tends to be constructivist, holistic, and process oriented (Everhart, 1988). In our view, quantitative methods are essential for the first part of this research program -- assessing the extent to which administrative effects seem to be present. The use of qualitative approaches, however, is also essential if we are to understand the more complex processes that underlie this set of observed interactions (Dwyer et al.,1983).

Research Design

Almost all of the studies identified in our search used some form of cross-sectional, correlational design, often employing surveys or interviews as methods of gathering information. Studies of this type have been labeled under the broad design type of "non-experimental" research (Pedhazur & Schmelkin, 1991). In this research approach, the independent variables are not manipulated as they are in experimental or quasi-experimental designs. Also, as in quasi-experimental designs the subjects are not assigned to groups through randomization. None of the studies in this review were classified as experimental or quasi-experimental in design.

Some may view this as problematic. Experimental and quasi-experimental designs employ stronger procedures for controlling extraneous variables that can confound measurement of the effects of the independent variable. Thus, such designs are often better suited to determinations of causation.

At the same time, the usefulness and feasibility of conducting experiments in the social sciences, often under relatively isolated conditions, has been debated with no clear conclusion(Pedhauzer & Schmelkin, 1991). Experimental research comes with its own set of problems and design limitations, particularly when applied to relatively large sample of schools. Thus, as we shall elaborate throughout this chapter, the over-riding issue is not design type per se, but the extent to which the chosen research design is guided by a strong theoretical model and the data are analyzed using appropriate methods.

As has been discussed elsewhere, the use of non-experimental designs to study causal relationships is a dauntingly complex task (Pedhazur & Schmelkin, 1991; Pitner, 1988). A point of departure in understanding distinctions among the three broad classes of designs is the role of theory to help specify the models for testing. In experimental and quasi-experimental designs causal inferences are made from the independent variables to the dependent variables. In contrast, in non-experimental designs causal inferences are generally made in the opposite direction (Pedhazur & Schmelkin, 1991). In non-experimental research attempts are made to account for a dependent variable by "uncovering" relevant independent variables. Appropriate model specification readily becomes a problem because one must recognize the need to include all relevant independent variables to specify the model properly.

Thus, the major threat to validity in non-experimental research stems from uncontrolled confounding variables. Major approaches to control include subject selection, statistical adjustments (Pedhazur & Schmelkin, 1991), and replication of results through varying conditions (Heck & Marcoulides, 1992). These are critical points to keep in mind when assessing the contribution to knowledge made by this set of studies.

We suggest two further points to consider in judging the contribution of the studies in this review. First, as implied above, in non-experimental research the complexity of relationships explicated in the tested model play an important role in interpreting the results. Empirical research grounded in overly simplistic conceptualizations of leadership effects is unlikely to yield results that are useful, practically or theoretically.

Second, the appropriateness of analytical techniques used by the researcher affect the strength of the conclusions that can be drawn about the effects of the principal's role. Certainly more rigorous analyses may lead to uncovering relationships in the data that are not revealed in more simplistic analyses. At the same time, however, they are also more likely to lead to fewer findings of substance than have often been "claimed" in studies that employ more simplistic analytical methods (Pedhazur & Schmelkin, 1991).

Classification of Administrator Effects

While granting that there exists a rather narrow philosophical perspective undergirding studies of school and administrator effects, there is still considerable variation in the conceptual modeling of leadership effects. Because the group of studies is essentially confined to one philosophical stance, we found it convenient to apply Pitner's (1988) framework of administrator effects in classifying the studies. In an earlier analytical review, Pitner (1988) sought to conceptualize the possible theoretical approaches that could be used in studying administrator impact through non-experimental research designs. She identified five theoretical approaches to portraying administrator effects: direct-effects, moderated-effects, antecedent-effects, mediated-effects, and reciprocal-effects models (pp. 105-108). These models offer one means of viewing both the impact of the school context on administrative behavior and the influence of administrative behavior on the school organization and its outcomes. Our adaptation of Pitner's classification system of administrative effects is depicted in Figure 1.

Insert Figure 1 About Here

A *direct-effects model* of administrator effects (Model A, Figure 1) proposes that the principal's leadership actually exerts an influence on students apart from other variables within the school (e.g., teacher behavior, curricular organization, school culture). The researcher using this model does not normally seek to account or control for the effects of these in-school mediating variables. The researcher may or may not control for the impact of external variables.ⁱ

In the past, studies employing a direct-effects model were, therefore, typically bivariate in nature (e.g., O'Day, 1983; Ruczieska, 1988; van de Grift, 1987, 1989ⁱⁱ, 1990; see Table 1). More recently, however, sophisticated analytic techniques such as structural equation modeling have also been employed to assess the relationship between leadership and school outcomes (e.g., Hallinger et al., 1990; Heck et al., 1990; van de Grift, 1990; Weil, Marshalek, Mitman, Murphy,

Hallinger, & Pruyne, 1984). Two direct-effects studies (i.e., Ramey, 1982; van de Grift, 1990) used structural equation modeling.

While direct-effects studies are quite common in the literature, they have been criticized for making untenable assumptions about the nature of leadership in organizations (Murphy, 1988; Rowan et al., 1982). Moreover, in such studies the process by which administrators achieve an impact is hidden by a "black box." Consequently, these studies reveal little that advances our theoretical or practical understanding of the critical school processes through which the principal achieves an impact on school effectiveness (Leithwood et al., 1990; Leitner, 1994).

Another conceptual framework for examining administrator effects is represented by the *moderated-effects* model (Model B, Figure 1). Here it is proposed that some condition in the school or the environment operates as a third variable *moderating* the relationship between leadership and school outcomes through its presence or absence. One begins by noting a relationship (e.g., a correlation) between the administrator and outcome of interest. A third, usually dichotomous, variable is then added to the model and its effects on the original relationship assessed. Typically, researchers using a moderated-effects model would hypothesize that the administrator effect occurs under one set of conditions but not under another, less favorable condition (Pitner, 1988).

This approach was not widely used by researchers in this set of studies. We would, however, note that several studies in our review did identify third variables that appeared to act as moderating forces on the relationship between principal leadership and school outcomes. The most frequently studied was community socioeconomic factors (e.g., high or low school socioeconomic status). This environmental variable was found to specify the type of leadership principals exercise in an effort to improve school outcomes (i.e., Andrews & Soder, 1987; Hallinger & Murphy, 1986; Rowan & Denk, 1984). This result has implications for the moderated-effects approach, but was derived from studies using the antecedent-effects model.

A third perspective for viewing administrator effects identified by Pitner (1988) is the *antecedent-effects* model. Conceptually this model is more complex than the previous

conceptualization. Unlike the moderated-effects model where the administrator is considered as an independent variable, in antecedent-effects research "the administrator variable stands as both a dependent and an independent variable" (Pitner, 1988, p. 106). As a dependent variable, administrative behavior is subject to the influence of other variables within the school and its environment. As an independent variable, the administrator is an agent who acts to influence the actions of teachers, the nature of the school organization and the learning of pupils (Bridges, 1970, 1977; Hallinger & Murphy, 1986a; Leithwood et al., 1990; Rowan et al., 1982).

When actually categorizing studies by model type, we found the distinction between moderated effects and antecedent effects often difficult to ascertain. In practice the distinction depends on several factors. These include the specific theoretical model proposed before the data are analyzed (sometimes not specified by researchers), the level at which the relevant variables are measured (i.e., often continuous for antecedent-effects and categorical for moderated-effects models), the techniques used to analyze moderated (i.e., correlations, t-tests and analysis of variance) and antecedent (i.e., structural equation modeling or regression with interaction terms) effects, and the findings as interpreted in the published study. Thus, both conceptual and empirical issues influence the choice of model.

As indicated, the two conceptual models imply different approaches to data analysis. For example, the researcher who conceives of socioeconomic status as exerting a *moderating* effect might build data tables to portray and compare the relationship between leadership and outcomes in light of the moderating variable. The first table would analyze the relationship between leadership and outcomes in schools with high socioeconomic status. The other would show the relationship between leadership and outcomes in schools with low socioeconomic status (e.g., Andrews & Soder, 1987; Andrews, Soder & Jacoby, 1986; Hallinger & Murphy, 1986b). The researcher then looks at both tables to determine how the moderating variable affects the original relationship. Although this approach to analyzing moderating effects most clearly demonstrates the concept, it should be noted that other analytical approaches can be used as well (Rowan & Denk, 1984).

If we conceive of the antecedent variable(s) (e.g., socioeconomic status) as exerting both a direct effect on school outcomes as well as an indirect effect through intervening variables (e.g., on classroom composition, teacher behavior, school leadership), an antecedent-effects model may be preferable. This approach is considerably more flexible in that it allows the researchers to include multiple antecedents in the model and to analyze the whole constellation of relationships simultaneously rather than separately.

In fact, only a few analytical techniques are able to measure the indirect effects implied in antecedent-effect models. As we shall elaborate later in the chapter, path analysis and structural equation modeling are the preferred methods for this type of model. There are, selected instances where such relationships can be implied using interaction terms associated with analytical techniques such as analysis of variance or covariance and regression analysis.

An important conceptual point to keep in mind in differentiating these two models is the relative importance attached to the third (moderating or antecedent) variable. In moderated-effects studies this variable remains of secondary concern to leadership. In antecedent-effects studies this external variable is conceived to play a key role in understanding the system of relationships among the variables.

An example serves to illustrate this distinction. Andrews and Soder (1987) tested the direct effects of leadership on school outcomes. They concluded that a significant relationship existed between principal leadership and student outcomes across all schools in their sample for reading and math improvement. Then they introduced school socioeconomic status, conducting separate analyses for high and low SES schools. All significant relationships between leadership and math and reading outcomes disappeared for high SES schools, but some remained for low SES schools. The same general pattern also held when the sample was dichotomized by predominant ethnicity. They were unable to consider the possible interactions between ethnicity and SES, or between these variables and leadership, in their analysis, however, because of the small sample size.

In this study, therefore, the more complete conclusion is that socioeconomic status and ethnicity *moderate* the effects of leadership on school outcomes. Within this model, the direct and indirect effects of environmental conditions could not be optimally tested on both leadership and outcomes simultaneously (as implied in an antecedent-effect model). Therefore, linkages between the antecedent contextual variables and leadership and their resultant combined impacts on school outcomes are hinted at but left unclear. The study illustrates the distinction between the models and the limitations of the moderated-effects approach.

A *mediated-effects* model (Model C, Figure 1) assumes that some or all of the impact attained by administrators on desired school outcomes occurs through manipulation of, or interaction with features of the school organization (e.g., Biester et al., 1984; Crawford, Kimball, & Watson & 1985; Eberts & Stone, 1988; Jackson, 1982; Silins, 1994; see Table 1). This is consistent with Bridges' (1977) contention that managers achieve their results through other people. Mediated-effects studies, therefore, are more useful for theory building than direct-effects studies. They are also of potentially greater value for revealing avenues by which administrators achieve practical results.

A fifth approach Pitner proposed is the *reciprocal-effects* model. Scholars have noted that it is not necessarily the case that the effects of administrative move in only one direction. It is possible, even likely, that the relationship between the administrator and features of the school are interactive. Administrators may cause changes which then cause reciprocal effects in the originating variable (i.e., leadership). Alternatively, it could be hypothesized that the administrator adjusts his or her leadership behavior to processes and characteristics of the school (e.g., current and changing states of student outcomes).

Where doubts about the direction of causality are expressed, cross-sectional data are unable to resolve the ambiguity inherent in correlations and other measures of association (Davies, 1994). As Davies (1994) argues, "the duration in current state" data often collected in cross-sectional studies are not sufficient to overcome this type of problem. Moreover, with cross-sectional data, one cannot characterize the inertial properties of the assumed reciprocal

relationship. A more complete representation, therefore, would require the assumption that the reciprocal effects will only become apparent over time. Thus, in order to specify such models properly, longitudinal data are preferred.

Notably, only two studies tested for reciprocal effects and these are actually more properly conceived of as C_1 studies (Hallinger et al., 1990; Heck et al., 1990). Yet, the approach holds much promise for future investigations that seek to employ dynamic models of principal effects.

We contend that a comprehensive framework for viewing the principal's role in school effectiveness must locate principal leadership within both organizational and environmental contexts. To adapt this view to the framework proposed by Pitner, the effects of the principal on the organization can be modeled by combining antecedent variables with either direct or mediated-effects models (see Figure 1 and Table 1). When antecedent variables are incorporated into the direct-effects model, we will refer to this new model as Model A_1 (e.g., Blank, 1987; Dilworth, 1987; Krug, 1986; see Table 1). When antecedent variables are included in a mediated-effects model, we will term this combination as Model C_1 (e.g., Goldring & Pasternak, 1994; Hallinger et al., 1990; Heck et al., 1990, 1993, Leithwood et al., 1993; Leitner, 1994; Scott & Teddlie, 1987; Weil et al., 1984; see Table 1).

To illustrate two contrasting conceptualizations within this analytic scheme, some studies in the review simply seek to establish whether certain relationships exist between the principal's leadership and other dependent variables including, but not limited to school achievement. The purpose of these studies is often to address a question of practical interest. Thus, the researcher might ask, how do principals in high- or low-achieving schools differ in terms of their attitudes, beliefs, or behavior? Most of these studies are not primarily concerned with theory in the sense of contributing to our understanding of larger theoretical issues. These studies tend to cluster into what we will later describe as direct-effects studies (see Models A and A_1 in Figure 1).

In contrast, quite a few researchers were explicit in their attempt to link their empirical analysis to theoretical issues involving principal role and the relationship of leadership to the organization and its environment (e.g., Goldring & Pasternak, 1994; Heck et al., 1990; Jones,

1987; Leithwood et al., 1993; Leitner, 1994; Ogawa & Hart, 1985; Silins, 1994). Such studies developed and tested more complete theoretical models. Occasionally, they also tested competing theoretical models about the effects of leadership (e.g., Cheng, 1994; Silins, 1994) and gave greater attention to the replication of findings across a variety of contextual conditions. In general, these studies also tended to employ more sophisticated designs and analytical techniques by which to test their data against the proposed theoretical relationships.

By our measure, approximately one-half of the studies could be classified as being driven by broad theoretical and methodological issues. These studies were disproportionately represented under Model C₁, C, and A₁. In these reports, the researchers addressed issues that went well beyond the common empirical concern about the principal's impact that draws these investigations to our attention. This represents significant improvement over the picture depicted by Bridges in 1982, who termed research in the field as intellectual random events.

Table 1 presents our full classification of the studies included in our review. As summarized in the table, the studies reviewed most frequently incorporated features of the antecedent-effects, direct-effects, and mediated-effects models. Less frequent were studies that used the moderated- or reciprocal-effects models in studying the relationship between principal leadership and school effectiveness (see Hallinger & Heck, in press for a more complete discussion of the conceptual models in use). It may be further noted that there is a general trend over time (i.e., from 1980 to 1995) in terms of the complexity of the model tested. Research tends to be moving from more simple, direct-effects models to more comprehensive models (e.g., mediated effects with antecedents).

Insert Table 1 About Here

Conceptualizing and Operationalizing Variables

Within any of the previously outlined theoretical models, researchers still have numerous choices as to the selection and operationalization of variables for studying the role of the principal in school effectiveness. Foremost, perhaps, for this review is the theoretical orientation and steps taken to model and measure the principal's leadership role. As suggested by our previous discussion, differences in how the theoretical models are conceptualized and variables operationalized have important implications for the ways in which the models are tested. This in turn has an impact on the types of conclusions that we can draw about the nature of principal leadership and its effects. Here we will briefly examine the predominant conceptualizations of variables included in the studies. Consistent with our notion of comprehensive models, we will briefly examine principal leadership, antecedent/context variables, mediating variables, and school outcomes.

Principal leadership. The conceptualization of principal leadership has evolved considerably over the past 25 years (Duke, this volume; Glasman & Heck, 1992; Hallinger, 1992; Leithwood et al., 1992). Predominant notions of the principal's role have evolved from manager, to street-level bureaucrat, to change agent, to instructional manager, to instructional leader, to transformational leader. Even within the past decade, we have observed a discernable shift in emphasis in the conceptualization of the principal's role. Thus, we have begun to see less emphasis given to the instructional leadership role and more to various models construed as more consistent with school restructuring such as transformational leadership (Hallinger, 1992; Leithwood & Jantzi, 1993).

As Glasman and Heck (1992) argue, the changing role conceptualizations may result from increasing external demands and reflect the reform of an educational system that is moving from a closed to a more open system of governance. This has implications for management, with a decreased emphasis on centralized, directive management and an increased emphasis on participatory leadership and decentralized decisionmaking. This evolving view of leadership continues to reflect a recognition of the importance of the principal as a leader, but with different expectations of both stylistic and substantive nature.

The evolving conceptualization of the principal's role is captured in the chronological progression of the studies. Studies from the early to late 1980's were dominated by an instructional leadership conceptualization drawn from the effective schools literature (e.g., Andrews & Soder, 1987; Biester et al., 1984; Hallinger & Murphy, 1986; Jackson, 1982; Jones, 1987; Krug, 1984; O'Day, 1983; Scott & Teddlie, 1987). This was not unexpected given the policy context of the past decade or so. Though defined in a variety of ways, 30 of the 40 studies conceptualized the principal's role in school effectiveness in terms of instructional leadership.

When an instructional leadership construct was employed, researchers most frequently drew on comprehensive conceptualizations of the principal's role based on the theoretical models of Bossert and his colleagues [(1982) e.g., Hallinger et al., 1990; Heck et al., 1990; Heck, 1993], Hallinger and Murphy [(1985; Weil et al., 1984) e.g., Andrews & Soder, 1987; Andrews, Soder & Jacoby, 1986; Dilworth, 1987; Jones, 1987; Krug, 1986; Leitner, 1994; O'Day, 1983; Ruczieska, 1988], or the Connecticut State Education Department (Sirois, & Villanova, 1982; Villanova, Gauthier, Proctor, & Shoemaker, 1981). Others also focused on instructional leadership, but were less comprehensive in their approach to conceptualizing instructional leadership. For example, Biester and his colleagues (1984) focused on the effects of "achievement-directed leadership." Braughton and Riley (1991) investigated principals' knowledge and involvement in the reading process with school outcomes in reading. In two studies, Glasman (1983, 1984) explored how principals use achievement data to stimulate teachers towards instructional improvement. Both Bamburg and Andrews (1990) and Goldring and Pasternak (1994) employed instructional leadership conceptualizations that relied heavily on the principals' use of goals as an agent for achieving an impact on the school.

The conceptualizations embedded in more recent studies tend to take a less confined approach to the role. Three studies examined transformational and transactional leadership constructs (Leithwood et al., 1993; Leithwood, 1994; Silins, 1994). One used Bolman and Deal's (1992) four frames of organization as the basis for their study of principal leadership (Cheng, 1994). The remaining studies were either not explicit in their conception of principal leadership

(e.g., Glasman & Binianimov, 1981; Ramey et al., 1982) or relied solely on the principal's role as a proxy for determination of leadership effects (e.g., Ogawa & Hart, 1985; Rowan & Denk, 1984).

Personal antecedents and school context. Our analysis suggests that the complexity of the relationship between principal leadership and student outcomes overmatched the conceptual and methodological tools being used by researchers. Beyond the simple question of whether principals make a difference, researchers have since sought to further understand how the context in which the principal works influence the expectations and requirements for leadership and the subsequent responses of principals. Our analysis indicates that some progress has been made in increasing our understanding of how contextual factors influence school leadership. Notably, this has been achieved primarily through studies using Model B and C₁ theoretical frameworks.

That said, the surface has only been scratched. There has been a fairly wide range of variation in the operationalization of environmental variables. A few (e.g., socioeconomic status and school level) have received considerably more attention than others (e.g., district structures, community political conditions, rural/urban, cultural distinctions). Similarly, past personal demographics, (e.g., experience, education) personal variables have not received much, if any, consideration in the past group of studies (e.g., gender differences, professional and organizational socialization). From our perspective, there is certainly room for future research that incorporates various socialization processes and gender-related research (and alternative theoretical perspectives) into the impact of principals upon school processes and outcomes.

Mediating variables. As a group, the studies offer several hints at promising mediating variables between principal actions and school effectiveness. A multi-level perspective on schools suggests that principals work primarily with teachers across classrooms and teachers with students within classrooms. This makes untangling principal effects relatively complex.

Because organizations are themselves socially constructed, defining potential in-school processes that are affected by leadership depends heavily on the perceptions of those in the organization who contribute to the creation of an organizational reality. The processes we refer

to as intervening variables (e.g., student opportunity to learn) are therefore often more perceived than real. Thus, we can only observe these processes indirectly.

Recent methodological innovations such as structural equation modeling provide an important means for investigating the school processes that comprise important mediating variables in our conceptual models. The system of structural relations explicated in a model, (and tested empirically) serves as an intermediary between visible manifestations of the process and the abstract theories or models through which we interpret and understand these processes. The types of in-school processes that these researchers have studied can be referred to in structural equation modeling as latent (or unobserved) variables. These are defined by any number of observed manifestations of the process. Together the observed manifestations serve as a proxy for the hypothesized dimension (e.g., see Hallinger et al., 1990; Heck, 1993; Heck et al., 1990; Leithwood et al., 1993).

Leithwood (1994), for example, focused on psychological dispositions of teachers with respect to leadership actions of the principals. He studied teacher perceptions about school culture, decision-making processes, programs and instruction, school goals, resources, commitment and attitudes toward change, and organizational development. Other studies operationalized aspects of the Bossert et al. (1982) framework. These examined facets of school climate, or culture, and instructional organization.

Weil et al. (1984) concluded that principals in outstanding elementary schools had a strong effect on teachers' perceptions of the learning environment. They explored the effects of high expectations and success with students, student motivation, alignment of curriculum relative to principals in average-achieving schools. The researchers noted differences between effective and average schools in terms of the social environment (sense of community, student involvement, orderly environment) and leadership. Many of these perceptions were triangulated with data from teachers, students, and parents.

Several studies found positive relationships between principals' supervision of instructional processes and outcomes (e.g., Brewer, 1993; Hallinger et al., 1990; Heck et al., 1990; Heck et al.,

1991). Our own reanalysis of the Braughton and Riley(1991) study demonstrates that principals' active involvement in supervising the work of teachers has a substantial positive indirect effect on students' reading outcomes. A number of studies also produced consistent findings in terms of the principal's role in setting and communicating school goals (e.g., Brewer, 1993; Heck et al., 1990).

Interestingly, when the studies that report positive findings are reviewed, only one mediating variable shows up with consistency as a significant factor interacting with principal leadership: school goals (e.g., Bamberg & Andrews, 1990; Brewer, 1993; Goldring & Pasternak, 1994; Hallinger et al., in press; Hallinger & Murphy, 1986; Heck et al., 1990; Leithwood, 1994; Silins, 1994). This variable was measured differently in various studies. In some studies the goal variable was measured in terms of goal consensus; in others it was operationalized in terms of the simple presence of school goals, the degree of academic focus, principal vision or focus, or the principal's role in communicating a mission. For the purposes of this chapter, we cannot do more than note that this function of the principal -- establishing and sustaining a school-wide purpose -- does receive empirical support. Clearly this will represent an area for further investigation within mediated-effects models.

School outcomes or effectiveness. The question that often guides discussions of principal impact -- "Do principals make a difference?" -- is subject to widely varying interpretations. While research that examines impact on school outcomes is highly attractive from a policy perspective, it has been fraught with conceptual and methodological problems that are not easily untangled. One of the most salient strengths of this research-- its reliance on standardized test scores as the outcome measure-- has also come to be viewed as one of its greatest limitations.

During the 1980's many policymakers, preoccupied with quantitative measures of educational productivity, cast the question of principal impact in terms of its effects on a single variable--student achievement. Research adopting this perspective first began to appear in 1984 (e.g., Biester et al., 1984; Glasman, 1984; O'Day, 1984; Rowan & Denk,1984). These early

studies of principal impact on student learning foreshadowed the increasing interest among researchers that would emerge over the next decade.

In the mid-1990's it is increasingly accepted internationally that effective education is represented by a wider range of cognitive and affective variables than are measured by the common standardized achievement tests (Leithwood et al., 1990). We believe that the continued reliance on narrow standardized measures for assessing administrative impact distorts the meaning of the question, "Do principals make a difference? Perhaps more importantly, because of the research shortcomings mentioned previously, the empirical findings that accumulated have often been viewed as ambiguous. At times, they are even contradictory. This may be at least in part a function of how school outcomes have been conceived.

When we inquire into the impact of principal leadership, a wide range of dependent variables is available for study. The dependent variables may be broadly grouped into at least three categories of impact: school and environmental effects (e.g., parent satisfaction, community participation, perceptions of school functioning), intra-organizational processes (e.g., staff morale, curricular organization, instructional effectiveness), and student effects (e.g., student achievement, attitudes, retention). The priority assigned to these variables varies widely, both within and across countries, and is implicitly value-laden.

As restructuring has become a more popular policy solution during the 1990's, we note a corresponding shift in thinking toward a broader view of potential principal impact. The later studies explore the relationship between principal leadership and a wider range of intervening classroom and school level variables in the belief that the impact of principal leadership is likely to be indirect in nature (e.g., Goldring & Pasternak, 1994; Hallinger et al., 1990), or look more specifically at leadership aimed at "second order," or restructuring, changes (e.g., Leithwood et al., 1993; Leithwood, 1994; Silins, 1994). These studies tend to group into the second category of concern--intra-organizational processes. As evidenced in Table 1, however, little attention has been directed toward environmental outcomes such as parent satisfaction or community participation.

While the assessment of higher-order cognitive outcomes is not in evidence in these studies, the researchers have sought to move away, to some degree, from exclusive reliance on standardized tests to assess student performance. As shown in Table 1, there is some variation in the ways in which school outcomes have been conceptualized and measured (e.g., test scores, school effectiveness, improving or declining schools, teacher-perceived program, school, and student outcomes).

Most often the dependent variable has tended to be student scores on standardized tests or an overall school effectiveness variable. These measures may or may not include controls for contextual differences in schools [see Hallinger & Murphy, (1986) or Heck et al., (1990) for examples]. The validity and reliability of this range of dependent variables is an important issue in interpreting the findings from these studies. One must consider the extent to which a standardized test score or any number of definitions of effectiveness represent a valid school "outcomes."

We believe continued effort must be directed toward broadening the school outcome construct and in considering the findings of individual studies in light of the particular dependent constructs investigated. For example, the psychometric properties of any number of dependent constructs can be reported, but this issue was seldom addressed in this set of studies. While standardized tests, for example, are typically highly reliable, their validity has been open to increasing debate.

Sampling

There are several important issues in sampling. For the purposes of our review we will concern ourselves with the representativeness of samples (i.e., the extent to which the sample represents the populations from which they are chosen) and the precision of estimates based on sample size (Pedhuzer & Schmelkin, 1991).

Nature of the samples. These reports reveal a wide range in the nature of the samples studied. They indicate a distinct preference among researchers for studying elementary school principals. Of these 40 reports, over two-thirds either studied elementary schools exclusively or in

combination with schools at other levels. Nine studies included principals at the high school level along with elementary schools, and three additional studies explored leadership solely at the high school level.

The predilection for studying the impact of elementary school principal's leadership is not surprising. Researchers often base their preference for studying the impact of the principal at the elementary level because of the greater likelihood of detecting effects in the smaller, less organizationally complex, venue of elementary schools. Aside from this factor, exploring the relationship between principal leadership and school effectiveness at the high school level also poses greater conceptual and methodological challenges (Firestone & Herriott, 1982; Hallinger & Murphy, 1986a).

Given the non-experimental nature of this research, the sampling design of the studies becomes an important issue when we seek to detect the effects of principal leadership quantitatively (Bridges, 1982; Eberts & Stone, 1988; van de Grift, 1990). How the sample is obtained is central to the validity of results. Samples are broadly of two types: nonprobability samples, which are based on convenience or accessibility; and probability samples, where assumptions are made about the criteria and procedures for obtaining the sample. Often, however, the use of intact schools, "statistical tails" consisting of the best and worst schools in a population, or various added criteria regarding the stability of staff and principal precluded the use of true random samples in this body of research.

While several studies employed elements of randomization (e.g., selecting teachers at random within a nonprobability sample of schools), few studies employed truly randomly selected samples. Notable exceptions are Goldring and Pasternak (1994) and van de Grift (1990). Others addressed the sampling issue by attempting to be exhaustive in focusing on a single school system or complex of schools and including every school (or a sample of schools) and their personnel within the unit (e.g., Heck et al., 1991; Leitner, 1994). This approach approximates cluster sampling (Pedhazur & Schmelkin, 1991). There is considerable flexibility in how the samples were selected.

Size of samples. Decisions about sample size are relatively complex and subject to many concerns. These can include sampling strategy, economic, and practical considerations. Of considerable importance in interpreting the findings of any study is the concept of effect size; that is, the degree to which the phenomenon is present in the population (Cohen, 1988). For observing the principal's impact on school processes or outcomes, we must assume that the effect size is relatively small. To detect small effects, therefore, implies the need for larger samples. Otherwise, one may fail to reject the null hypothesis because the research was not designed with sufficient ability to detect the effect due to the small sample size. Power analysis concerns the ability of the statistical analysis conducted to detect the presence of an effect, should one exist, given various effect sizes, levels of significance, and sample size (Cohen, 1988). Other things being equal, small samples, therefore, are only capable of detecting relatively large effects in the population.

The size of the principal or school samples studied ranged widely from a low of 75 teachers and principals in a complex of four schools (Heck et al., 1991) to a high of about 300 principals (Eberts & Stone, 1988; Glasman, 1984). The median sample size was 34 schools or principals studied. Much of the discussion of sample size also depends upon the purposes of the research, the unit of analysis, and analytical techniques used to investigate the data. Our judgment is that many of the studies reviewed had samples that were not selected through optimal sampling methods and were too small to detect leadership effects if they were present. Thus, caution must be employed in evaluating the findings of individual studies in light of the sampling methods used.

In the absence of preferred methods of sampling, however, replication across educational settings and with various instruments becomes an important means of increasing confidence in findings. From this standpoint, while the majority of studies reviewed have sampling problems to some extent when viewed in isolation, the convergence of findings from Model C studies that principals do affect school outcomes at least indirectly is encouraging. This is especially true since they were conducted in a variety of educational settings, with varied samples, using

different instruments to operationalize leadership constructs, and a variety of techniques used to analyze the data. Thus, despite the sampling problems of individual studies, as a group these problems become less severe.

Unit of Analysis

Researchers have recently given increased attention to the measurement of variables across levels of organizations. In education, these distinctions are readily apparent, but have been difficult to investigate empirically. Students bring individual abilities to their classrooms. Teachers shape the children's classroom environment. Principals monitor teachers within their schools. Superintendents develop improvement plans for their districts. Theoretically, this has been referred to as a “nested structure” and represents one example of a hierarchical data structure (Bossert et al., 1982).

When studying the interrelationships among principal, teacher, and student-level variables, this structural feature of educational organizations takes on particular importance. This is the case because principals are likely to influence the school level of the organization more directly than the classroom (e.g., how teachers organize instruction) or student levels (the motivation of particular students). We immediately run into problems in attempting to decide where the appropriate unit of analysis is when looking at data intended for school-level analyses.

One illustration of this concerns whether constructs such as school climate or principal leadership are basic properties of the organization or merely perceptions of the individuals. If we accept the former conceptualization, perceptions should be measured at the school level. In the latter case, they would more appropriately be measured at the individual teacher or student level.

Earlier studies of schools as organizations seldom addressed the problem of variables that are impacted by multiple levels of the organization (e.g., student achievement). In this group of studies, for example, the unit of analysis tended to be either the individual level (e.g., teachers and principal as individuals without regard to their school setting) or the school level. In the latter case, teacher responses were summed to create school means. The researchers would then compare *schools* within the sample.

Neither solution is completely satisfactory. A limitation of a school-level analysis is that every individual in the school is assumed to hold the same perception about the principal. Individual level analyses do not allow us to assess accurately the effects of different levels of the organization on the relevant outcome of interest. For example, if quality of teaching is hypothesized to affect student learning, then we know that some students in our sample of individuals have the same teacher and, further, differ in the quality of teaching that they receive. To ignore this by disaggregating the data, therefore, we violate the assumption of independence of observations that is basic for classical statistical techniques. That is, systematic groups of students in our sample would have the same value on all variables at the classroom level (Bryk & Raudenbush, 1992).

Several promising analytic techniques have been developed over the past few years (e.g., structural equation modeling, hierarchical linear modeling) that allow the investigation of these effects across organizational levels. One such study (Rowan, Raudenbush, & Kang, 1991) serves as a good example of how this approach can yield useful information about administrative effects. This study was not included in our review because it did not use school achievement as a dependent variable.

Rowan et al. (1991) looked at differences in principal leadership and other organizational processes using a multilevel analysis. Hierarchical linear modeling (Bryk & Raudenbush, 1992) was first employed to separate the variance accounted for in principal leadership practices into within-school and between-school components. From this analysis, they determined that most of the variance in teacher perceptions about principal leadership (roughly 75 percent) was actually the result of within-school variation in how teachers view the principal's actions. Of course, this variation would be lost if measurements were simply aggregated to school level means, and those school-level means used in a between-school analysis. Therefore, studies that do not account for within-school variation can run the risk of over-emphasizing differences in leadership between schools.

Now that the analytic technique is available, one alternative solution proposed through hierarchical linear modeling is to develop a within-school model to determine the relative effects of various factors (e.g., gender, backgrounds of teachers) on teachers' perceptions about the principal's leadership. In this manner, we can also determine whether leadership has been measured with sufficient reliability and validity within each school to permit a school-level analysis of differences in principal leadership. If there is, the adjusted leadership mean for each school can then be used as the dependent variable in a between-school comparison.

The smaller proportion of variance in principal leadership due to differences between schools (in Rowan et al.'s 1991 study this was about 25 percent), for example, might be additionally affected by the level of the school, whether it was public or private, and the socioeconomic status of the community. Thus, the estimated effects of independent variables at the individual level of the school can be adjusted simultaneously for effects that may be present at other levels of analysis (Rowan et al., 1991). The approach is illuminating for many research situations in that it allows the variability in important dependent variables (e.g., outcomes, leadership) to be decomposed across different levels of the organization. Although not without its own limitations as an analytical technique (e.g., assessing indirect effects), this application should prove to be useful in instances where researchers are attempting to answer questions about the principal's impact across various layers of the school.

Data Collection Methods

As noted a dozen years ago, surveys continue to be the data collection strategy of choice among researchers who examine the impact of school administrators (Bridges, 1982). Of the 40 studies analyzed here, all but four relied on a survey for all or part of the data collected in connection with the school-level variables (i.e., leadership, in-school processes). Six studies incorporated interviews as all or part of their data collection schemes (Biester et al., 1984; Blank, 1987; Jackson, 1982; Glasman, 1992; Glasman & Fuller, 1992; Hallinger & Murphy, 1986).

While most of the studies drew on demographic data to supplement interviews and/or surveys, two investigations appeared to rely solely upon demographic data for their analysis

(Ogawa & Hart, 1985; Rowan & Denk, 1984). Although it was hardly a shock to find surveys comprising the predominant mode of data collection, we were surprised by the relative dearth of mixed method studies. Only six of the studies could truly be classified as mixed qualitative and quantitative studies. As we suggested previously, however, our own bias in identifying studies for the review was weighted towards quantitative studies because of our interest in determining the trend of effects uncovered by those who have investigated the impact of the principal's leadership on student achievement.

Given the fact that the majority of studies of principal impact are quantitative and used data collected with surveys, issues of instrument reliability and validity take on great importance. In his review, Bridges (1982) was highly critical of the state-of-the-art when it came to the care with which researchers in educational administrator developed, used, and tested their instrumentation. Issues of instrument reliability are highly germane in the field of principal effects research. Research instruments with low (or undocumented) reliability will have difficulty accurately detecting differences among subjects without samples larger than those that typify most of the research in this field.

Judging from our reading of these studies, some progress has been made in developing reliable instrumentation for measuring principal leadership. Instrumentation for measuring the instructional leadership construct developed by Villanova and his colleagues in Connecticut (1982) was used with documented reliability (e.g., Hallinger et al. 1990; Scott & Teddlie, 1987). Likewise, the reliability of the *Principal Instructional Management Rating Scale* (Hallinger, 1983) was documented in several investigations (Hallinger, Taraseina, & Miller, 1994; Jones, 1987; Leitner, 1994; O'Day, 1983).

The analytic approach taken by Rowan et al. (1991), for example, can also be used to determine how reliably leadership can be measured both within and between schools. Marcoulides and Heck (1992) also presented a procedure for establishing the reliability and validity of observations on the principal. They analyzed the variance components (i.e., leadership behavior, teachers as raters, occasions of measurement) of teachers' perceptions of principal

leadership within schools and found that teachers as raters were a substantial source of error variation (in this case representing about 20 percent), but that occasions of measurement were a small source of error (only about 1 percent). The finding indicates increasing the number of teachers sampled within each school reduces the source of error observed, but that measurements about principal leadership can be reliably gathered from each school on one occasion.

Still of concern in this domain is the less frequent documentation of the validity of instrumentation. Very few of the studies that we reviewed gave explicit attention to issues of instrument validity. It should also be noted, however, that while these instruments achieved consistently high levels of reliability, most of the studies were conducted at the elementary level.

To illustrate the importance of this issue, we need only examine one study that did systematically examine the validity of instrumentation as a step in its data analysis. Jones (1987) used the PIMRS (Hallinger, 1983) to collect data on principal instructional leadership in a sample of Canadian high schools. While the instrument possesses documented validity at the elementary school level, Jones' (1987) findings cast doubt on whether the same level of validity is may be expected at the high school level. Given differences in the principal's role at the secondary level, as well as in other community and cultural settings, researchers must continue to be vigilant in testing their instruments across a variety of conditions (Firestone & Herriott, 1982).

Data Analysis

Beyond issues of sampling, unit of analysis, and data collection, we can examine the extent to which the analyses conducted are appropriate for the expressed purpose of the research and the particular theoretical model being investigated (Tatsuoka & Silver, 1988). A wide variety of qualitative and quantitative analytical techniques are currently available for research in educational administration. Some excellent reviews of the application of those techniques have already been provided elsewhere (e.g., Everhart, 1988; Tatsuoka & Silver, 1988). Because of the exclusive use of non-experimental designs in work that is being done in this sub-field of educational administration, we will confine our comments about analytical techniques to those

commonly used with this type of research (see Pedhauzer & Schmelkin (1991) for a more complete discussion of appropriate techniques for use with non-experimental research).

Tatsuoka and Silver (1988) categorized available methods of data analysis into four broad categories -- descriptive (e.g., means, percentages, cross tabulations), analysis of variance (i.e., ANOVA, MANOVA, analysis of covariance), correlational (e.g., correlation, regression), and causal inference methods (e.g., path analysis, structural equation modeling). Pitner (1988) argued that of these categories of techniques, only causal inference methods are an appropriate means of testing the types of models proposed in Figure 1. She made this claim because these methods are able to provide estimates of both the direct and indirect effects implied in the models. After analyzing the studies in this review, however, we have come to believe that the issue is more complicated than a simple right-approach, wrong-approach.

For the most part, the techniques used in analyzing the impact of principals over the past 15 years demonstrate greater variation and considerable improvement over those reported by Bridges in 1982. The studies analyzed here used techniques in all four broad categories outlined by Tatsuoka and Silver (1988). We do, however, note differences in the extent to which the analytic techniques employed were able to provide a complete test of the models implied in the studies.

Analytical techniques used to test direct-effects models. Any analytic technique must provide an adequate test of the theoretical model implied in the study. Bridges (1982) reported that earlier research on school administrators was inclined to use single factor, correlational analytic techniques for examining relationships between principal leadership and organizational outcomes (see Model A studies, Figure 1). Often such tests were conducted without including relevant control variables. Moreover, Bridges noted that it was often the case that the theoretical models being studied were not fully explicated. These represent severe constraints on the construction of knowledge. Without controlling for relevant variables or explicating the theoretical relationships among variables, analytic techniques offer little hope of shedding light on causal relationships.

As Table 1 indicates, researchers investigating Model A frameworks (direct-effects models) tended to use descriptive statistics, bivariate correlational analysis, and tests of significance between groups of principals. A common approach has been to use common t-tests for two groups or analysis of variance (ANOVA) when there are more than two categories. While several studies isolated some positive (or mixed) differences in principals' actions, these techniques were not fully capable of testing for the direct effects of principal leadership on school outcomes. In some cases, it should be noted, this was not the researcher's primary intent (e.g., Glasman, 1983, 1992).

The frequent use of t-tests, ANOVA and correlational methods in direct-effects research (A and A₁) represents a continuing drag on the accumulation of knowledge in our field. Unfortunately, these are often the only kinds of analytic approaches to which many researchers are exposed (Pedhauzer & Schmelkin, 1991), including those in educational administration (Bridges, 1982; Pitner, 1988). The explanation for this is partially historical. ANOVA and related techniques have been strongly associated with experimental research in psychology. Regression analysis is commonly used for research in natural settings as in sociology and economics (Amick & Walberg, 1978). These disciplines represent the most common methodological bases for studies in educational administration over the past several decades.

The historical context of usage has implications for the appropriateness in which the analyses are carried out in non-experimental research. A bit more discussion of the application of available analytic techniques may be useful here. In terms of non-experimental research, techniques such as correlation and t-tests are not "causal" in the term's limited sense of referring to an integrated system of relationships. For example, with respect to correlation, any two variables may be correlated but still not exist in a theoretical system of relations. Similarly, because the results of a t-test indicate that two groups of principals are not the "same" (i.e., they differ "significantly") in terms of their perceived leadership, it does indicate what might account for this observed difference. The conceptual leaps that often follow from such findings of differences among groups are often made on very shaky analytical (and theoretical) grounds.

Similar arguments can be made about other tests of group differences (e.g., ANOVA, MANOVA) when they are not used in an experimental setting. In experimental research, relevant variables are controlled and the independent variable is actually manipulated as a means of testing for its effects on a dependent variable. Such is not the case in non-experimental studies. Depending upon the interpretation made by the researcher, this can represent a misuse of techniques for analysis of data in non-experimental studies.

The misuse of these analytical techniques is related to the assumptions they make concerning the process of data analysis. In non-experimental designs we work backward from the dependent variable to relevant predictor variables hoping to explain possible outcome variation. These techniques often give the unfortunate impression of having conducted an "experiment" when in fact one has not done so. This can lead to incorrect conclusions being drawn from the results.

ANOVA-type designs (including analysis of covariance) also make assumptions that the independent variables are truly "independent" of each other. In fact, in studies of principal and school effects, the variables investigated are quite highly correlated. Thus, the use of such analytical methods in non-experimental research can lead to serious errors and misspecifications when seeking to understand causal relationships [see Pedhazur & Schmelkin (1991) for a full discussion of this problem].

Moreover, while continuous independent variables (e.g., SES) are common in this field, researchers often "carve" up such variables into categories and employ an ANOVA design. This represents another misuse of a design that is better suited to analyzing the effects of varying levels of a treatment administered than arbitrary categories of SES (e.g., high and low). Moreover, within the A_1 model, analysis of variance cannot analyze both the direct effects of antecedents on leadership and their corresponding indirect effects on outcomes (i.e., through leadership). This is a case where the statistical tests were simply inadequate to the demands of the theoretical model and the research design. Approaches based on analysis of variance are, therefore, severely limited. This conclusion extends even to their ability to shed light on the theoretically simple relationships proposed in Model A and A_1 studies.

Model A₁ studies show somewhat more variation in terms of analytic techniques. Recall that the theoretical model implied in these studies includes antecedent variables, leadership, and outcomes (see Figure 1). In addition to correlational analyses and those emphasizing differences between groups of principals, some of these direct-effects studies used regression analysis. The goal of research in regression studies -- to explain variation in the dependent variable -- is generally more consistent with the logic of non-experimental research designs than is analysis of variance (i.e., that groups are not the same). In multiple regression analysis the relationship among independent and dependent variables must be fixed in advance. Here the categorical independent variables (e.g., strong, average, weak leadership) are treated as a set of coded vectors, and the dependent variable (e.g., outcomes) is continuous.

The multiple regression approach can accommodate categorical, continuous, and combinations of two types of independent variables in the same analysis. Moreover, several independent variables can be considered more efficiently within the same analysis. The lesser risk, therefore, in Model A and A₁ studies, where some type of theoretical framework is at least inferred, is to use regression analysis.

As can be seen in Table 1, regression was used frequently in Model A₁ (antecedent/direct-effects) studies. Where dependent variables (e.g., school effectiveness) were categorical, discriminant analysis was also used occasionally (e.g., Goldring & Pasternak, 1994; Heck, 1993). We classified these two studies as regression studies because their intent was to investigate several sets of variables related to the school outcome variables (i.e., effective and ineffective schools). We make note of this because discriminant analysis is conceptually closer to multivariate analysis of variance than to regression in that its main purpose is the description of group differences.

In terms of its use in the explanation of categorical outcomes, discriminant analysis is now being replaced by newer and more flexible regression programs designed for categorical data (e.g., logistic regression, hierarchical log-linear models). The analysis of categorical dependent

variables requires these newer techniques because the observations are generally not from populations that meet the assumption of being normally distributed.

Exploring principal effects through regression analysis. Pitner (1988) argued that causal inference techniques are the most appropriate to use with all of the models in Figure 1. While we do not disagree with this view, we believe the regression model is flexible enough to encourage its use in many, though not all circumstances. This judgment includes those procedures used for categorical dependent variables and nonlinear relationships.

The ultimate value of any regression approach is, of course, dependent on the substantive model it is meant to represent. While in regression analysis it is not necessary to make assumptions about the causal structure linking predictor variables, interpreting the results of these analyses does require some assumptions about the underlying causal relationships (Mueller, Schuessler, & Costner, 1977). Conclusions about the relative importance of predictors must therefore take into consideration both their direct and indirect effects (i.e., effects mediated by other variables). This is true whether or not the full range of effects is actually investigated. Again, this points to the importance of model specification through a proposed theory of the phenomenon under investigation *before* actual testing with data. Analysis should then focus on the extent to which the results are consistent and logical given the model.

Regression-type analyses are, however, not without problems. One of the most serious problems is model misspecification. A common habit among researchers using multiple regression analysis is to fish for explanations. This is especially dangerous in non-experimental research, in which prior theoretical explication of the model is critical to overcoming inherent limitations of the research design. Regression models are also limited in that they are not fully suited to test for indirect effects. At best, such effects can sometimes be implied through the cautious use of varying the entry of variables into the model (or sometimes by specifying interaction terms). An example will serve to illustrate.

In a direct-effects model (Model A), leadership variables are hypothesized to affect school outcomes. When an antecedent variable is added (as in Model A₁), the two independent variables

(antecedent and leadership) can be treated in two ways. First, as a single-stage model, the antecedents and leadership can be treated as correlated independent variables, where the correlation is treated as given but not explained. This, for example, is how the van de Grift (1990) study treated the relationship between leadership and SES.

In contrast, in a two-stage model, the antecedent (e.g., SES) can be hypothesized to affect leadership. Now the two independent variables are no longer merely correlated. Order of entry into the equation makes a difference (i.e., earlier entry generally accounts for more variance in the dependent variable). In explanatory research, the decision must be made on the basis of one's underlying theory. One can make a valid case for entering SES first, as it is now hypothesized to affect both leadership and outcomes. Leadership would be entered second and it would be expected that SES would account for the most variance in outcomes (because of its combined direct and indirect effects) and leadership a smaller, incremental amount. It would not make sense to enter leadership first because it would not be hypothesized to affect SES.

As our discussion of research in this chapter has attempted to illustrate, specifying the appropriate theoretical model in advance (as opposed to upon examination of the results) is of utmost importance to determining and interpreting what is found, no matter which analytical technique is used. Regression, and other path-analytic approaches, are therefore most effectively used when guided by strong theory. This discourages the use of procedures such as stepwise entry of independent variables, because such techniques emphasize statistical significance as opposed to an a priori set of defined theoretical relationships.

Interestingly, the same problem arose in the limited instances where structural equation modeling was used in direct-effects studies (e.g., van de Grift, 1990). Here the method was adequate to the test, but the theoretical model was not sufficiently developed. The result in this instance was the inverse: the statistical test was underused. That is, it was not used to reveal the full set of relationships, though it might have (see below). This again highlights the importance of starting with a well explicated theoretical model before choosing and applying the statistical test.

Analytical techniques used to test mediated-effects models. More complex theoretical models require the use of what Tatsuoka and Silver (1988) termed causal inference techniques. These include path-analytic methods, which are increasingly being used to replace multiple regression analysis in Model C and C₁ studies. There are a wide variety of terms used to describe these types of analytic techniques.

Path analysis has been used to refer to models where single (observed) indicators are used to represent the variables in the theoretical model. Covariance structure models, latent variable models, structural equation models (SEM) are all terms that refer to models that have observed and underlying (latent) variables. They are erroneously referred to by the computer programs used to analyze the data (e.g., LISREL, EQS). The growing popularity of these techniques is due to their flexibility in handling a wide variety of theoretical models. These include direct, indirect, and total effects, as well as reciprocal (nonrecursive) and hierarchical relationships.

At the time of Pitner's (1988) review, she could find no study that employed causal inference techniques. From this standpoint, the studies in our review are much improved over those in Pitner's (1988) analysis. The majority of model C studies use some type of causal inference technique. This bodes well with respect to the technical requirements needed to understand how administrative leadership influences organizational processes and performance.

As our discussion of the conceptual and methodological progress in this sub-field of educational administration shows, the formulation of a theoretical model is a long process entailing a great deal of creativity, critical thinking, insight, and empirical validation. While we caution against their indiscriminant use, we also believe that these techniques have the potential to open up exciting new possibilities for research. In order to demonstrate how theoretical and methodological advances combine to create a convergence of findings we will provide two examples from the set of studies.

Two new analyses. To test the potential of path modeling, we decided to reconceptualize and reanalyze two Model A studies (Braughton & Riley, 1991; van de Grift, 1990). Both studies concluded that there was no positive relationship between principal leadership and school

outcomes. However, the studies gave little attention to possible antecedent variables and mediating school processes.

Broughton and Riley (1991) used an appropriate technique (i.e., regression analysis) to investigate the effects of a variety of leadership and teacher variables on reading scores. Previous reading scores was included as a control, but not considered in terms of direct or indirect effects. All 14 variables were considered as observed and entered into the regression equation. Within their regression model only direct effects were considered. No relationships were considered between principals and teachers, nor between previous reading grade and principal or teacher behavior. As we suggested, in a regression analysis (that is not conceived of as a path analysis), such effects might be partially considered through interaction effects or variable entry (e.g., enter principal set, then teacher set).

We recast the study as Model C₁ to consider possible antecedent variables (i.e., previous reading outcomes) on teachers and principals, and the effects of principal leadership mediated by teachers' classroom practices on school reading outcomes. Thus, our redesigned conceptual model considered several paths of theoretical interest that were not investigated in the previous study. Principal leadership, as defined by principal knowledge, attitudes, and supervisory behavior, consisted of eight observed indicators that formed the latent construct. Teacher classroom practices were also conceived of as a latent variable consisting of four observed indicators of teacher classroom reading practices. We then tested the new proposed theoretical model using structural equation modeling (LISREL 8) and Broughton and Riley's correlation matrix of observed variables as input.

The resulting findings indicated no direct effects of principal leadership (with a standardized path coefficient of .04) on outcomes. Teacher practices were significantly (although negatively) related to outcomes. Our interpretation of this finding is that where student reading scores were lower, teachers were more actively involved in developing students' skills. This is essentially where the previous analysis stopped.

Drawing upon the additional capabilities of structural modeling, we found further that principal actions were significantly (and negatively) related to certain intervening variables: teacher knowledge and practices. The finding indicates that where teachers were less knowledgeable and competent about reading practices, principals were more involved in supervision and direct intervention. Moreover, the total effects of leadership on student outcomes, which considers *both the direct and indirect effects* (i.e., through teacher practices) were significant and positive (.38).

Thus, almost all of the effects of principal leadership on school outcomes were indirect (.34). The effects resulted from the principals' knowledge and skills in monitoring and supervising teachers, as well as improving the reading curriculum. It was also noted that previous reading scores (an antecedent variable) was found to affect subsequent teacher classroom behavior significantly, but not to affect principal behavior significantly. Our results, therefore, supported the theoretical model (Model C, Figure 1) in which principals' actions have an indirect impact on school outcomes through teachers as a mediating factor. The particular antecedent used in the study, however, did not influence leadership practices. The finding demonstrates both the importance of theory in guiding development of empirical model and why it is critical to select the correct statistical method for analysis of the data.

Our reanalysis of one of van de Grift's (1990) studies was similarly illuminating. This was especially true in that this study has often been pointed to as evidence that leadership does not affect school outcomes, at least in The Netherlands (with a reported standardized path coefficient of .002). Van de Grift had previously investigated the direct effects of socioeconomic status and leadership on school outcomes using structural equation modeling (Model A). Therefore, in this instance, we needed only to set up a competing theoretical model and reanalyze his data.

We believe that there is ample evidence to suggest that SES affects leadership practices (e.g., Andrews & Soder, 1987; Hallinger & Murphy, 1986a, Rowan & Denk, 1984). The van de Grift study did not, however, hypothesize any effect between the antecedent variable (SES) and leadership. SES was used simply as a control variable in relation to student achievement. In our

reanalysis, we hypothesized that the antecedent variable, socioeconomic status, should affect leadership directly and outcomes both directly and indirectly (i.e., through leadership).

Consequently we set up a C_1 model portraying antecedent- and mediated-effects.

In addition, van de Grift considered leadership as one variable comprised of four observed subscales (initiating innovations, stimulating teachers, supervising teachers, fostering climate). We chose to create two leadership constructs (instructional organization and school climate), following the Bossert et al. (1982) model of mediated effects. The measurement model in our reanalysis, which estimates relationships between the observed variables and latent constructs, was virtually identical to the various parameters reported previously.

The new structural relationships tested, however, indicated that socioeconomic status affected instructional organization leadership significantly (but weakly), but not leadership involving school climate. Socioeconomic status also affected outcomes significantly, as was found previously. Leadership in monitoring the instructional organization of the school was found to affect leadership in developing school climate significantly, but not the other way around [similar to Heck et al.'s (1990) finding]. Principal leadership aimed at developing school climate was positively related to school outcomes, but not significantly. Moreover, instructional organization leadership was found to be significantly (but weakly) negatively related to outcomes. Thus, the substantive conclusions that can be drawn from the data are more extensive and complex than the conclusion of “no effects” reached in the original study.

The negative coefficient observed between principal leadership aimed at school improvement and teacher supervision (i.e., instructional organization) was somewhat puzzling. One interpretation of the negative coefficient in this study [and the Braughton & Riley (1991) study we reanalyzed] is that where school achievement has been consistently low, teachers perceive that principals are actually working harder to create innovation, stimulate teachers to improve teaching, and supervise school's teaching and learning processes. Interestingly, this path is also similar to Heck's (1993) study in Singapore, where a negative coefficient was also observed between instructional monitoring and outcomes.

It is our belief that this consistency in finding across several studies relates back to how researchers define the dependent variable in their studies. For example, Heck et al.'s (1990) study investigated school effectiveness longitudinally (i.e., over several years) and in relative terms (i.e., relative school outcomes controlled for socioeconomic status and language background), suggesting that "effective" urban schools might produce lower actual outcomes than effective suburban schools because of student composition factors. The studies in our reanalysis [and the Heck (1993) Singapore study] investigated actual outcomes in terms of the school outcome means unadjusted for composition differences.

In both examples we chose to reanalyze, therefore, we found that our proposed structural models fit the data quite well. As is the case with structural models, a model that is not rejected (i.e., that fits the data reasonably well) is only one of many possible models that might fit the data. This emphasizes the importance that theory plays in developing a model (or set of competing models) to be tested (see also Hallinger, Bickman, & Davis in press; Silins, 1994).

Van de Grift's (1990) and Braughton and Riley's (1991) determination that leadership does not affect school outcomes are correct with respect to the direct-effects models that were tested. Our models, conceived to be relatively consistent with the Bossert et al. (1982) model (i.e., antecedent and mediated effects), also fit the data presented in each study, yet they reveal more about the interrelationships between the theoretical constructs comprising each model. They also yield quite different conclusions concerning the nature of principal impact on school achievement. We believe that these reanalyzes demonstrate the greater potential that emerges when more comprehensive theory is combined with sufficiently powerful methods of research.

Construct Validity and Generalizability

Overall, the most important aspect of any of the studies conducted is the interpretation of results (Tatsuoka & Silver, 1988) in light of the theoretical grounding of the model tested and method employed to test the model. There is a close interplay between substantive theory and statistical method in yielding explanations of results (Tatsuoka & Silver, 1988). This is especially true in the social sciences, where one must often rely on quasi-experimental or non-

experimental designs. It is not necessarily the sophistication of the analytic method that primarily determines whether, and to what extent, a study verifies a particular set of theoretical relations. It is rather the manner in which the study is designed and conducted that matters the most (Tatsuoka & Silver, 1988). Ultimately, this process relates to the confidence and validity we can place in the observed outcomes of the investigation.

Of the many approaches to validity that have been discussed in the literature, construct validity is the most general and can be considered to include all others (Cronbach, 1971). Construct validity occurs when the researcher evaluates a set of operations in light of a specified construct (National Council on Measurement in Education, 1984) or set of theoretical relations. The proposed interpretation generates specific testable hypotheses, which are a means of confirming or disconfirming the claim (Cronbach & Meehl (1955). Construct validity is in fact an ongoing process (Heck & Marcoulides, 1992). A variety of statistical procedures have been developed to address this issue [see Pedhazur & Schmelkin (1991) for a complete discussion].

As a group, the studies provide support for the construct validity of the antecedent with mediating-effects model and, more specifically, the mediating-effects model. First, several studies used confirmatory factor analysis, one established means of investigating construct validity (Heck & Marcoulides, 1992) to validate their specific models. This process assesses the extent to which observed variables in a model are related to the theoretical constructs they are hypothesized to measure. The procedure itself forces researchers to define proposed theoretical relationships prior to testing models with data.

Another approach used to establish construct validity is to conduct tests of model invariance across different groups. When group differences are being considered, an assumption is made that the construct(s) being investigated are similar for all groups being examined. If the constructs measured are not similar, this failure to may be due to a lack of construct validity or an indication that the groups are different (Heck & Marcoulides, 1992). The value of a proposed model of leadership is greatly enhanced if the same model can be replicated in subsamples from

different populations (Cattell, 1962). This in fact was observed in several studies that reached similar conclusions in different settings and with different subgroups of the population tested.

Finally, there is evidence that researchers are beginning to test competing theoretical model, which is seen as another means of studying construct validity. Given the empirical evidence of effects from the antecedent with mediated-effects model, as well as our further reanalyses, it would appear that this model represents a promising conceptualization. Our brief discussion suggests that researchers should pay greater attention to issues of validity as well as reliability in their investigations of principal leadership constructs. This is reflected in retrospective reviews of the prior work of researchers that focus on examining the construct validity of particular leadership models (e.g., Heck & Marcoulides, 1992; Leithwood, 1994).

Conclusion

We began this review of the literature on principal effects with uncertainty as to whether the results would warrant the effort. Several respected reviews had already been conducted of this research literature over the past 15 years (Bossert et al., 1982; Bridges, 1982; Leithwood & Montgomery, 1982; Leithwood et al., 1990; Murphy, Hallinger, & Mitman, 1983; Murphy, 1988). Although the rationale for reviewing the empirical literature once again seemed sound, we were unsure what new information of value it would yield to researchers in this domain.

In our judgment, the results have more than warranted the effort. The review reveals what we would characterize as a "leap forward" in the quality of empirical research being conducted in this sub-field of educational administration since 1980. In this most recent generation of principal effects research we found substantial progress on both conceptual and methodological dimensions (see also Hallinger & Heck, in press). The evidence of progress is especially notable in that this particular domain presents formidable challenges to the researcher. In our opinion, the methodological issues are as complex as those posed in any other topical area within educational administration. While we cannot generalize the gains found here to the field at-large, our results

certainly suggest marked improvement in the research being conducted by scholars who have been working on this particular set of issues.

In contrast to Bridges' in 1982, we found a clear accumulation of knowledge in both methodological and conceptual domains. The recommendations of earlier reviewers of this literature (e.g., Bossert, Boyan, Bridges, Cuban, Glasman, Hallinger, Leithwood, Murphy, Pitner, Rowan, van de Grift) as well as the incremental suggestions advanced by who have conducted programmatic empirical research (e.g., Andrews, Cheng, Glasman, Goldring, Hallinger, Heck, Leithwood, Marcoulides, Miskel, Ogawa, Rowan, van de Grift) are interwoven throughout the studies we reviewed. Not surprisingly, this accumulation of knowledge in the processes of research is also, to a degree, reflected in an increasing consistency of findings that emerge from the studies. As noted in our discussion of construct validity, this type of programmatic effort is necessary to developing the foundations on which knowledge is built.

Drawing upon prior research reviews, we focused our lens upon this one subset of the educational administration literature: empirical studies of administrator effects on student learning. The chapter had three primary aims:

- 1) to analyze the theoretical frameworks that have been employed by researchers in the
- 2) to examine the methodological features of this literature;
- 3) to propose a theoretical framework and appropriate methodological approaches that

In this last section of the chapter we will summarize the major findings of the review with respect to the first two goals. In doing so, we will use the findings from Bridges' (1982) review as a benchmark both for identifying the progress that has been made to date and key targets for future improvement. Woven throughout the discussion are the implications of our assessment for future investigations in this domain of educational administration.

Theoretical Frameworks for Conceptualizing Principal Impact

This chapter has reviewed research on the principal's role in school effectiveness conducted between 1980 and 1995. Although the studies vary in their conceptualizations of principal impact, all of them included measures of school outcomes, most frequently student achievement.

We analyzed 40 studies identified by our search on both conceptual and methodological dimensions.

We drew upon a framework proposed by Pitner (1988) for categorizing non-experimental studies of principal impact as a conceptual framework for the review. The framework includes five distinct models plus two variations for viewing principal impact: direct-effects, moderated-effects, antecedent-effects, mediated-effects, reciprocal-effects models. Each suggests a different theoretical perspective for viewing the principal's leadership role in school effectiveness. In addition, the models require different types of analytic techniques for empirical investigation.

When applied to the studies, we found that the most frequently represented models were the mediated-effects, direct-effects, and combined antecedent-effects models. A few studies could be viewed as falling within the moderated-effects category, but this distinction was not always clear. Despite the intuitive logic in support of the reciprocal-effects model, no studies had been explicitly designed to test for these potentially important effects.

One reason for the paucity of reciprocal-effects studies seemed to be the absence of longitudinal data. Such data enhance our ability to validly test reciprocal-effects models. For example, principal leadership can be hypothesized both to depend on features of the school's culture (e.g., teacher resistance to change) and also to shape it (e.g., by promoting collaboration). While reciprocal relationships can be implied at one point in time (i.e., in cross-sectional data), these dynamic relationships are best observed over time. Common approaches to this problem are to pool cross-sectional data or to use time series (panel) data. Unfortunately, longitudinal data on principal impact appears to be in short supply. Given the theoretical importance of the issues that flow from viewing principal leadership effects as reciprocal rather than uni-directional, we see this as a prime target for future study (also see Bridges, 1982; Rowan et al., 1982).

A second pattern illuminated by the classification scheme was the clear trend over time away from simple direct-effects models and towards the use of more complex, comprehensive conceptualizations of the principal's leadership role. In the past half-dozen years, the most popular approach used to study principal impact has been the antecedent/mediated-effects model.

As noted above, this model requires researchers to be more explicit and comprehensive in their conceptualization of leadership effects. It also requires more sophisticated analytical methods. This trend suggests definitive progress in the field. This is a very different story than was found in the prior generation of research as reported by Bridges in 1982.

The framework for classification also yielded a striking and rather unexpected result concerning the role of the school principal in school effectiveness. When the studies were grouped in terms of these theoretical models, a clear trend emerged in the direction of their substantive results. The studies that utilized more sophisticated theoretical models yielded more consistently positive findings concerning the positive impact of the principal on school outcomes than did the less sophisticated studies (see Table 1). The studies reporting evidence of principal effects tended to fall into the antecedent/mediated-effects category.

This model hypothesizes that principal leadership is simultaneously a dependent and independent variable. The principal's leadership is exercised in response to features of the school organization and its environment and is aimed towards influencing internal school processes that are more directly linked to student learning. These internal processes range from school policies and norms to the practices of teachers.

Studies based on this model frequently uncovered positive *indirect* effects of principal leadership on student achievement. The finding of positive indirect effects was particularly significant because that such studies put the leadership construct to a more rigorous test than is posed by direct-effects studies. While the findings across studies were neither uniform in direction nor overly powerful in effects, the trend was clear.

Moreover, the studies that reported positive indirect effects of principal leadership on student achievement consistently found those effects acting on the school organization through its goals. Principal leadership that was geared towards the development and sustenance of a school-wide purpose or focus seemed to make a difference for student learning. That said, additional investigation is needed that explores this interaction in more detail.

This substantive finding reinforces our view of simultaneous progress on the dual fronts of theory and methodology. The finding that principal effects, when they occur, are indirect in nature is, in fact, conceptually consistent with accepted notions of how leadership is actually exercised in educational organizations (Cuban, 1988). However, in the absence of theoretical models that force empirical tests to include mediating variables, we are left with the conceptually weak and empirically ambiguous findings that necessarily result from studies that employ a direct-effects model.

Thus, our first task in this review was a fruitful one. The framework used to guide our analysis of the studies (Pitner, 1988) illuminated important theoretical and methodological dimensions of the research in this domain. Moreover, it highlighted the critical interplay between theory, method and results so important in the conduct of valid research. In the next section, we summarize the main findings with respect to the specific methodological issues that emerged from the review.

Conceptual Progress

As a group, the studies demonstrate several conceptual advances over the prior generation of principal impact studies. First, whereas Bridges (1982) noted an absence of theoretically-oriented research, we found a distinctly different trend among these studies. From our reading, almost all of the studies could be counted as theoretically informed. The authors were uniformly explicit about the particular lineage in which their studies were located. They also tended to be quite careful in defining their constructs, particularly the leadership variable.

A more strict definition of theoretical orientation would require the studies to link their empirical efforts to larger conceptual frameworks and issues. It would also require the authors to use theory to guide both selection of variables and placement within hypothesized models. By this standard, well over half of the studies reviewed in this data-set were quite sophisticated in terms of their theoretical orientation. Particularly within the C and C₁ groups, authors took pains to discuss, in advance, how the leadership construct was theoretically linked to the intervening variables and student outcomes (e.g., Goldring & Pasternak, 1994; Hallinger et al., 1990, in

press; Heck et al., 1990; Jones, 1987; Leitner, 1994; Leithwood et al., 1994; Silins, 1994; Weil et al., 1984). This theoretical groundwork proved critical in light of the non-experimental research designs that predominate in studying the principal's role in school effectiveness.

The review further demonstrates the importance of beginning with a theoretically informed model of leadership and how it influences school outcomes. If the impact of principal leadership is achieved through indirect means -- for example through the school's culture -- we need to advance our understanding of how such linkages and norms are shaped by the principal. The studies offer some limited guidance as to the types of intervening variables that may potentially yield fruit. Yet, more work remains to be done in uncovering the nature of this particular path.

We would like to step back and make several recommendations concerning the conceptualization of the principal's role in school effectiveness. First, we strongly recommend that researchers continue using *comprehensive* models of leadership in studies of principal effects. The abbreviated direct-effects model represented in Models A and A₁ simply cannot be defended in light of current theory. Given the availability of appropriate analytical techniques such as structural equation modeling, researchers would be better served by building upon the model C₁ studies conducted to date.

Second, we would argue for researchers to incorporate more comprehensive notions of administrator outcomes into their conceptual models in future studies in this domain. Our own review would appear to "buy into" the notion that student achievement is the most desirable measure of principal effectiveness. In fact, while we remain committed to understanding how principals influence student learning, achievement tests only represent one proxy for this outcome.

Thus, we would call for more diverse conceptualizations of the goals of administrative behavior. The current group of studies begins to suggest other outcomes of principal leadership that may be worthwhile exploring. For example, studies conducted by Leithwood and colleagues (1993) and by Silins (1994) analyzed effects on a wide range of school improvement variables that should be of interest to scholars of administrator effectiveness. These included program

implementation, teacher professional development, and school improvement indices. Where possible, we would suggest that these types of measures be used to supplement direct measures of student learning.

Although it was not a focus for the review, we would be remiss in not returning to an important and related issue that has lurked in the shadows of this chapter. Earlier we noted that this set of studies reflects an exclusively functionalist and instrumental view of the principal's leadership role. This is linked to the criteria we used in defining the purpose of the review: to examine research on the impact of principals on student achievement and school effectiveness. This approach to studying administrator effectiveness reflects a dominant perspective in our educational culture as well as in the field of leadership research (Anderson, 1990; Blase, Dedrick, & Strathe, 1986).

This conceptual lens for viewing principal impact narrowed the issue, thereby allowing us to conduct a technically sound literature review. At the same time, it also clouds potentially important discourse about the normative purposes of administrative behavior. We have seen within this body of literature how conceptual progress can "work its way into" empirical investigation over a period of time. Thus, we would explicitly urge researchers to undertake studies that are conceived from alternative philosophical frameworks that are gaining currency in this era (e.g., critical and feminist theory, postmodernism, chaos and complexity theory).

Finally, although the studies included in this review were conducted in a variety of countries and cultures, the conceptualizations of leadership were all based on Western notions of how leadership is exercised in organizations. This covers up an assumption that characterizes empirical research as well as theory in educational administration. There has been surprisingly little research that is either cross-cultural in nature or that employs indigenous conceptions of leadership in non-Western cultures (Hallinger, 1995).

This is a glaring shortcoming that needs to be addressed. Such research will need to take into account not only potentially different conceptualizations of leadership, but also different views on the desired outcomes of leadership in other cultures. The serious consideration of non-

Western conceptions of leadership and effectiveness have the potential to open our eyes to very different theoretical treatments of this domain. We believe that this will be an increasingly important area of inquiry in future years.

Methodological Progress

Methodological progress reflects a similarly positive trend over time, though some important and quite fundamental issues remain to be addressed. Foremost among our concerns is the continued over-reliance on a single type of research design: non-experimental research designs (primarily cross-sectional survey research). Cross-sectional designs -- even ones of high quality -- limit our ability to understand the causal relationships involved in studying the impact of school administrators. Interpretation of data from studies of principal impact continues to be further hindered by the absence of longitudinal, experimental, quasi-experimental and qualitative research. We encourage future studies employing greater diversity in research design.

In particular, this should include both quasi-experimental designs and qualitative approaches. We found none of the former in our search of the literature. With respect to qualitative studies, several were uncovered in our initial search, though fewer than we might have expected. We would also argue for more mixed method and two-stage studies. In the latter approach, the researcher engages the basic question of administrator effects issues at a broad level of study through quantitative analysis and then focuses on specific issues through more flexible qualitative methods (e.g., Hallinger & Murphy, 1986b; Jackson, 1982; Leitner, 1994). We see this as a potentially fruitful means of uncovering the more subtle processes that underlie expertise in leadership behavior (e.g., Dwyer et al., 1983; Leithwood et al., 1992).

We would also note that most of the studies included in this data-set tended to focus on the impact of elementary school principals. This was not a function of predetermined choice on our part; rather, it resulted from the relative paucity of studies of principal impact being conducted at the secondary level. This is an issue of both theoretical and practical importance since secondary schools differ from elementary schools on important contextual dimensions such as size and complexity. Theory, as well as preliminary research in high schools, suggest that these

contextual variables influence both how principals exercise leadership and the results of their leadership (e.g., Firestone & Herriott, 1982; Hallinger & Murphy, 1986a; Heck, 1993; Jones, 1987; Saavedra, 1987). While findings from elementary level studies provide clues as to the direction researchers may take at the secondary level, additional theoretical and empirical work is needed to describe and account for the impact of these contextual differences.

Another issue concerns how studies are conducted and analyzed. With respect to quantitative studies, sampling remains a problematic area. Probability sampling and adequate sample size were highlighted as key sampling issues of concern. Obtaining an adequate sample size is necessary if we hope to detect effects in this domain where we expect measurable impact to be relatively weak. As we noted, probability sampling is central to non-experimental research. Most of the studies we reviewed fell far short of ideal sampling conditions both on sample size and the nature of the sampling procedures. Improvement is needed in this area.

In the realm of instrumentation, we found clear progress on two fronts. First, researchers showed increased concern and care in assessing and describing the characteristics of their measurement instruments. Many of the studies that we reviewed included some discussion of instrument reliability, at least reporting alpha coefficients or other relevant statistics if not the details of obtaining them. Second, it appears that several instruments have emerged over the past decade with a reasonable track record for use in studies of school administrators. These cover several constructs including instructional leadership, transformational leadership, as well as several perspectives on leadership derived from the work of Bolman and Deal (1992) and Sergiovanni (see Cheng, 1994).

At the same time, while the reliability of measurement scales seems to be less of an issue today than a dozen years ago, researchers must attend with greater care to assessing the validity of their measurement instruments. We illustrated the basis for this concern in pointing out how an instrument that demonstrated high degrees of validity at the elementary failed to replicate similar validation at the secondary level (Jones, 1987). More complicated sets of issues arise in attempting to conduct studies of principal leadership in diverse cultural settings. This instance

serves as a cautionary note for researchers to exercise greater attention to validity as well as reliability in the development and use of their research instruments.

To extend the process of validation further, we noted that although these studies were conducted in a variety of countries, the philosophical perspectives and methodological approaches undergirding them has not really opened the field up to solid comparative research. The bulk of the studies were conducted in contexts where Western forms and systems of education predominate. Moreover, even when research has been conducted outside of Western contexts, Western conceptions of leadership and schooling have, in most cases been overlaid onto the other culture.

We would emphasize here that a more culturally diverse orientation to the study of leadership and its effects also has quite specific methodological implications. First it suggests the need to begin with a culturally appropriate definition of leadership. This may be gained most effectively through the use of qualitative methods that stress the inductive generation of culturally grounded theory. The actual investigation of leadership effects from a cross cultural perspective raises a number of interesting methodological issues including obtaining samples across diverse settings, validation of instrumentation, and using analytical techniques that are appropriate for multi-group samples (e.g., structural equation modeling).

The range of analytical techniques for addressing complex research questions and theoretical models has expanded greatly over the past 25 years. At the same time, however, researchers in the field of educational administration have not always taken full advantage of the techniques available (Bridges, 1982; Tatsuoka & Silver, 1988; Willower, 1987). When new analytical techniques have been utilized, researchers have not always fully understood the implications of those approaches for the pursuit of knowledge. As our review bears out, sometimes researchers have applied the wrong analytical techniques to the right theoretical models; at other times they have applied the right techniques to the wrong models.

We believe that researchers of administrative activity will also profit greatly from adopting a multi-level perspective towards schools as organizations (Bossert et al., 1982; Rowan et al.,

1991). Treating data within its hierarchical structure may assist in building theory about the nature of administrator impact across levels of the organization. It will also facilitate more refined investigations into a wider variety of theoretical perspectives on how impact is obtained in different types of organizational structure (e.g., restructured schools).

That said, we can also conclude that substantial progress is evident on the analytical front. Even among the less sophisticated Model A and A₁ studies, we found an increased use of control variables, whose relative scarcity was bemoaned in Bridges' (1982) review. Moreover, we also observed a distinct trend towards the use of increasing levels of sophistication in the data analysis techniques in more recent years.

Comprehensive modeling of principal effects requires more powerful analytical tools than have been applied in the past. Several new approaches (e.g., structural equation modeling, hierarchical linear modeling) show particular promise in terms of their appropriateness for addressing key conceptual issues in this domain of research. The frequency of use of these approaches has increased dramatically since 1990 and bodes well for future research in this area.

As our reanalyses of others' data demonstrated, however, the use of these tools is double-edged. It is important to keep in mind that strong theoretical explication must be used to guide the specification of models when applying these techniques or it becomes very easy to fall into the trap of drawing incorrect or over-stated interpretation of the findings. When using these tests, the failure to reject one model is not an indication that there are no other models that could fit the data equally well or better. Thus, it is wise to consider competing structural models in light of theoretical propositions and previous empirical work (see Leithwood et al., 1993; Silins, 1994).

This highlights a major theme that emerged from the review. It returns us to the notion that theory and method play a mutually reinforcing role in the creation of new knowledge. This review suggests that although much work remains in filling out our understanding of the principal's role in school effectiveness, as a field we are making progress in the application of theory and methods of research in attacking this potentially important problem of practice.

Thus, in conclusion, the review supports *both* the potency of Bridges' (1982) highly critical review of methodology in educational administration and the conceptual promise hinted at by other reviewers of this literature (Bossert et al., 1982; Leithwood & Montgomery, 1982; Murphy, Hallinger, & Mitman, 1983). While readers may find the substantive conclusion of the review of interest, we believe that this finding is actually of only secondary importance. For the purposes of researchers, the most salient result is the demonstration of how substantive progress in a field can be achieved when headway occurs simultaneously on methodological and conceptual fronts. If in fifteen years, methodological and conceptual advancements of a similar magnitude can be claimed, we are confident that the field will have made much more significant headway in addressing important substantive problems of interest to practitioners, policymakers and researchers.

References

- Amick, D. & Walberg, H. (1975). Introductory multivariate analysis. Berkeley, CA: McCutchan Publishing.
- Anderson, G. 1990. Toward a critical constructivist approach to school administration: Invisibility, legitimation, and the study of non-events. Educational Administration Quarterly, 26(1) 38-59.
- Andrews, R. & Soder, R. (1987). Principal instructional leadership and school achievement. Educational Leadership, 44, pp. 9-11.
- Andrews, R., Soder, R., & Jacoby, D. (1986, April). Principal roles, other in-school variables, and academic achievement.
- Bamburg, J. & Andrews, R. (1990). School goals, principals and achievement. School Effectiveness and School Improvement, 2(3), pp. 175-191.
- Benham, M. & Heck, R. (1994). Political culture and policy in a state-controlled educational system: The case of North Carolina.
- Biester, T., Kruse, J., Beyer, F., & Heller, B. (1984, April). Effects of administrative leadership on student achievement.
- Blank, R. (1987). The role of the principal as leader: Analysis of variation in leadership in urban high schools. Journal of Educational Administration, 25(1), pp. 1-11.
- Blase, J., Dedrick, C., & Strathe, M. (1986). Leadership behavior of principals in relation to teacher stress, satisfaction, and organizational commitment. Journal of Educational Administration, 24(1), pp. 1-11.
- Bolman, L. & Deal, T. (1992). Reframing organizations. San Francisco, CA: Jossey-Bass.
- Bossert, S., Dwyer, D., Rowan, B., & Lee, G. (1982). The instructional management role of the principal. Educational Administration Quarterly, 18(3), pp. 34-64.
- Boyan, N. (1988). Describing and explaining administrative behavior. In N. Boyan (Ed.), Handbook of research in educational administration. New York: Longman.
- Broughton, R. & Riley, J. (1991). The relationship between principals' knowledge of reading processes and elementary school reading achievement. ERIC: ED341952.
- Brewer, D. (1993). Principals and student outcomes: Evidence from U.S. high schools. Economics of Education Review, 12(4), pp. 281-292.
- Bridges, E. (1970). Administrative man: Origin or pawn in decision-making? Educational Administration Quarterly, 16(1), pp. 1-11.
- Bridges, E. (1982). Research on the school administrator: The state-of-the-art, 1967-1980. Educational Administration Quarterly, 18(3), pp. 12-33.
- Bridges, E. (1977). The nature of leadership. In L. Cunningham, W. Hack, & R. Nystrand (Eds.), Educational Leadership: A Review of Research and Practice. New York: Praeger.

- Brookover, W. & Lezotte, L. (1977). Changes in school characteristics coincident with changes in student achievement. New York: Harper & Row.
- Brookover, W., Schweitzer, J., Schneider, J., Beady, C. Flood, J., & Wisenbaker, J. (1978). Elementary school characteristics and student achievement. New York: Harper & Row.
- Bryk, A. & Raudenbusch, S. (1992). Hierarchical linear models: Applications and data analysis methods. New York: Guilford Press.
- Burns, J. M. (1978). Leadership. NY:Harper & Row.
- Cantu, M. (1994). A study of principal instructional leadership behaviors manifested in successful and nonsuccessful urban elementary schools. Unpublished doctoral dissertation, University of Tennessee.
- Cattell, R. (1962). The basis of recognition and interpretation of factors. Educational and Psychological Measurement, 22, 667-669.
- Cheng, Y. C. (1991). Leadership style of principals and organizational process in secondary schools. Journal of Educational Research, 84(2), 107-114.
- Cheng, Y.C. (1994). Principal's leadership as a critical factor for school performance: Evidence from multi-levels of primary schools. School Effectiveness and School Improvement, 5 (3), pp. 299-317.
- Cohen. M.R. (1988). Statistical power for the behavioral sciences (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crawford, J., Kimball, G., & Watson, P. (1985). Causal modeling of school effects on achievement. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Cronbach, L. (1971). Test validation. In R. Thorndike (Ed.) Educational measurement (2nd edition). Washington, DC: American Psychological Association.
- Cronbach, L. & Meehl, P. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 28-35.
- Cuban, L. (1988). The managerial imperative and the practice of leadership in schools. Albany: Suny Press.
- Davies, R. (1994). From cross-sectional to longitudinal analysis. In A. Dale & R. Davies (Eds.), Analyzing social research. London: Falmer Press.
- Dilworth, R. (1987). A study of the relationship between student achievement and the variables of teacher-perceived instructional leadership behaviors of principals and teacher attendance. Unpublished doctoral dissertation, University of Tennessee.
- Duke, D. (this volume). Perception, prescription and the future of school leadership. In K. Leithwood (Ed.). The international handbook of research in educational leadership and administration. London: Falmer Press.
- Dwyer, D. (1986). Understanding the principal's contribution to instruction. Peabody Journal of Education, 63(1), pp. 3-18.
- Dwyer, D., Lee, G., Rowan, B., & Bossert, S. (1983). Five principals in action: Perspectives on instructional management. San Francisco: Far West Laboratory for Educational Research and Development.

Eberts, R. & Stone, J. (1988). Student achievement in public schools: Do principals make a difference? Economics of Education Review, 7(3), pp. 291-299.

Edington, E. & Dibenedetto, R. (1988). Principal leadership styles and student achievement in small and rural schools of New Mexico. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April.

Edmonds, R. (1979). Effective schools for the urban poor. Educational Leadership, 37, pp. 15-24.

Eisner, E. (1993). Forms of understanding and the future of educational research. Educational Researcher, 22

Everhart, R. (1988). Fieldwork methodology in educational administration. In N. Boyan (Ed.) The handbook

Glasman, N. (1983). Increased centrality of evaluation and the school principal.

Administrator's Notebook, 30(1), pp. 1-4.

Glasman, N. (1984). Student achievement and the school principal. Educational Evaluation and Policy Analysis, 6(3), pp. 283-296.

Glasman, N. & Binianimov, I. (1981). Input-output analyses of schools. Review of Educational Research, 51(4), pp. 509-539.

Glasman, N. & Fuller, J. (1992). Assessing the decision-making patterns of school principals. The International

Glasman, N. & Heck, R. (1992). The changing leadership role of the principal: Implications for principal assess

Glasman, N. & Heck, R. (this volume). Role-based evaluation of principals: Developing an appraisal system

Goldring E. & Pasternak, R. (1994). Principals' coordinating strategies and school effectiveness.

School Effectiveness and School Improvement, 5 (3), pp. 239-253.

Hallinger, P. (Spring, 1995). Culture and leadership: Developing an international perspective in educational administration. UCEA Review.

Hallinger, P. (1992). Changing norms of principal leadership in the United States. Journal of Educational Ac

Hallinger, P. (1983). Principal instructional management rating scale. Palo Alto: Stanford University.

Hallinger, P., Bickman, L., & Davis, K. (1990). What makes a difference? School context, principal leadership and student achievement. Occasional paper #3. The National Center for Educational

Hallinger, P., Bickman, L., & Davis, K. (in press). School context, principal leadership and student achieve

Hallinger, P. & Heck, R. (in press). Reassessing the Principal's Role in School Effectiveness:

A Review of Empirical Research, 1980 -1995. Educational Administration Quarterly.

Hallinger, P. & Murphy, J. (1985). Assessing the instructional management behavior of principals. Elementary School Journal, 86(2), pp. 217-247.

Hallinger, P. & Murphy, J. (1986a). Instructional leadership in school contexts. In W.

Greenfield (Ed.), Instructional leadership: Concepts, issues and controversies. Lexington, MA: Allyn & B

Hallinger, P. & Murphy, J. (1986b). The social context of effective schools. American Journal of

Education, 94(3), pp. 328-355.

Heck, R. (1993). School context, principal leadership, and achievement: The case of secondary schools in Singapore. The Urban Review. 25(2), pp. 151-166.

Heck, R. (1992). Principal instructional leadership and the identification of high- and low-

achieving schools: The application of discriminant techniques. Administrator's Notebook, 34(7), pp. 1-4.

Heck, R. & Marcoulides, G. (1992). Principal assessment: Conceptual problem, methodological problem, or both?

Heck, R. & Marcoulides, G. (1989). Examining the generalizability of administrative allocation decisions. The

Heck, R., Larson, T., & Marcoulides, G. (1990). Principal instructional leadership and school

achievement: Validation of a causal model. Educational Administration Quarterly, 26, pp. 94-125.

Heck, R., Marcoulides, G., & Lang, P. (1991). Principal instructional leadership and school

achievement: The application of discriminant techniques. School Effectiveness and School Improvement, 2

Hunter, C. (1994). Los Angeles Unified School District Middle School principals' instructional leadership behaviors

Jackson, S. (1982, April). Instructional leadership behaviors that differentiate effective and

ineffective low income urban schools. Paper presented at the convention of the International Reading Association

Jones, P. (1987). The relationship between principal behavior and student achievement in

Canadian secondary schools. Unpublished doctoral dissertation, Stanford University, Palo Alto, CA.

Kaplan, A. (1964). The conduct of inquiry: Methodology for behavioral science. San Francisco: Chandler.

Krug, F. (1986). The relationship between the instructional management behavior of

elementary school principals and student achievement. Unpublished doctoral dissertation, U. of San Francisco

Leitner, D. (1994). Do principals affect student outcomes? An organizational perspective.

School Effectiveness and School Improvement, 5 (3), pp. 219-239.

Leithwood, K. (1994). Leadership for school restructuring. Educational Administration

Quarterly, 30(4), pp. 498-518.

Leithwood, K., Begley, P., & Cousins, B. (1992). Developing expert leaders for future schools.

Bristol, PA: Falmer Press.

Leithwood, K., Begley, P. & Cousins, B. (1990). The nature, causes and consequences of

principals' practices: A agenda for future research. Journal of Educational Administration, 28(4), pp. 5-31

Leithwood, K. & Hallinger, P. (1993, Fall). Cognitive perspectives on educational

administration: An introduction. Educational Administration Quarterly, 24 (3), pp. 296- 301.

Leithwood, K. & Jantzi, D. (1990). Transformational leadership: How principals can help

reform school cultures. School Effectiveness and School Improvement, 1, (1), pp. 249-280.

Leithwood, K., Jantzi, D., Silins, H., & Dart, B., (1993). Using the appraisal of school leaders as an instrument for

Leithwood, K. & Montgomery, D. (1982). The role of the elementary principal in program

improvement. Review of Educational Research, 52(3), pp. 309-339.

Marcoulides, G. & Heck, R. (1992). Assessing instructional leadership effectiveness with "g" theory. International

Marcoulides, G. & Heck, R. (1993). Organizational culture and performance: Proposing and testing a model

Mueller, J., Schuessler, K. & Costner, H. (1977). Statistical reasoning in sociology (3rd edition). Boston: Houghton

Murphy, J. (1988). Methodological, measurement and conceptual problems in the study of

instructional leadership. Educational Evaluation and Policy Analysis, 10(2), pp. 117-139.

Murphy, J., Hallinger, P., & Mitman, A. (1983). Research on educational leadership: Issues

to be addressed. Educational Evaluation and Policy Analysis, 5(3), pp. 297-305.

National Council on Measurement in Education. (1984). Standards for educational and

psychological testing. Washington, DC: American Psychological Association.

O'Day, K. (1983). The relationship between principal and teacher perceptions of principal

instructional management behavior and student achievement. Unpublished doctoral

dissertation, Northern Illinois University, Normal, Illinois.

Ogawa, R. (1992). Institutional theory and examining leadership in schools. International Journal of Educational Management, 6(3), 14-21.

Ogawa, R. & Hart, A. (1985). The effect of principals on the instructional performance of schools. Journal of Educational Administration, 22(1), pp. 59-72.

Pedhazur, E. & Schmelkin, L. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale: NJ: Lawrence Erlbaum Associates.

Pitner, N. (1988). The study of administrator effects and effectiveness. In N. Boyan (Ed.), Handbook of research in educational administration. New York: Longman.

Ramey, M., Hillman, L., & Matthews, T. (1982, March). School characteristics associated with instructional effectiveness. Paper presented at the annual meeting of the American Educational Research Association, New York.

Rowan, B. & Denk, C. (1984). Management succession, school socioeconomic context and basic skills achievement. American Educational Research Journal, 21(3), pp. 17-537.

Rowan, B., Dwyer, D., & Bossert, S. (1982). Methodological considerations in the study of effective principals. Paper presented at the annual meeting of the American Educational Research Association.

Rowan, B. Raudenbush, S. & Kang, S. (1991). Organizational design in high schools: A multilevel analysis. American Journal of Education, 99(2), pp. 238-266.

Ruczieska, J.K. (1988). The relationships among principals' sense of efficacy, instructional leadership and school improvement.

Saavedra, A. (1987). Instructional management behaviors of secondary school administrators. Unpublished masters dissertation, Bukidnon State College, Malaybalay, Bukidnon, Malaysia.

Scott, C. & Teddlie, C. (1987, April). Student, teacher, and principal academic expectations and attributed responsibility.

Silins, H. (1994). The relationship between transformational and transactional leadership and school improvement.

Sirois, H. & Villanova, R. (1982). Theory into practice: A theoretical and research base for the characteristics of effective principals.

Tatsuoka, M. & Silver, P. (1988). Quantitative research methods in educational administration. In N. Boyan (Ed.), Handbook of research in educational administration.

Teddlie, C., Falkowski, C., Stringfield, S., Desselle, S., & Garvue, R. (1983, April). The study of principal and teacher inputs to student achievement. Paper presented at the annual meeting of the American Educational Research Association.

van de Grift, W. (1990). Educational leadership and academic achievement in elementary education. School Effectiveness and School Improvement, 1(3), pp. 26-40.

van de Grift, W. (1989). Self perceptions of educational leadership and mean pupil achievements. In D. Reynolds, B.P.M. Creemers, & T. Peters (Eds.). School effectiveness and improvement. Cardiff/Groningen: School of Education/RION, pp. 227-242.

van de Grift, W. (1987). Zelfpercepties van onderwijskundig leiderschap. In F.J. Van der Krogt (Ed.). Schoolle

Villanova, R., Gauthier, W., Proctor, P., & Shoemaker, J. (1981). The Connecticut school effectiveness q

Weil, M., Marshalek, B. Mitman, A., Murphy, J., Hallinger, P., & Pruyn, J. (1984, April). Effective and t

Willower, D. (1987). Inquiry into educational administration: The last twenty-five years and the next. Journal o

Philip Hallinger is Professor at Vanderbilt University and Chiang Mai University, Thailand. His research focuses on principal leadership, the professional development of school leaders, and international issues in school improvement. His recent work includes publications on the use of problem-based learning in educational administration and on the changing role of principals in efforts to decentralize education in the United States and abroad.

Ronald Heck is Associate Professor at the University of Hawaii, Manoa. His research interests include the role of the principal and educational policy and politics. His recent publications focus on administrative socialization, educational assessment, and the effects of state political culture on educational policy.

Acknowledgement

The authors would like to acknowledge useful comments made on earlier drafts of this manuscript by Edwin Bridges, Larry Cuban and Ken Leithwood.

ⁱ. Two issues arise here. The first involves control of the dependent variable. Some studies include control variables designed to account for other influences on students achievement such as student SES or students' prior achievement. In other cases, the control variables may also be manipulated as a means of studying the influence of context on the principal's leadership. For our purposes, inclusion of control variables that are used exclusively in connection with the dependent outcome measure is indicated in Table 1 with an asterisk. When the control variable is also used to inform our understanding of antecedent-effects on leadership, this is classified in either Model A₁ or C₁ in Table 1 and Figure 1.

ii. Van de Grift's 1987 and 1989 studies were analyzed based upon information provided in van de Grift, 1990, not from the reports of the original studies.