Measurement Properties of the Principal Instructional Management Rating Scale

Technical Report Ver. 5.11
November 2, 2013
Professor Philip Hallinger[1]

[1] Portions of this research were his conducted with support from the University Grants Commission of Hong Kong through the General Research Fund # GRF 841711. The author may be contacted at Hallinger@gmail.com

Table of Contents

Introduction

This report describes the development of the PIMRS, a rating instrument for appraising the instructional leadership behavior of school principals. The initial goal in defining the PIMRS was to develop an instrument that met the following requirements:

1. The instrument would focus on specific job related behaviors of school principals that concerned leading and managing teaching and learning in schools.
2. The behavioral components of the instrument would be drawn from research related to principal effectiveness as well as from current practice.
3. The instrument would be useful for a variety of purposes including principal evaluation, staff development, research, and district policy analysis.

Over the ensuing years since the development of the original form of the scale in 1982 (Hallinger, 1983; Hallinger & Murphy, 1985), the PIMRS has been used in more than 175 studies (i.e., as of early 2013; see Appendix A). The methodologies used in these studies were reviewed in Hallinger, 2011a). In 2011-12, the author undertook a review of the scale's reliability (Hallinger, Wang, & Chen, 2012, 2013) and refinement of the scale as detailed in this Technical Report. This resulted not only in a more comprehensive picture of the scale's measurement properties based upon numerous studies, but also resulted in the development of a short form of the scale for use with teachers (see Chapter 7).

As indicated in the Table of Contents this report has the following purposes:

1. *Chapter 1* provides background information on the historical development of instructional leadership and purpose and organization of this report.
2. *Chapter 2* introduces the conceptual framework that underlies the development of the PIMRS instrument. The framework consists of 3 dimensions and 10 instructional leadership functions.
3. *Chapter 3* focuses more specifically on identifying the steps employed in development of the PIMRS instrument.
4. *Chapter 4* examines the reliability of the instrument. It first discusses the concept of reliability. Then it presents data on the original reliability study (Hallinger, 1983; Hallinger & Murphy, 1985). Finally, it presents the results of the updated meta-analysis of reliability studies conducted in 2012 (Hallinger et al., 2012, 2013). These are the reliability estimates that users should refer to in their reports.
5. *Chapter 5* examines the validity of the scale. Again, the results of the original validation study are presented, followed by an updated assessment of validity conducted in 2012.
6. *Chapter 6* reports on the development of a Teacher Short Form of the PIMRS. The basic form of the PIMRS has consisted of 50 items measuring 3 dimensions and 10 instructional leadership functions (see Chapters 2 and 3). This form has traditionally been used with all three forms (i.e., teacher, principal, supervisor) of the PIMRS. However, our recent studies of the instrument have enabled the development of a shorter version consisting of 23 items that can be used with teachers while retaining a high rate of reliability. The Teacher Short Form, however, only yields information on the 3 instructional leadership

dimensions, not the 10 functions. For many researchers, however, this is sufficient. The Principal Form remains 50 items. These issues are discussed at length in Chapter 7.

7. *Chapter 7* presents a variety of ways in which the PIMRS has been used for research and practice.
8. *Appendices A and B* contain complete lists of PIMRS studies (Appendix A) as well as a general reference list of articles, books, chapters, and conference papers related to principal instructional leadership (Appendix B).

Chapter 1

Evolution of Instructional Leadership

Among the global trends in educational leadership and management that have emerged over the past 50 years, few have been more significant, widespread or persistent than the focus on understanding linkages between school leadership and learning (Bell, Bolam, & Cubillo, 2003; Bridges, 1967; Gross & Herriot, 1965; Hallinger & Heck, 1996a, 1996b; Leithwood, Anderson, Mascall & Strauss, 2011; Robinson, Lloyd & Rowe, 2008; Witziers, Bosker, & Kruger, 2003). The "elusive search" (Witziers et al., 2003) for understanding the nature of leadership that makes a difference for student learning has engaged scholars in studying a wide variety of leadership models. These include instructional leadership (e.g., Bossert, Dwyer, Rowan & Lee, 1982; Hallinger & Murphy, 1985a), transformational and transactional leadership (e.g., Leithwood & Jantzi, 2000), strategic leadership (e.g., Davies, Ellison & Bowring-Carr, 2005), teacher leadership (e.g., Barth, 2001; Lambert, 2002; York-Barr & Duke, 2004), collaborative leadership (Hallinger & Heck, 2010) and distributed leadership (Spillane, 2006).

Recent research syntheses support the conclusion that, among these competing models, instructional leadership has demonstrated the greatest impact on student learning (e.g., Hallinger, 2011b; Leithwood, Day, Sammons, Harris, & Hopkins, 2006; Robinson et al., 2008). This conclusion has further enhanced the prominence of instructional leadership as a focus for policy and practice, and provides a rationale for why school personnel should focus on strengthening instructional leadership as a lever for school improvement (Hallinger, 2003; Leithwood et al., 2011; Printy, n.d.).

In this chapter, we trace the evolution of the instructional leadership construct. We begin with its emergence in the literature in the mid-20th century in the USA. The we highlight its maturation during the 1980s with the advent of the effective schools movement. During this decade, instructional leadership held the high ground as the most influential leadership model in the educational leadership literature, at least in the USA. Then we examine the waxing, waning, and eventual transcendence of this leadership model over the ensuing 30 years up to the present.

Historical Background: 1950s to 1980

Instructional leadership is a practice-based rather than a theory-driven construct, with wide, if not deep roots in American education (Bridges, 1967; Lipham, 1961; Uhls, 1962). More than 50 years ago, James Lipham (1961) asserted that effective principals were associated with effective schools. During the early and middle years of the 20th century, practical wisdom shared by principals, school superintendents, teachers and parents in the United States conveyed the belief that 'good schools have good principals' (e.g., Grobman & Hynes, 1956; Gross & Herriott, 1965; Lipham, 1961; Miller, 1960; Stuart, 1950; Tyack & Hansot, 1982; Uhls, 1962). The definitive example was a book published by James Lipham in the early 1960s, aptly titled, *Effective Principal, Effective School* (Lipham, 1961). As stated by Lipham:

> In summarizing findings on the principal's role in the school, this monograph assumes that the principal is a pivotal figure in the school and is the one who most affects the quality of teacher performance and student achievement. The author concludes that the studies

> reviewed demonstrate that the principal is a key factor in the success of the school. (Lipham, 1961, p. 3)

Yet, despite formulating this conclusion, Lipham (1961) also acknowledged that the 'studies' on which his conclusions were based consisted largely of opinion surveys and case studies rather than 'scientific' inquiry. Several years later, Edwin Bridges' offered a more pointed critique of the practical wisdom of the times.

> Of the seven major task areas for which principals have responsibility, curriculum and instruction has generated the most sound and fury. On the one hand, the principal has been exhorted to exert instructional leadership, while on the other hand, he has been told flatly that such a role is beyond his or any other human being's capacity. The problem with these disputations is that the exponents of a given position have neither defined sharply what is signified by the concept of instructional leadership nor made their assumptions explicit. (Bridges, 1967, p.136)

Consequently, this practical wisdom, though widely accepted, lacked anything approaching a sound empirical knowledge base. It was, therefore, unable to offer reliable guidance for policymakers, educators of leaders, or school leaders themselves (Bridges, 1967, 1982; Erickson, 1967). Nonetheless, a perusal of the professional and scholarly literatures of the ensuing era suggests that support for this practical wisdom continued unabated (Bridges, 1982).

1980 to 1990s: Instructional Leadership in Effective Schools

The next significant point in the historical evolution of this construct came at the dawn of the effective schools era in the USA in the early 1980s (Edmonds, 1979). Researchers studying instructionally effective schools identified 'strong instructional leadership' by the principal' as a hallmark of effective urban elementary schools in the United States (Bossert et al., 1982; Edmonds, 1979; Purkey & Smith, 1983). Although this conclusion found a ready reception among American policymakers, there were significant limitations in the research designs employed in these studies.

Consequently, the research finding of "strong instructional leadership by the principal" continued to yield considerable ambiguity concerning both the nature of the role as well as its contribution to school improvement (Barth, 1986; Barth & Deal, 1982; Bossert et al., 1982; Cuban, 1984; Leithwood & Montgomery, 1982; Murphy, Hallinger & Mitman, 1983; Rowan, Bossert & Dwyer, 1983). Instructional leaders were described as strong, directive leaders who had been successful at "turning their schools around" (Bamburg & Andrews, 1990; Bossert et al., 1982; Edmonds, 1979; Hallinger & Murphy, 1985a, 1985b, 1986). There were relatively few descriptions of effective instructional leaders working in typical schools. Yet schools differ widely in terms of their needs, resources as well as in the type of leadership required to move them forward.

Instructional leaders were viewed as culture builders. They sought to create an "academic press" that fostered high expectations and standards for students, as well as for teachers (Barth, 1990, 2002; Bossert et al., 1982; Mortimore, 1993; Glasman, 1984; Hallinger et al., 1996; Hallinger & Murphy, 1985a, 1985b, 1986; Heck et al., 1990; Purkey & Smith, 1983). Notably, instructional

leaders were viewed as a minority of principals who somehow managed to overcome the multiple pressures that push principals away from curriculum, instruction and the classroom.

Instructional leaders were goal-oriented. As leaders they were able to define a clear direction for the school and motivate others to join in its achievement. In instructionally effective schools, this direction focused primarily on the improvement of student academic outcomes (Bamburg & Andrews, 1990; Glasman, 1984; Goldring & Pasternak, 1994; Hallinger & Murphy, 1986; Heck et al., 1990; Leithwood, Begley & Cousins, 1990; Leitner, 1994; O'Day, 1983). Vision, goals, and mission became strongly situated in the vocabulary of principals who wished to succeed in the evolving environment of school reform.

The effective instructional leader was able to align the strategies and activities of the school with the school's academic mission. Thus, instructional leaders focused not only on leading, but also on managing. Their managerial roles included coordinating, controlling, supervising, and developing curriculum and instruction (Bamburg & Andrews, 1990; Bossert et al., 1982; Cohen & Miller, 1980; Dwyer, 1986; Glasman, 1984; Goldring & Pasternak, 1994; Hallinger et al., 1996; Heck, 1992, 1993; Heck et al., 1990; Jones, 1983; Leitner, 1994).

Instructional leaders led from a combination of expertise and charisma. These were hands-on principals, hip-deep in curriculum and instruction (Cuban 1984) and unafraid of working directly with teachers on the improvement of teaching and learning (Bossert et al., 1982; Cuban, 1984; Dwyer, 1986; Edmonds, 1979; Hallinger et al., 1996; Hallinger & Murphy, 1986; Heck et al., 1990; Leithwood, Begley & Cousins, 1990).

Moreover, the wave of excitement surrounding this affirmation of the importance of principal leadership was not without its skeptics and critics. Skeptics focus on the same 'gap' between prescription and normative reality noted 20 years earlier by Bridges (1967). The skeptics included respected practitioner-scholars such as Larry Cuban (1984, 1988) and Roland Barth (1980, 1986, 1990) who evinced discomfort with placing such high expectations on the school principal. Descriptions of these instructional leaders in the effective schools literature tended towards a heroic view of their capabilities. These often spawned feelings ranging from inadequacy to guilt among the vast majority of principals who wondered why they had such difficulty fitting into this role expectation (Barth, 1986; Donaldson 2001; Marshall, 1996).

Thus, even as the spotlight on instructional leadership intensified, these 'skeptical friends' called attention to a less obvious but equally powerful set of constraints that shape the role behavior of school principals. Scholars have, for many years, described forces that draw principals away from rather than towards engagement in instructional leadership (e.g., Barth, 1990; Cuban, 1988; Goldring, Huff, May& Camburn, 2008; Horng, Klasik, & Loeb, 2010; Marshall, 1996, 2004; May, Huff, & Goldring, 2012; Murphy, Hallinger, Lotto & Miller, 1987). My own dissertation advisor, Larry Cuban, was one of these friendly skeptics. His historical analysis (Cuban, 1988) of past efforts to press principals into the instructional leadership role highlighted the many forces that combine to create a 'force-field' (Marshall, 1996) around the classroom (see also, Barth, 1980, 1986, 1990; Barth & Deal, 1982; Hallinger & Murphy, 2013; Murphy Hallinger, Lotto & Miller, 1987).

Any policy-driven effort to foster sustainable instructional leadership in American schools must take these forces into account, or accept the predictable consequences of principals who suffer from unfulfilled expectations, disappointment, guilt and burnout (Barth, 1990; Bridges, 1967; Donaldson, 2006; Horng et al., 2010; Marshall, 1996, 2004). Thus, even as America's

policymakers were poised to use principals as the engine for education reform, the skeptics worried that this would leave the principals themselves 'running on empty' (see Barth, 1986, 1990; Donaldson, 2006; Hallinger & Murphy, 2013). They questioned whether instructional leadership represented a leadership model that could be broadly applied to the principalship in all schools (e.g., Barth, 1986; Barth & Deal, 1982; Cuban, 1984, 1988).

While the skeptics focused primarily on the gap between prescription and reality of schools, another group of critics focused on 'technical limitations' of the emerging literature on instructional leadership (Bridges, 1982; Erickson, 1979; Rowan, Bossert & Dwyer, 1983; Rowan, Dwyer & Bossert, 1982; Murphy, Hallinger & Mitman, 1983). Limitations noted by reviewers of this literature included:

- Lack of clearly explicated conceptual frameworks;
- Lack of valid and reliable instrumentation for studying the role;
- Lack of theoretical models that articulated how this role influenced student learning;
- Reliance on weak research designs, ill-equipped to test for causal effects.

For example, in 1982 Bridges' review of the more general literature on educational administration of the preceding period noted the following.

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

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

During this same period, Stephen Bossert and colleagues (1982) at the Far West Lab in San Francisco published a seminal literature review that synthesized findings from empirical studies that had focused more specifically on investigating school leadership and learning. While the authors acknowledged these methodological limitations, they also claimed to see the foundation within this literature for a productive program of research targeting instructional leadership and its effects on learning.

Bossert's "instructional management framework" subsequently became a valuable lens used by other scholars for conceptualizing how leadership for learning is enacted in schools. We note that the findings from the Bossert review were largely supported by other contemporary reviews of this literature (e.g., Leithwood, Begley & Cousins, 1990; Leithwood & Montgomery, 1982; Hallinger & Murphy, 1985a; Murphy et al., 1983; Purkey & Smith, 1983). Concurrent efforts were undertaken to develop new conceptual frameworks (e.g., Bossert et al., 1982; Hallinger, Murphy,

Weil, Mesa, & Mitman, 1983) and the first research instruments (e.g., Hallinger & Murphy, 1985a; Villanova, Gauthier, Proctor, & Shoemaker, 1981) developed to assess instructional leadership.

The *Principal Instructional Management Rating Scale* was one such tool (Hallinger, 1983; Hallinger & Murphy, 1985a). Since its development in 1983, the PIMRS has been used in over 200 studies in more than 25 different countries (e.g., USA, Canada, Mexico, England, Israel, Germany, UAE, India, Pakistan, Iran, Turkey, Malaysia, Vietnam, Thailand, Australia, New Zealand, China, Portugal, Philippines, Taiwan, Maldives, Kenya, South Africa, Cameroon, Nigeria).[i] More generally, during the 1980s and 1990s a growing number of largely North American scholars began to undertake more intentionally designed empirical investigations of the principal's instructional leadership role (e.g., Andres & Soder, 1987; Bamburg & Andrews, 1990; Blasé, 1987; Braughton Riley, 1991; Brewer, 1993; Dwyer et al., 1983; Eberts & Stone, 1988; Glasman, 1984; Goldring & Pasternak, 1994; Goldring & Sullivan, 1996; Hallinger, Bickman & Davis, 1996; Hallinger & Murphy, 1985a, 1985b; Hallinger, Taraseina & Miller, 1994; Heck, 1992, 1993; Heck, Larson & Marcoulides, 1990; Howe, 1995; Jones, 1987; Krug, 1986; Leitner, 1994; Leithwood & Montgomery, 1982; Leithwood & Stager, 1989; O' Day, 1986; Pounder, Ogawa & Adams, 1995; Sheppard, 1996; Snyder & Ebmeier, 1992; van de Grfift, 1989, 1990).

These developments signaled the emergence of practice-oriented conceptions of instructional leadership as a research-based construct. Findings generated from this body of empirical research further highlighted the construct's potential for contributing to the profession's understanding of how principal leadership impacts student learning (Bridges, 1982; Hallinger & Heck, 1996a, 2011a, 2011b; Leithwood et al., 1990). As a result, by the mid-1990s, Hallinger and Heck (1996a, 1996b, 1998) observed that instructional leadership had become the most prevalent perspective adopted by researchers engaged in the study of school leadership effects in North America.

The Paradigm Wars of the 1990s

This overview of the evolution of instructional leadership highlights the linkage between the socio-political context of education and the role expectations proposed for school leaders. With the advent of school restructuring in North America during the 1990's, the notion of transformational leadership began to eclipse the popularity of instructional leadership's as a guiding vision for the principal's work in leading schools. Transformational leadership originated in studies of political leaders (Burns, 1978), and has subsequently been adapted for application in education organizations by Leithwood and colleagues (e.g., Leithwood, 1994; Leithwood & Jantzi, 1999, 2000, 2005; Leithwood & Sun, 2012; Silins, 1994). This leadership model focuses on the leader's role in inspiring others towards a collective vision of change and motivating members of the organization to towards higher levels of capacity and performance (Bass, 1985; Leithwood, 1994;).

A core feature of instructional leadership emphasized the principals direct engagement with processes concerned with the 'technical core' of the school (Bossert et al., 1982; Hallinger & Murphy, 1985a). Moreover, although the instructional leadership model highlighted the role of a collective vision, it 'assumed' that continuously improving academic performance of all students was the preeminent goal for America's schools. In contrast, transformational leadership emphasized the leader's role in vision-building and capacity development, but without any specific assumptions concerning what those goals should be (Leithwood, 1994). Moreover, transformational leadership models did not posit any direct engagement with teaching and learning

by the principals (Hallinger, 2003). Cuban (1984, 1988) referred to these contrasting leadership foci as "first-order' versus "second-order" changes.

The emergence of transformational leadership models in education not only reflected the changing reform context of schools, but also a broader recognition that the concerns evinced by the 'skeptics' of instructional leadership model had some degree of validity. Moreover, in an era of 'teacher empowerment' the focus on principals as the driver for school improvement almost seemed out of place. Transformational leadership soon began to dominate the school leadership landscape, as instructional leadership receded into the background, at least in the professional literature (e.g., Leithwood, 1994; Leithwood & Jantzi, 1999, 2000; Silins, 1994). Thus, a status report on school leadership at the turn of the millennium would have highlighted the waxing status of transformational leadership and the waning status of instructional leadership.

Continuing Evolution of Instructional Leadership: 2000 to the Present

However, around this time, the pendulum of educational goals began to shift once again. Policies embedded in No Child Left Behind and Race to the Top represent the culmination of three decades of almost continuous education reform in the USA. Starting with the *A Nation at Risk* report in the early 1980s, federal policymakers increasingly sought to direct the improvement of America's education system from Washington. A central factor mediating the success of Federal and State policy efforts at educational reform lies in the leadership capacity of the nation's school principals and teachers (Leithwood, 2001; Neumerski, 2012; Schoen & Fusarelli, 2008;). While effective leadership cannot guarantee successful education reform, research affirms that sustainable school improvement is seldom found without active, skillful instructional leadership from principals and teachers (Fullan, 2006; Hall & Hord, 2002; Hallinger, 2011; Hallinger & Heck, 2010; Leithwood, Harris & Hopkins, 2008; Leithwood, Louis, Anderson, & Wahlstrom, 2004).

In recent years, acceptance of this tenet among policymakers has refocused the spotlight on principal instructional leadership (Hallinger, 2011; Neumerski, 2012; Nettles & Herrington, 2006; Schoen & Fusarelli, 2008; Silva, White & Yoshida, 2011). This has, for example, resulted in accountability policies that mandate more comprehensive systems of teacher and principal evaluation and raised the bar in terms of standards of performance (Leithwood, 2001; Murphy & Shipman, 2003; Silva et al., 2011). Indeed, these policies require the replacement of principals (and teachers) in underperforming schools that fail to demonstrate improvement. Observers assert that these school accountability policies have transformed instructional leadership from an option into a necessity for America's school administrators (Murphy, 2008; Nettles & Herrington, 2007; Schoen & Fusarelli, 2008; Silva et al., 2011).

In the 1980s when instructional leadership emerged as a new construct, some scholars questioned both its relevance and viability as a guiding metaphor for school leadership (e.g., Barth, 1986; Cuban, 1984). Thirty years later, 'instructional leadership' and its global cousin, 'leadership for learning' are widely accepted by policymakers and practitioners as essential elements of management practice in schools. Indeed, recent reviews of research largely confirm early assertions concerning the relationship between instructional leadership and student learning (see Hallinger, 2011a; Leithwood et al., 2012; Robinson et al., 2008). Thus, contrary to early predictions, instructional leadership has demonstrated impressive staying power as a core concept guiding research, policy and practice in the field of educational leadership and management.

At the same time, however, it is interesting to note that prior to the turn of the millennium interest in instructional leadership was a largely North American phenomenon. Indeed, it is only in the last decade that the term instructional leadership and its kissing cousin, 'leadership for learning', have gained broad international currency. This is reflected in research and policy publications from the UK (Bell et al., 2003; Day, Sammons, Hopkins, Harris, Leithwood, Gu, Brown, Ahtaridou, & Kingston, 2009; Hunter Foundation, 2005; MacBeath & Cheng, 2008; Southworth, 2002), continental Europe (Krüger, Witziers, & Sleegers, 2007; Witziers et al., 2003), East Asia (Chan & Cheng, 1993; Kim, 1988; Hallinger et al., 1994; Ratchaneeladdajit, 1997; Poovatanikul, 1993; Wongtrakool, 1995) and Australia/New Zealand (Mulford & Silins, 2009; Robinson et al., 2008). This reflects the growing global interest in understanding the ways in which school leaders contribute to school improvement and student learning (Hallinger & Heck, 2011a).

This growing global interest in instructional leadership has subsequently generated an expanding body of empirical research and continuing advances in clarifying its contribution to improvements in teaching and learning (e.g., see Bryk, Sebring, Allensworth, Luppescu, & Easton, 2009; Datnow & Castellano, 2001; Hallinger & Heck, 2010; Krüger et al., 2007; Knapp, Copland, Honig, Plecki, & Portin, 2009; Marks & Printy, 2003; May & Supovitz, 2010; Mulford & Silins, 2009; Nettles & Herrington, 2007; Opdenakker, & Van Damme, 2007; Printy, Marks, & Bowers, 2009; Silva, White, & Yoshida, 2011; Spillane, 2006; Wahlstrom & Louis, 2008; Wiley, 2001). Thus, we conclude that 50 years after publication of *Effective Principal, Effective School* (Lipham, 1961), instructional leadership is become increasingly accepted globally as a normative expectation in the principalship.

Thus, we assert that even as fads and fashions in leadership have waxed and waned, scholarly interest in instructional leadership has remained surprisingly consistent and strong. Over the ensuing decades scholars have generated a substantial body of empirical research on instructional leadership. This research has been the subject of analytical reviews conducted by scholars in North America (Hallinger & Heck, 1996a, 1996b; Leithwood et al., 2006, in press), Europe (Bell et al., 2003; Southworth, 2002; Witziers et al., 2003), Asia (Hallinger, 2011a, 2011b, 2012) and ANZ (Mulford, & Silins, 2003; Robinson et al., 2008). Indeed, data reviewed in recent years (e.g., Hallinger, 2011a, 2011b; Leithwood & Jantzi, 2005; Robinson et al., 2008) affirm that instructional leadership has become firmly entrenched in the firmament of professional practice and gained currency as a focal construct in the eyes of scholars. One can conclude that instructional leadership is possibly even more relevant in today than thirty years ago and certainly in a wider array of context around the world.

As will be detailed later in this report, the PIMRS instrument appears to provide reliable and valid data on instructional leadership when the assessments come from teachers. With respect to desirable foci for research, we suggest that scholars who use the PIMRS more squarely accept the challenge of investigating the linkages between instructional leadership and school-level variables that mediate effects on teacher effectiveness, and student learning. Numerous scholars have noted the need to shed light on the "black box" which contains the processes through which leadership contributes to the improvement capacity of schools to create a positive impact on student learning (Hallinger & Heck, 1996a; Heck & Hallinger, 2005; Leithwood et al., in press). We noted increased interest in studying this issue over the last decade, and wish to encourage it further using comprehensive conceptual models, multivariate statistics, as well as through mixed method studies.

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

As suggested above, we believe that impact should be studied in terms of student learning. However, worthy research may also target other intermediate and distal variables such as teacher collective efficacy, satisfaction and commitment, school health, organizational learning, teacher change, and student engagement. Useful models for conducting empirical research on the relationship between school leadership and these variables using comprehensive models and robust statistical methods exist in the literature (e.g., see Hallinger, Bickman & Davis, 1996; Heck & Hallinger, 2009; Heck, Larson & Marcoulides, 1990; Leitner, 1994; Leithwood & Jantzi, 2000; Marks & Printy, 2003). Doctoral students are encouraged to draw upon these models and methods, rather than simply citing findings from these studies.

Chapter 2

Conceptual Framework[ii]

The quotation from Bridges (1967) noted earlier, highlighted the importance of starting with a sound definition of what is meant by instructional leadership. Bridges had asserted that coherent discussions about the instructional leadership role were invariably hindered by the lack of a common definition and language about the construct. This chapter first introduces two of the most salient conceptual models of instructional leadership. Then, these models are placed in a broader perspective of leadership for learning. More specifically, we examine the various ways in which leadership has been conceptualized to impact learning.

Models of Instructional Leadership

Two predominant conceptual models of instructional leadership emerged during the 1980s in the USA. These were developed by Bossert and colleagues (1982) at the Far West Lab for Research and Development in San Francisco, and a complementary model developed by Hallinger and Murphy (1985a). We examine each of these in turn.

Bossert's Instructional Leadership Model

Thirty years ago, in their seminal review of the literature, Bossert and colleagues (1982) defined the construct of instructional management. They selected the term instructional management because they inferred that this role of the principal revolved around managerial functions concerned with the coordination and control of curriculum and instruction. Their instructional management framework (see Figure 2.1) became an influential model that, to this day, continues to guide researchers in this field.

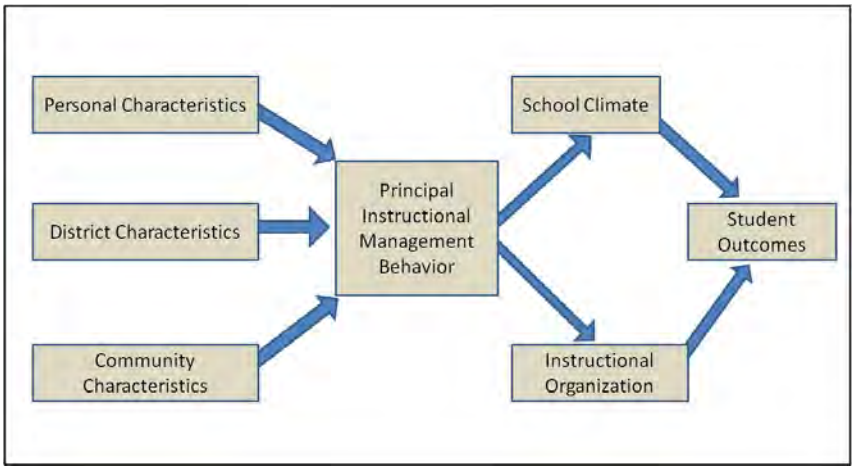


Figure 2.1. Instructional management framework (Bossert et al., 1982)

There are several distinctive features of the Bossert framework worthy of note:

- The model gives priority to a specific domain of the principal's activities, instructional management.

- Principal leadership is frame within a context, thereby recognizing that leadership itself is influenced by both personal and organizational features (e.g., Belchetz & Leithwood, 2007; Bridges, 1977; Goldring et al., 2008; Hallinger & Murphy, 1986). Leaders do not operate in a vacuum.
- The principal's effects on student outcomes are mediated by features of the school (i.e., school climate and instructional organization). This is consistent with what Bridges (1977, 1982) termed, "achieving results through people.
- The ultimate effectiveness of the principals efforts are based upon the impact achieved on students.

While Bossert and his colleagues coined the term instructional management, over time instructional leadership came to be accepted as the term commonly used by scholars and practitioners. In our view, the formal distinction between these conceptual terms lies in the sources of power and means proposed to achieve results. Instructional leadership become the preferred term due to the recognition that principals who operate from this frame of reference rely more on expertise and influence than on formal authority and power to achieve a positive and lasting impact on staff motivation and behavior and student learning (e.g., Blasé, 1987; Hallinger, 2003; Hallinger & Heck, 1996a; Leithwood et al., 1990).

Hallinger and Murphy's Instructional Leadership Model

Another early attempt to provide a clear definition of instructional leadership came from the author and Joseph Murphy in the early 1980s (Murphy, Hallinger, Weil, & Mitman, 1983; Hallinger et al., 1983; Hallinger & Murphy, 1985a). Our proposed conceptual framework incorporated three dimensions in this role: Defining the School's Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate (Hallinger et al., 1983; Hallinger & Murphy, 1985a; see Figure 2.2). These dimensions were further delineated into 10 instructional leadership functions.

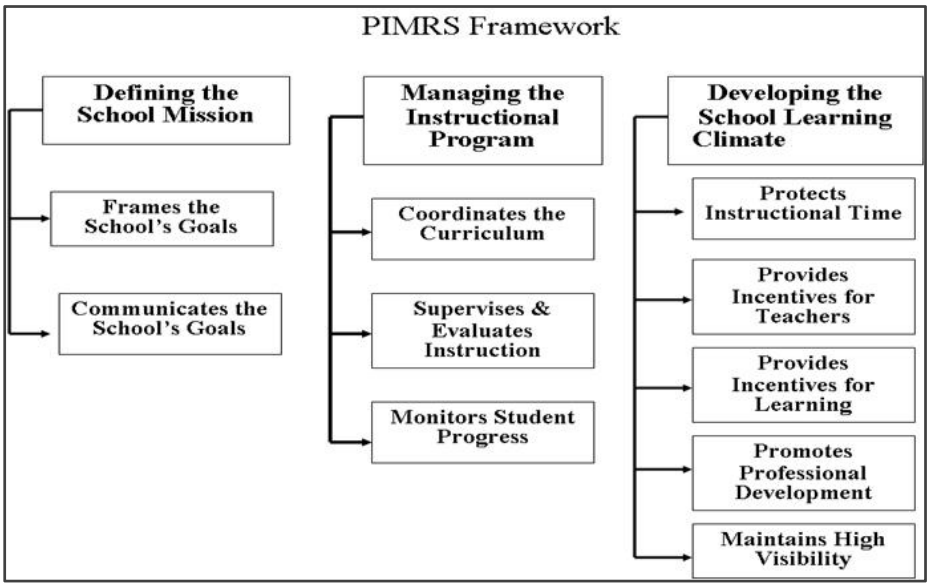


Figure 2.2. PIMRS conceptual framework

Defining a School Mission

Two functions, Framing the School's Goals and Communicating the School's Goals, comprise the dimension, Defining the School's Mission. These functions concern the principal's role in working with staff to ensure that the school has a clear mission and that the mission is focused on academic progress of its students. While this dimension does not assume that the principal defines the school's mission alone, it does propose that the principal is responsible for ensuring that such a mission exists and for communicating it widely to staff. This dimension is the starting point for creating a learner-centered school.

Within this model, the process of goal development was considered less critical than the outcome. Goals could be set by the principal or in collaboration with staff. The bottom-line, however, was the school should have clear, academic goals that staff support and incorporate into their daily practice. This picture of goal-oriented, academically-focused schools contrasted with the typical situation in which schools were portrayed as pursuing a variety of vague, ill-defined, and sometimes conflicting academic and non-academic goals.

Framing School Goals

This refers to principal's role in determining the areas in which the school will focus its resources during a given school year. Instructionally effective schools generally have a clearly defined mission or set of goals which student achievement. The emphasis is on fewer goals around which staff energy and other school resources can be mobilized. A few coordinated objectives, each with a manageable scope, appear to work best. The goals should incorporate data on past/current student performance and include staff responsibilities for achieving the goals. Staff and parent input during the development of the school's goal seem important. Performance goals should be expressed in measurable terms (Brookover, Schweitzer, Schneider, Beady, Flood, & Wisenbaker, 1982; Edmonds, 1979; Venezky & Winfield, 1979; Clark, 1980; Bossert et al., 1982).

Communicating School Goals

This function is concerned with the ways in which the principal communicates the school's most important goals to teachers, parents, students etc. Principals can ensure that the importance of the school's goals is understood by discussing and reviewing them with staff on a regular basis during the school year, especially in the context of instructional, curricular, and budgetary decisions. Both formal communication channels (e.g., goal statements, staff bulletins, articles in the principal or site council newsletter, the school handbook, assemblies) and informal ones (e.g., parent conferences, teacher conferences, curricular meetings, other discussions with staff, can be used to communicate the school's primary purpose (Brookover et al., 1978; Brookover & Lezotte, 1979 Edmonds, 1979; Brookover et al., 1982; Hallinger et a1., In press).
The instructional leader's role in defining a school mission was captured in a study of effective California elementary schools conducted by Hallinger and Murphy (1986). In the course of their study, they observed teachers in their classrooms for several days. One teacher had an affective education activity center entitled "I am. . ." in the back of the room. However, they never saw students working at it. When queried about this, the teacher observed:

> Yes, the affective activity center is something I really like to use with my students. However, this particular class has not made the usual progress in basic subjects, so I've had less time for affective activities. Our focus in the school is on

> ensuring that every one of our students has mastered basic subjects. We really try to make time for optional subjects as well. However, our principal expects us to spend as much time on reading, writing, spelling, and math as is necessary to achieve this objective (emphasis added). So I adjust the time accordingly. (Hallinger & Murphy, 1986)

Later during one of his interviews, the principal repeated this expectation almost word for world. It was obviously something that had been discussed with and among the staff many times.

This comment captures several characteristics of the instructional leader's role in defining a clear mission. First, at this school the mission was absolutely clear. It was written down and visible around the school. Second, it was focused on academic development appropriate to the needs of this particular school population. Third, the mission set a priority for the work of teachers. Fourth, it was known and accepted as legitimate by teachers throughout the school. Fifth, the mission was articulated, actively supported, and modeled by the principal.

Managing the Instructional Program

The second dimension Managing the Instructional Program focuses on the coordination and control of instruction and curriculum. This dimension incorporates three leadership (or what might be termed management) functions: *Supervising and Evaluating Instruction, Coordinating the Curriculum, Monitoring Student Progress*. This dimension focuses on the role of the principal in "managing the technical core" of the school. In larger schools, it is clear that the principal is not the only person involved in monitoring and developing the school's instructional program. Yet this framework assumes that coordination and control of the academic program of the school is a key leadership responsibility of the principal.

This dimension requires the principal and other leadersto be deeply engaged in stimulating, supervising and monitoring teaching and learning in the school. Obviously, these functions demand that the principal to have expertise in teaching and learning, as well as a commitment to the school's improvement. It is this dimension that requires the principal to become "hip-deep" in the school's instructional program (Bossert et al., 1982; Cuban, 1984; Dwyer, 1986; Edmonds, 1979; Marshall, 1996).

By way of example, I would again recall the principal in the example cited above. In discussions of how they monitored student progress, several different teachers at this school observed that the principal "knew the reading level and progress of all 650+ students in this primary school" (Hallinger & Murphy, 1985b, 1986). This particular behavior is not a requirement for instructional leadership. However, it reflects the degree of this principal's involvement in monitoring student progress and in managing the school's instructional program.

Supervising and Evaluating Instruction

A central task of the principal is to ensure that the goals of the school are being translated into practice at the classroom level. This involves coordinating the classroom objectives of teachers with those of the school and evaluating classroom instruction. In addition, it includes providing instructional support to teachers and monitoring classroom instruction through numerous informal classroom visits (Levine, 1982; Lipham, 1981; New York State, 1974).

Coordinating Curriculum

A characteristic which stands out in instructionally effective schools is the high degree of curricular coordination. School curricular objectives are closely aligned with both the content taught in classes and the achievement tests used by the school. In addition, there appears to be a fairly high degree of continuity in the curricular series used across grade levels. This aspect of curricular coordination is often supported by greater interaction among teachers within and across grade levels on instructional and/or curricular issues (Brookover et al., 1982; Clark, 1980; Cohen & Miller, 1981; Cooley & Leinhardt, 1980; Levine, 1982; New York State, 1974; Venezky & Winfield, l979; Wellisch, MacQueen, Carriere, & Duck, 1978).

Monitoring Student Progress

Instructionally effective schools place a strong emphasis on both standardized and criterion referenced testing. The tests are used to diagnose programmatic and student weaknesses, to evaluate the results of changes in the school's instructional program, and to help in making classroom assignment. The principal plays a key role in this area in several ways. He/she can provide teachers with test results in a timely and useful fashion, discuss test results with the staff as a whole, with grade level staff and individual teachers, and provide interpretive analyses for teachers detailing the relevant test data in a concise form (Brookover et al., 1982; Edmonds, l979; Hallinger et al., In press; Stallings 1980; Purkey & Smith, 1982; Stallings & Mohlman, 1981; Venezky & Winfield, 1979).

Developing a Positive School Learning Climate

The third dimension, Developing a Positive School Learning Climate includes several functions: Protecting Instructional Time, Promoting Professional Development, Maintaining High Visibility, Providing Incentives for Teachers, and Providing Incentives for Learning. This dimension is broader in scope and intent than the second dimension and overlaps with dimensions incorporated into transformational leadership frameworks (Hallinger, 2003; Leithwood et al., 2006). It conforms to the notion that successful schools create an "academic press" through the development of high standards and expectations and a culture that fosters and rewards continuous learning and improvement.

Instructionally effective schools develop a culture of continuous improvement in which rewards are aligned with purposes and practices (Barth, 1990; Glasman, 1984; Hallinger & Murphy, 1986; Heck et al., 1990; Leithwood & Montgomery, 1982; Mortimore, 1993; Purkey & Smith, 1983). Finally, the principal must model values and practices that create a climate and support the continuous improvement of teaching and learning (Dwyer, 1986; Hallinger & Murphy, 1985b).

Protecting Instructional Time

The work of Jane Stallings and others on allocated learning time has called attention to the importance of providing teachers with blocks of uninterrupted work time. Improved classroom management and instructional skills are not used to the greatest effect if teachers are frequently interrupted by announcements, tardy students, and requests from the office. The principal has control over this area through the development and enforcement of school - wide policies related to the interruption of classroom learning time (Bossert et al., 1982; Stallings, 1980; Stallings & Mohlman, l98l; Wynne, 1980).

Maintaining High Visibility

The contexts in which the principal is seen provide one indicator to teachers and students of his/her priorities. Although a significant portion of the principal's time may be out of his/her control, the principal can set priorities on how the remaining time is to be spent. Visibility on the campus and in classrooms increases the interaction between the principal and students as well as with teachers. This can have positive effects on student behavior and classroom instruction (Brookover et al., 1982; Casey, 1980; Clark, 1980; Wynne, 1980).

Providing Incentives for Teachers

Few monetary rewards are available principals to use with teachers. The single salary schedule and tenure system severely limit the alternatives open to principals with respect to motivating teachers. However, one study in which the relative effects of money, praise and public recognition were measured found that money was only slightly more effective than praise as an incentive. Clearly money is less cost effective. This suggests that the principal should make the best use of both formal and informal ways of providing teachers with praise when it is deserved (Latham & Wexley, 1981).

Promoting Professional Development

The principal has several ways of supporting teachers in the effort to improve instruction. He/she can arrange for, provide, or inform teachers of relevant opportunities for staff development. The principal also can encourage certain types of staff development which/ are closely linked to the school's goals (Brookover et al., 1982; Clark, 1980; Little, 1982; Rutter, Maugham, Mortimore, Ouston, & Smith, 1979).

Providing Incentives for Learning

The last function of the principal covered under the heading of instructional management is the area of promoting incentives for learning. It is possible to create a school learning climate in which academic achievement is highly valued by students by providing frequent opportunities for students to be rewarded and recognized for their academic achievement and improvement. The rewards need not be fancy or expensive; the recognition before teachers and peers is the key. Students should have opportunities to be recognized for their achievement both within the classroom and before the school as a whole (Brookover et al., 1978; Hallinger et al., In 1983; Rutter et al., l979).

The above dimensions of instructional leadership describe the scope of responsibilities of the principal and the school's leadership team with respect to leading learning. However, it is also useful to place these responsibilities into the broader context of how leadership achieves its effects in schools.

Modeling the Relationship Between Leadership and Learning

The phrase "school improvement leadership" implies the existence of a cause-effect relationship between the strategies of leaders, school improvement activities, teacher classroom practices, and growth in student outcomes. As noted above, it is only since the 1960s that scholars began to conceptualize and study school leadership as directed *explicitly* toward improvement in the quality of teaching (Gross & Herriott, 1965). Although progress has been made in defining the nature of

these relationships, scholars operating in the UK (Bell et al., 2003; Southworth, 2002, 2003), USA (Bossert et al., 1982; Hallinger & Heck, 1996, 1998), Canada (Leithwood et al., 2004, in press; York-Barr & Duke, 2004), Netherlands (Krüger et al., 2007; Sleegers et al., 2002; Witziers et al., 2003), and AnZed (Mulford & Silins, 2003; Robinson et al., 2008) continue to debate the meaning of empirical findings on school leadership effects.[iii] Moreover, the predominant assumption that leadership impacts school improvement understates the extent to which leaders are influenced by the organizational environment (Hallinger & Heck, 1996; Krüger et al., 2007; Leithwood et al., 2004; Southworth, 2002). Thus, we conclude that research on school leadership effects must take into account features of the organizational context and continue to approach issues of causal inference with caution.

In 1988, Pitner proposed several conceptual models that sought to explain the means by which leadership could impact student learning. The models included direct effects, indirect effects and reciprocal effects models of leadership for learning (see Figure 2.3). A decade later, Hallinger and Heck elaborated on these models in a review of empirical research on principal leadership and student learning (Hallinger & Heck, 1996a, 1996b, 1998).

- *Direct effects models* proposed that leadership effects could results directly from the actions of principals, and moreover, that these effects could be identified by analyzing the relationship between comparing measures of leadership and measures of student learning in samples of principals and students.
- *Indirect effects models* proposed that leaders obtained effects on students by impacting the structure, culture and people in the school organization (e.g., Bridges, 1977). The Bossert model show in Figure 2.1 represents one influential indirect (also referred to as mediated effects) model of leadership and learning. In the Bossert (1982) model, principal leadership influences learning through the principal's efforts to shape the school learning climate and instructional organization.
- *Reciprocal effects* models propose that leadership is a process of mutual interaction and influence both between leaders and followers and between the leader and his/her organizational context (e.g., school culture, community). In one sense reciprocal effects models tend to incorporate indirect interactions. However, they differ from standard indirect effects models by seeking measure the dynamic relationship of the leader within his/her school environment (Hallinger & Heck, 2011; Heck & Hallinger, 2001).
- As suggested in Figure 2.3, the comprehensiveness of any of these models can be enhanced through the inclusion of antecedent (e.g., personal characteristics of the principal) or context (e.g., school size, school level, student SES) variables. This, is shown for example in the distinctions between Model a and Model A-1 in Figure 2.3.

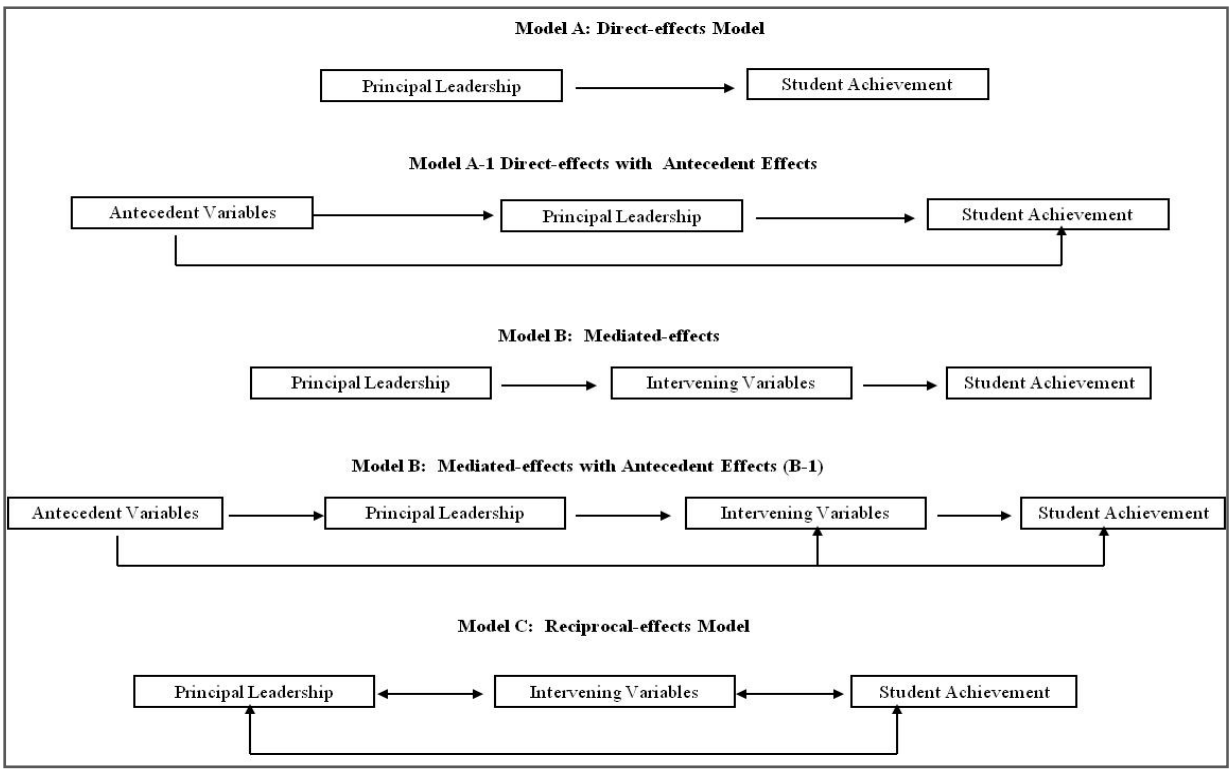


Figure 2.3 Conceptual models of leadership and learning (From Hallinger & Heck, 1996)

More recently researchers have tested these models as a means of furthering our understanding of how collaborative leadership contributes to school improvement and student learning (Hallinger & Heck, 2010, 2011; Leithwood & Patten, 2010; Mulford & Silins, 2009; Robinson et al., 2008; Witziers et al., 2003). Approaches to studying these models are discussed further in Chapter Seven.

Chapter 3

Developing and the PIMRS Instrument

Earlier it was noted that prior to the early 1980s, there were no validated tools available for measuring instructional leadership either for the purposes of research or practice.

Instrument Development

Three parallel forms of the PIMRS instrument have been developed and tested: a self-assessment form to be completed by the principal, a teacher form and a supervisor form. The items which comprise each form are identical; only the stems change to reflect the differing perspectives of the role groups. Early studies found significant differences in perceptions across role groups (Hallinger & Murphy, 1985; Krug, 1986; O'Day, 1984). Validation studies in the United States indicate that the PIMRS form that solicits teachers' perceptions provides the most valid data of the three forms. The Principal Instructional Management Rating Scale (PIMRS) provides a principal performance on 10 instructional leadership job functions associated with principal leadership in effective schools. The subscales are comprised of 50 items, which refer to specific principal behaviors or practices. The PIMRS has been used successfully at the elementary and secondary levels, and with both principals and assistant principals.

The original form of the PIMRS (Hallinger, 1982) contained 11 subscales and 72 "behaviorally anchored" items (See Hallinger (1982) and Latham and Wexley (1981) for discussions of behaviorally anchored rating scales and their development). Subsequent revision of the instrument reduced the instrument to 10 subscales and 50 items (Hallinger, 1983). As noted earlier, in 2012 the author developed a Teacher Short Form with 23 items (see Chapter Seven).

For each item, the rater assesses the frequency with which the principal enacts a behavior or practice associated with that particular instructional leadership function. Each item is rated on a Likert-type scale ranging from (1) almost never to (5) almost always (see Figure 3.1). The instrument is scored by calculating the mean for the items that comprise each subscale. This results in a profile that yields data on perceptions of principal performance on each of the 10 instructional leadership functions.

Sample *PIMRS* Rating Subscale: Teacher Form

To what extent does your principal. . . ?

I. FRAME THE SCHOOL GOALS

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

Figure 3.1 Sample Items from the PIMRS

Behaviorally Anchored Rating Scales

The methodology used to develop an instrument for measuring principals instructional management behavior generally followed steps prescribed by Latham and Wexley (1981) for constructing behaviorally anchored rating scales (BARS). Behaviorally anchored rating scales rely upon descriptions of critical job related behaviors for the development of scale items. The items are "behaviorally anchored" in the sense that they are statements of critical job related behaviors on which raters can base their appraisal of an individual's performance within a given dimension of a job.

The strength of the BARS approach lies in its specificity; the scales make explicit to both the appraiser and the employee exactly what is expected and what must be observed with respect to the employee's on the job behavior. The scales can also serve other functions within the organization: as the basis for a job description; as part of a performance feedback system for staff evaluation; as a blueprint for the development of staff training in the areas measured by the instrument; and, as an aid in manpower planning (Bernardin, 1977; Blood, 1974; Harari & Zedeck 1973; Latham, Fay, & Saari, 1979; Latham & Wexley, 1977, 1981; Smith & Kendall, 1963).

Developing the PIMRS

The first step in the development of the rating scales was to perform a careful job analysis of the principal's role as instructiona1 manager. The job analysis in this case drew heavily from research conducted on instructionally effective schools, schools in which students succeed beyond what would be expected given their socio-economic background (Bossert et al., 1982; Purkey & Smith, 1983). Eleven job functions which reflect the areas of responsibility of the principal in his/her role as instructional manager were abstracted from research effective schools and leadership (later reduced to 10 functions).

Item Construction

The methodology used to develop the instrument departed somewhat from the BARS approach in constructing the scale items. There are at least two ways to generate the specific behaviors which comprise the instructional management role of the elementary school principal. The BARS approach is inductive. It has a number of knowledgeable persons from the targeted role group(s) identify critical incidents in the relevant dimensions of the job being analyzed (Flanagan, 1954; Latham & Wexley, 1981; Smith & Kendall, 1963). A second approach is to deduce the critical instructional management behaviors from the descriptions of principal functions in the literature on instructionally effective schools. This study utilized a combination of these two approaches in order to construct scale items for the rating instrument.

The rationale behind this departure from the BARS methodology is related to a point made quite frequently in the literature on instructionally effective schools. These schools appear to differ substantially from schools in general in terms of the staff norms and role behavior which predominate (Brookover et al., 1978; Edmonds, 1979). A strict reliance on an inductive method such as the critical incident technique for generating the job related behaviors might have resulted in too narrow a range of principal behavior. The participating principals, if they fit the pattern of managerial behavior described earlier (i.e., only peripherally involved in instructional matters),

might have viewed sonic of job functions identified in the literature on instructionally effective schools as relatively unimportant.

The following steps were followed to generate the scale items:

1. First, as noted above, the literature on instructionally effective schools was reviewed in order to develop the job functions comprising instructional management.
2. Next, the "expert opinion" of a superintendent, his staff assistant and several principals was solicited in order to generate a list of critical job related behaviors within each of the job functions. This step was somewhat similar to the procedure described by Latham and Wexley for developing BARS, in the sense that it was done inductively.
3. The list developed in step two was then supplemented with behaviors deduced by the author within each of the job functions; in some cases other research findings were drawn upon if they shed light upon the critical behaviors constituting a particular job function.
4. This list of critical job related behaviors contained sixty behavioral statements concerning the principal's role as instructional manager. The behavioral statements were then rewritten so they described discrete behaviors for use as questionnaire items. This step resulted in a total of eighty-nine critical job related behaviors within the three general dimensions and eleven functional categories comprising instructional management. The additional behaviors resulted from the breakup of statements which contained more than discreet behavior.
5. Finally, each of the behavioral statements was adjusted grammatically so it would fit the same stem and response category. A "1" to "5" response scale accompanied each item with 1 representing "almost never"; 2, "seldom"; 3, "sometimes"; 4, "frequently"; and, 5, "almost always."

Survey Administration

The questionnaire that resulted from this set of procedures was used to collect information on the instructional management behavior of ten elementary school principals from a single school district. The raters were drawn from three role groups: 1) teachers at each of the schools (total of one hundred and four); 2) the ten elementary school principals; and, 3) supervisors from the district office (total of three). The district office supervisors included the superintendent, deputy superintendent, and the director of instruction. The same rating instrument was administered to each group, though the questionnaires were completed at different times and under different conditions. The computations contained in this article are derived solely from the teachers' questionnaire responses.

The PIMRS may be completed by teachers, the principal, and/or by a district office supervisor. The specific purpose(s) for which the results are to be used determine which role groups should be surveyed. Two discrete purposes exist for using the PIMRS: as part of a principal evaluation program or as part of a professional development program. Although there is overlap, it is useful to address separately the differing data needs associated with each of these purposes.

The original validation study found that the PIMRS met high standards of reliability (Hallinger, 1983). The reliability and validity of the instrument are discussed in Chapters 4 and 5. The PIMRS was designed to meet standards for use in research, professional development and principal

evaluation (Hallinger & Murphy, 1985a; Hallinger et al., 2013). We discuss the use of the instrument for each of these purposes in the next section of this chapter.

Scoring the PIMRS

The PIMRS has been designed so that it can be easily scored by the principal at his/her school site. One point of caution; do not mix scores from different role groups. That is, keep the ratings obtained from teachers separate from those of central office supervisors or the principal's own self-assessment. Although the process of scoring the instrument is straightforward, some thought is required concerning the types of scores that are desired by the user(s). Before discussing specific scoring methods it is necessary to understand a little about the construction of the instrument.

The PIMRS is divided into 10 subscales, each of which measures a different instructional leadership function (e.g., Framing School Goals, Providing incentives for Learning). Useful information can be obtained from ratings on the individual items within each subscale. Most commonly, however, the instrument is used to provide feedback on the subscales as components of instructional leadership. We do not generally suggest that users score the instrument as a single instructional leadership score. The PIMRS was not designed to provide a single score, and such a use would provide a distorted profile of a principal's performance in this role. Regardless of whether the PIMRS is used as part of an evaluation system or a professional development program, principals will only benefit from feedback on their performance that is usable. A single whole scale score does not provide usable information. Use of a single, whole scale score represents a misuse of the PIMRS.

Each instructional leadership subscale in the PIMRS consists of 5 items. Each item is scored on a "1" to "5" scale ("Almost Never" to Almost Always"), denoting the frequency with which the specific behavior is practiced. Several types of scores have proved' worthwhile in working with administrators. Think through the types of information you desire, and then choose the type(s) of scores that best get at that information. We also suggest portraying the results on bar graphs so that relative strengths and comparisons can be seen more easily (see attached).

1. **Item Averages** - These are obtained by averaging the scores from/the respondents on each item. Thus, if 25 teachers completed the assessment, their responses on item one would be averaged to obtain a mean score for that item.
2. **Item Distributions** - This score is only used where there are multiple respondents, such as the school faculty or several central office supervisors. Sometimes the mean score masks the perceptions of the various respondents. A mean score on an item of 3.5 may be obtained with a large distribution of teachers rating the principal at 2.5 and others at 4.5, or with most of the teachers rating the principal between 3.2 and 3.8. Thus, the interpretation of the same mean score can vary according to the distribution of responses on an item.
3. **Scoring the PIMRS** - In order to portray the distribution of responses on an item, simply place an X on a scoring sheet next to each of the possible responses from 1 to 5. Count the responses and report the number of each along with the mean score for the item.
4. **Subscale Averages and Distributions** - The subscale average is the basic score used with the PIMRS. This score portrays the administrator's performance within a given instructional leadership function. It is obtained by averaging the item scores within each instructional leadership subscale. Where there is more than one respondent, the score is obtained by averaging the averages". That is, in step one find the mean score on the subscale Framing

School Goals each of the teachers. Then average their mean scores on this subscale to obtain a grand mean" score. Again, it may be desirable, to portray the distribution of averages to get a sense of the spread of teacher perceptions.

5. **District Scores** - The above scoring methods focus attention on the scores of the individual administrator. At times, it may be useful to compare the scores of individual administrators to those of the administrators within the district. The average score of the principals as a group on each of the subscales constitute the district average. Again, some representation of the variability of scores across administrators is useful.
6. **Group Comparison Scores** - One of the uses of the PIMRS that can be eye-opening to administrators is the comparison of the perceptions of different role groups. This is also, in our opinion the place where administrators experience the greatest growth when done with care. Role group comparison scores simply portray the perceptions of the different role groups (i.e., teachers, supervisors and self-assessment) using any of the measurement methods listed under numbers 1., 2., and 3., above.
7. **District Comparison Scores** - The individual administrator's item or subscale averages can be compared to a district-wide average in larger districts (i.e., districts with more than 5 schools). There are no norms for the PIMRS scores. Thus, the scores obtained on the instrument must be compared the district's own sense of how it would like the administrators to perform the Instructional leadership role. In fact, we have found that the district support of the instructional leadership role is one of the best predictors of an individual principal's PIMRS scores.

Interpreting PIMRS Scores

It is important to note that the PIMRS does not measure an administrator's effectiveness. Rather, it assesses the degree to which a principal is providing instructional leadership in his/her school. Although higher item and subscale scores may suggest greater leadership activity by the administrator, the most effective principals do not necessarily score "5" on all subscales of the PIMRS. Contextual factors including school level and size, faculty age and experience, student background and levels of achievement all influence the type of instructional leadership that is appropriate in a given school. Thus, users of the PIMRS are encouraged to interpret the results in relation to: 1) the needs of their school; 2) the score of other administrators in the district; 3) changes in scores from the prior year(s). Thus, the results are designed primarily as a method of formative evaluation.

As noted above, three parallel forms of the instrument have been developed and tested: a self-assessment form to be completed by the principal, teachers or supervisors (Hallinger, 1982). The items that comprise each form are identical; only the stems change to reflect the differing perspectives of the role groups. Researchers have consistently reported significant differences between teacher and principal perceptions of the principal's instructional leadership, with principal self-report scores substantially higher than those obtained from teachers (e.g., Brown, 1991; Corkill, 1994; Dennis, 2009; Haack, 1991; Haasl, 1989; Hallinger, 1983; Henderson, 2007; Krug, 1986; Mallory, 2002; Marshall, 2005; Meek, 1999; Meyer, 1990; O'Day, 1984; O'Donnell, 2002; Reid, 1987; Shatzer, 2009; Smith, 2007; Stevens, 1996; Vinson, 1997). Notably these 'role set' (Merton, 1957) differences in PIMRS ratings obtained from teachers and their principals extend to contexts other than the USA. Empirical comparisons have yielded a similar pattern of

results in Thailand (Hallinger & Lee, 2012; Poovatanikul, 1993; Ratchaneeladdajit, 1997; Taraseina, 1993), Guam (San Nicolas, 2003), the Philippines (Saavedra, 1987; Salvador, 1999; Yogere, 1996), the Maldives (Wafir, 2011), Hong Kong (Chan, 1993), and Taiwan (Chi, 1999; Tang, 1993; Yang, 1996). It should be noted, in addition, that despite differences in the magnitude of ratings obtained by the two role groups, there is often a similar pattern in their ratings on the various sub-scales that comprise the PIMRS.

Chapter 4

Reliability of the PIMRS

This chapter introduces information concerning the reliability of the PIMRS as a tool for assessing principal instructional leadership. A key step in instrument development lies in establishing the reliability of the PIMRS was an important step in advancing our understanding of instructional leadership. The chapter presents information the reliability of the PIMRS based on the initial validation study (Hallinger, 1983; Hallinger & Murphy, 1985a) as well as a more comprehensive meta-analytic study 30 years after the instrument's development (Hallinger et al., 2013).

Background on the Concept of Reliability

Lang and Heiss (1998) defined reliability as the consistency with which an instrument yields the same or similar responses across settings and time. Several different approaches may be employed for assessing the reliability of a test instrument: test-retest, parallel forms, internal consistency (Kerlinger, 1966). Studies employing the PIMRS have relied exclusively on measures of internal consistency of the instrument. Internal consistency refers to the degree to which items that have been grouped together conceptually as subscales correlate with each other (Kerlinger, 1966).

The true score for each principal is his/her mean score across all possible observations, that is, observations by all teachers in the school on all occasions, using all possible measuring devices. Given this conception, a re1iabil1ity coefficient which estimates inter-rater reliability, correcting for they numbers of raters from school to school, is appropriate. If re1iabi1ity is viewed as the portion of observed score variance which is true variance, then the larger the coefficient, the less is the error and the more reliable is the instrument. When a high corre1ation exists among observations about the same object one can conclude that the instrument yields consistent or reliable data.

It is unlikely that perfect inter-rater reliability ever exists because no two raters ever see the object about which they are responding at the same time or doing the same things. One standard appropriate in situations where an instrument. is used for employee evaluation, is that the correlation among different raters should be at least .60. If agreement among rater less than .60 and the raters have had similar opportunities to view the person being examined, then it is like1y that the instrument is measuring different attitudes and biases of the rather than the performance of the person being observed (Latham & Wexley, 1981)

Approaches to Assessing Reliability

The PIMRS studies have varied with respect to the form of the scale that was used (i.e., teacher or principal), the level of the scale on which reliability was calculated (i.e., whole scale, dimension, function), as well as the school level (i.e., primary, middle, secondary) and cultural context (i.e., North America or East Asia) in which the study was conducted. Moreover, it is important to note that researchers have employed two different statistical tests to assess the scale's internal consistency. Some have used Cronbach's (1951) alpha, while others have employed Ebel's (1951)test.

Gathering data with the PIMRS directly from principals represents a type of self-assessment. The resulting score reflects a latent trait of the individual subject (Kerlinger, 1966). This is a typical case faced in measurement, and one in which researchers often employ Cronbach's alpha test of

the internal consistency among items in the scale (Cronbach, 1951). These studies have supported the original validation study in its conclusion that the scale provides reliable data on instructional management (e.g., Dunn, 2010; Fulton, 2009; Hallinger, Taraseina, & Miller, 1994; Harris, 2002; Howe, 1995; Jones, 1987; Leitner, 1994; Mercer, 2004; Moore, 2003; O'Day, 1984).

Many researchers have also employed Cronbach's test to examine the reliability of the Teacher Form of the PIMRS. For example, Hallinger's (1983) original validation study of the PIMRS applied Cronbach's alpha to data obtained with the PIMRS Teacher Form. It reported that the reliability of nine of the ten subscales exceeded .80. Numerous additional studies have employed the same test of reliability with data from teacher respondents (see Tables 1 and 5).

However, some researchers have suggested that this application of coefficient alpha violates a fundamental assumption of Cronbach's test (e.g., Jones, 1987; Leitner, 1994). When analyzing a PIMRS data set obtained from teacher respondents, Cronbach's test treats each teacher's response independently, as if each teacher was rating a different principal. In reality, however, teachers are 'nested' within schools, with each school's teachers rating their own principal. In this case, reliability estimates of internal consistency should be based on the combined ratings of teachers grouped together by their schools.

With this limitation in mind, several researchers (e.g., Howe, 1995; Jones, 1987; Leitner, 1990; Taraseina, 1993) employed Ebel's (1951) test of reliability. This test aggregates teacher ratings from a set of schools in which the respondents are grouped within their schools (e.g., teachers). Thus, the rating is treated as a feature of the school (i.e., the principal). Ebel's formula is: $r/k = (M/x – M) /M/x$, where r/k is the reliability of average ratings, M/x is the between groups variance, and M is the within groups variances. When employing this formula with the PIMRS, the researchers generally applied the test to each of the 10 leadership functions. These scholars asserted that that Ebel's formula provides a more accurate assessment of scale reliability.

For the purposes of the present background discussion of the PIMRS' reliability, it is sufficient to note that these different approaches have been employed. Thus, our data set of the reliability studies derived from research reports is diverse. The fact that the scale's reliability has been assessed from multiple perspectives and with a variety of methods represents a strength from the point of view of instrument development. W

Original Reliability Study (1983)

In Hallinger's (1983; Hallinger & Murphy, 1985)original validation study, the internal consistency of the instrument was chosen as the appropriate form of reliability. Internal consistency refers to the degree to which items that have been grouped together conceptually as subscales correlate with each other. Latham and Wexley (l981, p.66) indicate that a minimum standard reliability for behaviorally anchored rating scales should be set at .80 when assessing the internal consistency of the instrument.

Ten of the eleven functional categories or subscales met the standard of .80. The size of the Alpha coefficients for the subscales ranged from a low of .78 for the "Incentives to Improve Teaching, " to a high of .90 on three different subscales, "Supervision and Evaluation of Instruction, " Curriculum Coordination, " and Monitoring Student Progress." The reliability coefficients for the subscales are contained in Table 1. The reliability of the instrument as a whole was not measured since the individual subscales were conceptualized to represent related but discrete job functions.

Table 4.1. Reliability Estimates of the PIMRS (Hallinger, 1983)

Subscale	Reliability*	Sample Size
FRAME GOALS	.89	(77)
COMMUNICATE GOALS	.89	(70)
SUPERVISION/EVALUATION	.90	(61)
CURRICULAR COORDINATION	.90	(53)
MONITORS STUDENT PROGRESS	.90	(52)
PROTECTS INSTRUCTIONAL TIME	.84	(70)
VISIBILITY	.81	(69)
INCENTIVES FOR TEACHERS	.78	(70)
PROFESSIONAL DEVELOPMENT	.86	(58)
ACADEMIC STANDARDS	.83	(76)
INCENTIVES FOR LEARNING	.87	(61)

* Reliability estimates are Cronbach Alpha coefficients

2012 Reliability Meta-analytic Study

Hallinger's (2011a) review of research on the PIMRS found that the PIMRS continued to be an instrument of choice among scholars studying principal leadership through the first decade of the 21st century. The review study also found that the PIMRS had maintained a consistent record of yielding reliable and valid data. However, the review contained few details concerning the results of the cited studies. Given increasing global interest in instructional leadership and the continued widespread use of the PIMRS, Hallinger and colleagues (2012, 2013) undertook an updated assessment of its measurement properties.

The study consisted of a meta-analysis of reliability results derived from 43 independently conducted studies that employed the PIMRS as a research tool over the past three decades. The study's goal was to provide a comprehensive assessment of the reliability of the PIMRS. The study addressed several research questions:

1. Does the PIMRS provide reliable data for the purposes of assessing principal instructional leadership in research and practice?
2. How do reliability estimates differ based on the role group of the respondents (i.e., teachers or principals)?
3. Does the PIMRS yield reliable data when used in rating principals working at different school levels and cultural contexts?

The relevance of this information was inadvertently confirmed by a recent publication that critiqued the reliability of several leadership scales (Condon & Matthews, 2010). Yet, the information included on the PIMRS was incomplete and out-of-date. This highlighted the need for a comprehensive, up-to-date description of the reliability of the PIMRS. This would aid scholars and practitioners in choosing among instruments for assessing principal instructional leadership and in making methodological choices when using the PIMRS.

PIMRS Reliability Principal Form

We were fortunate to gain access to a large body of secondary data contained in 43 previously conducted PIMRS studies. The data consisted of two types. The first was statistical information extracted from published research reports and doctoral dissertations. The second consisted of raw data sets obtained directly from researchers who had used the scale.

There were 19 studies in which principals represented the data source (see Table 4.2). The studies had been conducted between 1991 and 2012. As shown in Table 4.2, this group consisted of 13 raw data sets and six data sets comprised of extracted information. We eliminated three studies that had surveyed fewer than 15 from our analyses (i.e., Carr, 2010; Gjelaj, 2010; Shafeeu, 2011). The remaining combined data set was comprised of 16 studies, consisting of four from East Asia and 12 from the USA (see Table 4.2). Eight studies had focused on secondary school principals, 3 on primary school principals, 1 middle schools, and four had collected data from principals across school levels. The sample size in these studies ranged 15 to 1,195 principals, with a mean of 157 principals per study, and a total sample of 2,508 principals.

It was standard procedure for the researchers to employ Cronbach's alpha in testing the reliability of the principal response data. After extracting alpha reliability estimates from the research reports, we proceeded to calculate Cronbach's alpha coefficients for the raw data sets that we had obtained. The latter data sets were preferred since access to item-level data made it possible to generate alpha reliability coefficients for the full scale, three dimensions and 10 leadership functions. We then combined the full set of alpha reliability coefficients into a single excel table in preparation for meta-analysis.

Table 4.2 Principal Data Sources

#	Author	Year	Data Type	Nation	N (P)	School Level
1	Anderson	2006	Extract	USA	190	Sec
2	Babcock	1991	Extract	USA	213	Pri
3	Carr	2011	Raw	USA	6	Pri
4	Carson	2011	Raw	USA	77	Sec
5	Dunn	2010	Extract	USA	128	Pri/Sec
6	Gjelaj	2010	Raw	USA	10	Pri
7	Goldring	2012	Raw	USA	58	Pri/Sec
8	Greb	2011	Raw	USA	31	Pri
9	Hallinger	2012	Raw	Thai	1195	Pri/Sec
10	Long	2008	Raw	USA	67	Sec
11	Lyons	2010	Raw	USA	15	Sec
12	Minus	2010	Extract	USA	62	Middle
13	Munro	2009	Raw	USA	35	Pri
14	Nogay	1995	Extract	USA	61	Sec
15	Peariso	2011	Extract	USA	36	Sec
16	Shafeeu	2011	Raw	Mald	10	Pri/Sec
17	Todd	2006	Raw	USA	122	Sec

18	Wang	2011	Raw	Chin	23	Sec
19	Wong	2010	Raw	Malay	195	Pri/Sec

The results of our efforts to summarize the reliability of the Principal Form of the PIMRS across studies are presented in Table 4.3. Cronbach's alpha was the statistic for estimating reliability employed in all of these studies. The full sample consisted of 2,508 principals. As indicated in Table 4.3, the whole-scale alpha reliability estimate was 0.96. Reliability estimates for the three dimensions were 0.88 for *Defining a School Mission*, 0.91 for *Managing the Instructional Program*, and 0.93 for *Developing a Positive School Learning Climate*. These all reflect a high standard of reliability.

Table 4.3. Meta-analysis of Reliability Derived from Principal Respondent Studies

#	Author	Year	Data Source	Nation	School Level	N (P)	Whole Scale	Create Mission	Instr'l Program	Develop Climate	Frame Goals	Comm Goals	Sup Inst'n	Coord Curric	Mon Prog	Highly Visible	Incent Tchers	Prof Deve	Incent Lrners	Inst Time
1	Anderson	2006	Extract	USA	Sec	190	-	0.90	0.92	0.87	-	-	-	-	-	-	-	-	-	-
2	Babcock	1991	Extract	USA	Pri	213	-	-	-	-	0.82	0.75	0.69	0.72	0.64	0.65	0.66	0.76	0.68	0.76
3	Carson	2012	Extract	USA	Sec	77	0.91	0.79	0.87	0.80	0.68	0.71	0.73	0.81	0.75	0.51	0.47	0.79	0.68	0.66
4	Dunn	2010	Extract	USA	Pri/Sec	128	0.94	-	-	-	0.86	0.86	0.57	0.88	0.85	0.79	0.86	0.86	0.85	0.60
5	Goldring	2012	Raw	USA	Pri/Sec	58	0.96	0.84	0.93	0.92	0.79	0.74	0.83	0.83	0.85	0.70	0.77	0.81	0.75	0.72
6	Greb	2011	Raw	USA	Pri	31	0.85	0.80	0.83	0.59	0.80	0.76	0.88	0.75	0.65	0.68	0.65	0.61	0.67	0.61
7	Hallinger	2012	Raw	Thai	Pri/Sec	1195	0.96	0.88	0.91	0.94	-	-	-	-	-	-	-	-	-	-
8	Long	2008	Raw	USA	Sec	67	0.99	0.95	0.97	0.97	0.92	0.88	0.93	0.92	0.92	0.90	0.92	0.92	0.91	0.90
9	Lyons	2010	Raw	USA	Sec	15	0.91	0.84	0.88	0.77	0.64	0.79	0.24	0.86	0.82	0.54	0.70	0.26	0.68	0.70
10	Minus	2010	Extract	USA	Mid	62	-	-	-	-	0.81	0.77	0.81	0.71	0.80	0.75	0.48	0.73	0.80	0.75
11	Munro	2009	Raw	USA	Pri	35	0.93	0.80	0.88	0.85	0.63	0.74	0.77	0.74	0.81	0.62	0.66	0.78	0.74	0.78
12	Nogay	1995	Extract	USA	Sec	61	0.93	-	-	-	0.86	0.83	0.60	0.83	0.72	0.67	0.74	0.72	0.71	0.72
13	Peariso	2011	Extract	USA	Sec	36	-	-	-	-	0.63	0.62	0.56	0.67	0.73	0.58	0.56	0.62	0.72	0.79
14	Todd	2006	Raw	USA	Sec	122	0.97	0.92	0.94	0.94	0.93	0.81	0.88	0.88	0.87	0.87	0.51	0.90	0.86	0.88
15	Wang	2011	Raw	China	Sec	23	0.98	0.90	0.91	0.96	0.88	0.84	0.85	0.80	0.92	0.88	0.82	0.80	0.92	0.80
16	Wong	2010	Raw	Malay	Pri/Sec	195	0.94	0.83	0.88	0.90	0.77	0.74	0.77	0.77	0.82	0.82	0.64	0.77	0.76	0.84
					Summary Statistics	2,508	0.96	0.88	0.91	0.93	0.85	0.79	0.80	0.83	0.82	0.78	0.74	0.82	0.80	0.80

Note: These analyses include all data sets comprised of 15 or more principals. All calculations are based upon Cronbach's alpha test of internal consistency. "Extract" refers to alpha coefficients extracted from research reports. "Raw" refers to our own analysis of data from secondary data sets.

PIMRS Reliability Teacher Form

In this section, the reliability of the PIMRS Teacher Form is discussed. Most previous researchers employed Cronbach's Alpha reliability for this form as well as for the Principal Form. Although this test is suitable when assessing self-report data gathered from principals, it violates a fundamental assumption of Cronbach's test when used to collect data from teachers.

The typical conceptualization of reliability is the ratio of true score variance to observed score variance.

$$\rho_{xx'} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} \quad \text{.....(1)}$$

In the computational formula of Cronbach's Alpha, the true score variance is observed score variance minus error variance. The notable point is that observed score variance is expressed by

the variance of every test takers' total score. It means the reliability is concerned with the stability of the individual test taker's score. For example, despite the teachers in the same school should be grouped together, the Cronbach's Alpha is concerned with the total variance of whole data without grouping.

We know that teachers in different school actually evaluate different principal, thus their score are affected by their principal's performance. Furthermore, any teacher has no chance to evaluate the principal in the other school. The structure for the teachers data is illustrated as figure 4.1.

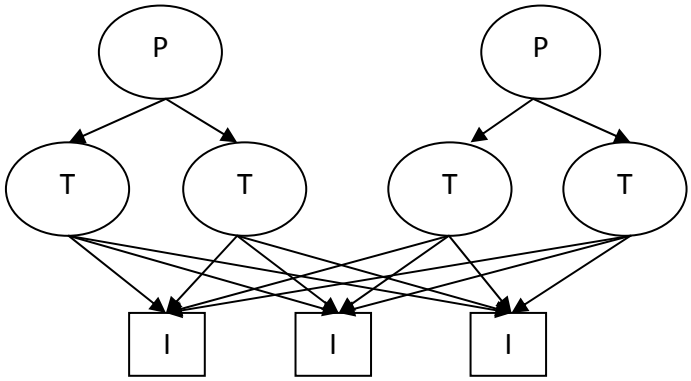


Figure 4.1. Structure of split-plot design where P is principal, T is teacher, and I is items.

The teachers' scores are nested within principal. All teachers evaluate their principals via the same items set, so the item effect is crossed with teachers. This design, with teachers nested within principals and crossed with items, is called a split-plot design. The purpose of PIMRS is assessing the principals' leadership performance by teachers rating. Thus, the reliability coefficient should be concerned with the dependability of principal means, rather than individual teacher ratings. If we still use Cronbach's alpha to represents the reliability, It is false to assume that each teacher response as if a different principal was being rated. Therefore, the reliability estimates should be based on assessments derived from teachers' ratings of their particular principals, school by school. According to generalizability theory, the random effects analysis of variance for split-plot design, with teacher (t) nested within principals (p), and crossed with items (i), yields estimates of the five components of variance presented in

Table 4.4. Summary of random effects ANOVA for split-plot design

Source of Variance	Sources as Confounded	df	E(MS)
P	p	$n_p - 1$!	$MS(p) = \sigma^2(e) + n_i\sigma^2(t,pt) + n_t\sigma^2(pi) + n_t n_i\sigma^2(p)$!
within p	!	!	
T Pt	t,pt	$n_p(n_t - 1)$!	$MS(p*t) = \sigma^2(e) + n_i\sigma^2(t,pt)$
I	i	$n_i - 1$!	$MS(i) = \sigma^2(e) + n_t\sigma^2(pi) + n_p n_t\sigma^2(i)$
Pi	pi	$(n_p - 1)(n_i - 1)$!	$MS(p*i) = \sigma^2(e) + n_t\sigma^2(pi)$
within pi	!	!	
Ti pti,e	ti,tpi,e	$n_p(n_t - 1)(n_i - 1)$!	$MS(p*t*i) = \sigma^2(e)$

The constants np, ni, and nt are , respectively, the numbers of principals, items, and teachers per school that are sampled. Because the number of teachers sampled from each principal are different, we suggest the harmonic mean can be used at nt.

The layout in Table 4.3 includes the main effect of principal, the main effect of items, the interaction between teacher and principal, the interaction between principal and items, and the error term. According to Kane (1976), the reliability of split plot design was referred as

$$\rho_p = \frac{\sigma_p^2 + \frac{1}{n_i}\sigma_{p\times i}^2}{\sigma_p^2 + \frac{1}{n_i}\sigma_{p\times i}^2 + \frac{1}{n_t}\sigma_{p\times t}^2 + \frac{1}{n_t n_i}\sigma_e^2} \quad \text{.....(2)}$$

, where the two interactions, $\sigma_{p\times i}^2$ and $\sigma_{p\times t}^2$, are supposed as the fixed number 0 in our testing situation. Firstly, the $\sigma_{p\times i}^2$ effect symbolizes the different items for different principals have different function or conception in their mind. For example, an item is used to assess the transformational leadership ability for one principal but to assess the transactional leadership for the other principal. This kind of different item function (DIF) should be examined and tested at the steps of making items, pretesting, and revising items. After the DIF examining, the items without DIF were retained and the items with DIF were revised or excluded in the test. Therefore, the $\sigma_{p\times i}^2$ effect in the finishing test will be assumed as zero. Secondly, the $\sigma_{p\times t}^2$ effect symbolizes the one teacher uses the different conception to evaluate the different principals. For example, one teacher used the leadership ability to evaluate the male principal, but used the attraction of her features to evaluate the female principal. For investigating this kind of interaction, it is needed to collect the group of teachers who evaluate many principals in different schools. There still are no relevant studies to demonstrate this effect. Therefore, we assume the effect does not exist. As the mention of the above two points, the Kane's function was revised as

$$\rho_{\mathrm{p}} = \frac{\sigma_{\mathrm{p}}^2}{\sigma_{\mathrm{p}}^2 + \frac{1}{n_{\mathrm{t}} n_{\mathrm{i}}}\sigma_{\mathrm{e}}^2} \quad \text{.....(3)}$$

We call this function as Gen reliability function. According to Table 4.4, the estimate of σ_{p}^2 and σ_{e}^2, $\hat{\sigma}_{\mathrm{p}}^2$ and $\hat{\sigma}_{\mathrm{e}}^2$, can be written as

$$\hat{\sigma}_{\mathrm{p}}^2 = \frac{MS_{\mathrm{p}} - MS_{\mathrm{p\times t}} - MS_{\mathrm{p\times i}} + MS_{\mathrm{e}}}{n_{\mathrm{t}} n_{\mathrm{i}}} \quad \text{.....(4)}$$

$$\hat{\sigma}_{\mathrm{e}}^2 = MS_{\mathrm{e}} \quad \text{.....(5)}$$

Therefore, the estimate of Gen reliability function was expressed as

$$\hat{\rho}_{\mathrm{p}} = \frac{[MS_{\mathrm{p}} - MS_{\mathrm{p\times t}} - MS_{\mathrm{p\times i}} + MS_{\mathrm{e}}]}{[MS_{\mathrm{p}} - MS_{\mathrm{p\times t}} - MS_{\mathrm{p\times i}} + MS_{\mathrm{e}}] + MS_{\mathrm{e}}} \quad \text{.....(6)}$$

We used this formula to compute the reliability in studies where we had obtained raw item-level teacher response data. This data set consisted of data gathered in 11 studies conducted between 2000 and 2012. The sample size of teacher respondents for the studies ranged from 95 to 1,610, with a mean of 329 teachers per study, and a total sample of 2313 teachers. Number of teacher in each school ranged from 3 to 28, and averaged 4.985 teachers were sampled from each school. The results of the Gen reliability of the raw data sets yielded a full-scale reliability of 0.99, with coefficients of .97 (*Defining a Mission*), .98 (*Managing Instruction*) and .98 (*Developing School Climate*) for the three dimensions. The combined reliability estimates for the 10 instructional leadership functions ranged from a low of .90 (*Maintaining High Visibility*) to a high of 0.95 on several functions. The reliability, $\hat{\sigma}_{\mathrm{p}}^2$ and $\hat{\sigma}_{\mathrm{e}}^2$ for whole scale and each dimension are listed in Table 4.5.

Table 4.5. Gen Reliability, Principal Variance, and Error Variance for PIMRS Dimensions

Dimension	$\hat{\sigma}_{\mathrm{p}}^2$	$\hat{\sigma}_{\mathrm{e}}^2$	$\hat{\rho}_{\mathrm{p}}$
Whole Scale	0.251	0.615	0.99
Defining a Mission	0.297	0.339	0.97
Managing Instruction	0.348	0.554	0.98
Developing School Climate	0.259	0.572	0.98

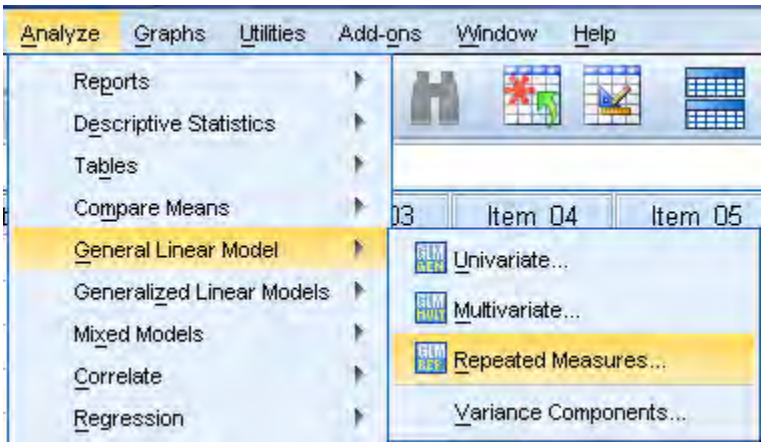
How to Use SPSS to Calculate Reliability of Teacher Data

In this section, we will describe the detail of the approach to calculate Gen reliability step by step. The software we used was the SPSS and Microsoft Excel. Here we have a data set as an example, which includes 107 teachers from 31 schools, averaging 2.38 teachers per school. The information required consists of the raw response data for 50 items and the teacher's school codes that identify the principal that each teacher rated. An example of table form of the data set is listed as Table 4.6.

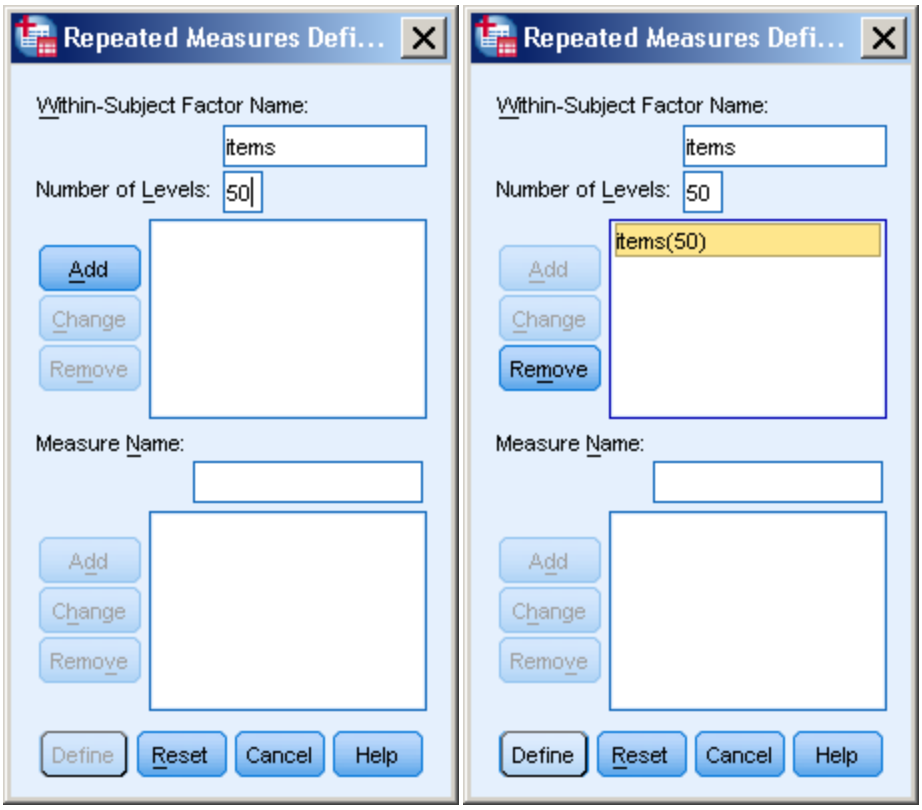
Table 4.6. Example of data requirements for calculating gen reliability

Teacher ID	School code	Item 01	Item 02	Item 03	…	Item 48	Item 49	Item 50
01	School 01	4	4	4	…	5	4	4
02	School 01	5	4	4	…	5	4	5
03	School 01	5	4	3	…	5	4	5
04	School 02	5	5	5	…	5	5	5
05	School 02	5	5	4	…	5	5	5
06	School 31	5	4	4	…	5	4	4
107	School 31	4	3	3	…	5	3	4

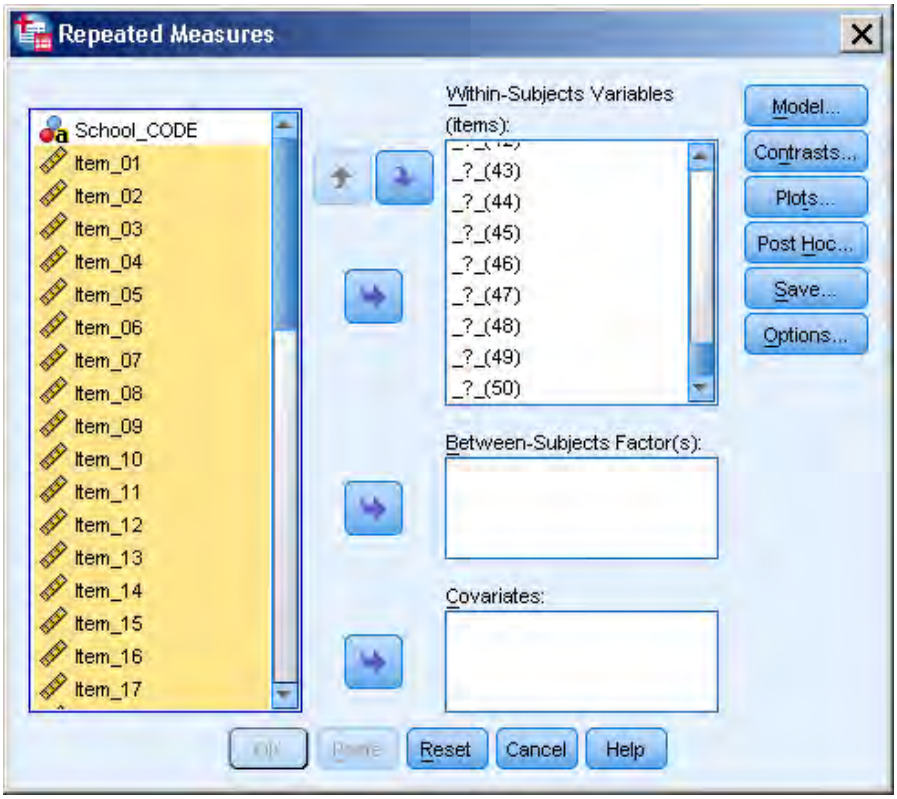
For calculating the materials of table 1, we should run the mixed design of ANOVA. In the tool bar of SPSS, we can click [Analyze] → [General Linear Model] → [Repeated Measures]



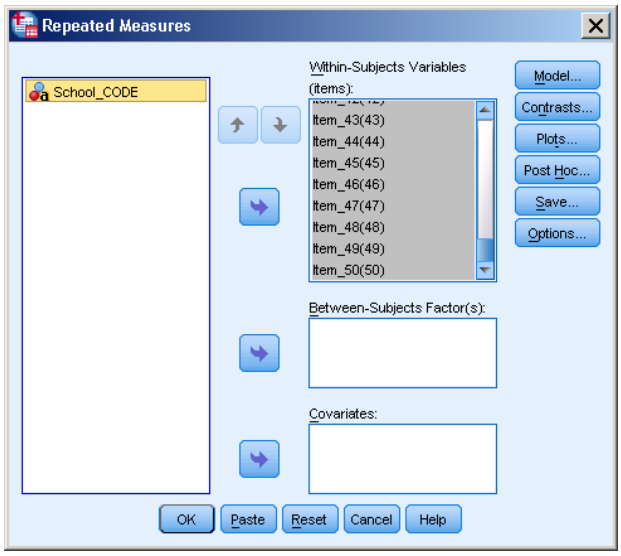
You will find the window as below. Type [items] in the Within-Subject Factor Name and [50] in the Number of Levels, then click [Add] and [Define] button.



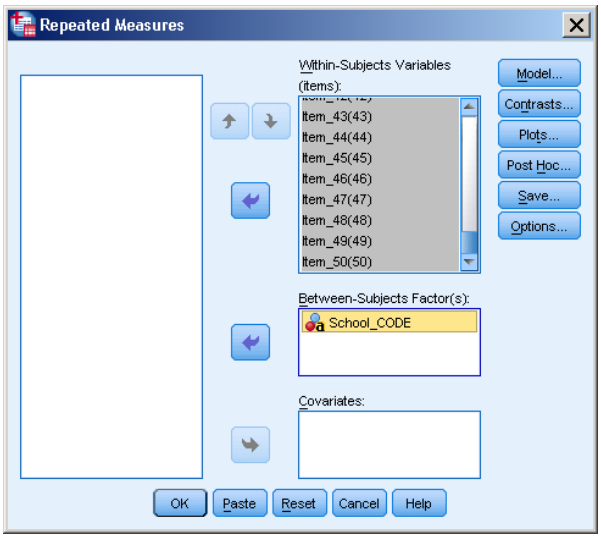
Then the following window will appear. Select all the items variable, [Item_01] ~[Item_50], and put them into the Within-Subjects Variables.



Select the variable [School_CODE], and put it into the Between-Subjects Factor(s):



Click [OK]. Start to run the SPSS.



Then we can obtain the result tables. We need the values in two tables, Test of Within-Subject Effects and Test of Between-Subject Effects. See Table 4.7 the Mean Square of items*School and Error(items) in the Within-Subject Table represent separately the $MS_{p\times i}$ and MS_e in the formula (6), and the Mean Square of School CODE and Error in the Between-Subject Table 4.8 represent separately the MS_p and $MS_{p\times t}$ in the formula (6). Therefore, in this example, we can obtain the materials, $MS_{p\times i} = 0.831$, $MS_e = 0.559$, $MS_p = 19.142$, and $MS_{p\times t} = 11.241$.

Finally we can calculate the Gen Reliability of this Data Set as:

$$\hat{\rho}_p = \frac{[19.142 - 11.241 - 0.831 + 0.559]}{[19.142 - 11.241 - 0.831 + 0.559] + 0.559} = 0.932$$

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
items	Sphericity Assumed	762.573	49	15.563	27.820	.000
	Greenhouse-Geisser	762.573	16.961	44.960	27.820	.000
	Huynh-Feldt	762.573	49.000	15.563	27.820	.000
	Lower-bound	762.573	1.000	762.573	27.820	.000
items * School_CODE	Sphericity Assumed	1140.169	1372	.831	1.486	.000
	Greenhouse-Geisser	1140.169	474.912	2.401	1.486	.000
	Huynh-Feldt	1140.169	1372.000	.831	1.486	.000
	Lower-bound	1140.169	28.000	40.720	1.486	.124
Error(items)	Sphericity Assumed	1096.438	1960	.559		
	Greenhouse-Geisser	1096.438	678.446	1.616		
	Huynh-Feldt	1096.438	1960.000	.559		
	Lower-bound	1096.438	40.000	27.411		

Table 4.7. Results of Mixed Design ANOVA from SPSS: Test of Within Subject Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	41591.012	1	41591.012	3699.898	.000
School_CODE	535.980	28	19.142	1.703	.060
Error	449.645	40	11.241		

Table 4.8. Results of mixed design ANOVA from SPSS: Test of between subject effects

Predicting the Effect of Sample Size on Reliability

This section will discuss how the sample size affects the Gen reliability, and we will show the suggested sample size for enough reliability. It can help you to decide the sampling design before you collect data. There are two kinds of sample size affecting the Gen reliability, one is the number of teachers within each school n_t, and the other is the number of items n_i. As the above section, the 0.99 reliability was based on 4.985 teachers within each school and 50 items tested. When those sample size changes, the different magnitude of reliability will be obtained. We can find that the larger nt and ni will make the influence of error term smaller in formula (3). It means that sampling more teachers in each school or putting more items in the exam will increase the Gen reliability.

According to Decision theory, we already obtained the estimate of principal variance and error variance equal to 0.251 and 0.615 from a large sample. Then we can predict the Gen reliability ρ'_p while the n_t and n_i change to n'_t and n'_i by the formula below.

$$\rho'_{\mathrm{p}} = \frac{\sigma^2_{\mathrm{p}}}{\sigma^2_{\mathrm{p}} + \frac{1}{n'_{\mathrm{t}} n'_{\mathrm{i}}}\sigma^2_{\mathrm{e}}} \quad \text{.....(7)}$$

When we decision to sample 3 teachers in each school and randomly sample 30 items as a short version scale from the total 50 items. We can give $n'_{\mathrm{t}} = 3$, $n'_{\mathrm{i}} = 30$, $\sigma^2_{\mathrm{p}} = 0.251$, and $\sigma^2_{\mathrm{e}} = 0.615$ to the formula (7) and then obtain the predictor of Gen reliability $\rho'_{\mathrm{p}} = 0.97$.

The succeeding series of figures list the predictive reliability for 10, 20, 30, 40, and 50 items for number of teachers in each school. For example, in Figure 4.4 we can predict that we will get the 0.919 reliability in the Managing Instruction dimension when we sample two teachers in each school and sample nine items as the questionnaire in this dimension.

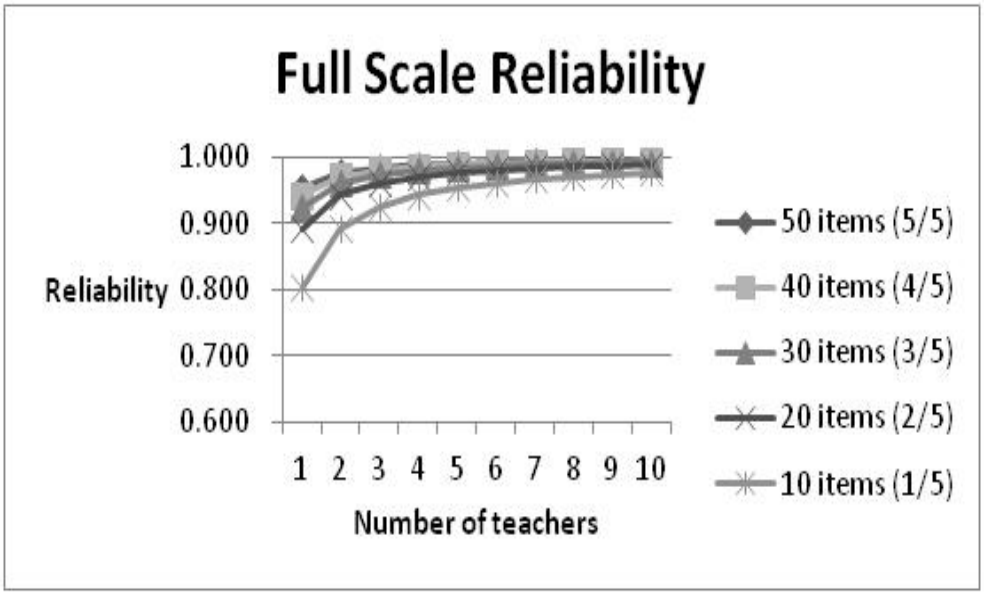


Figure 4.2. Effect of Sample Size on Reliability for Full PIMRS Scale

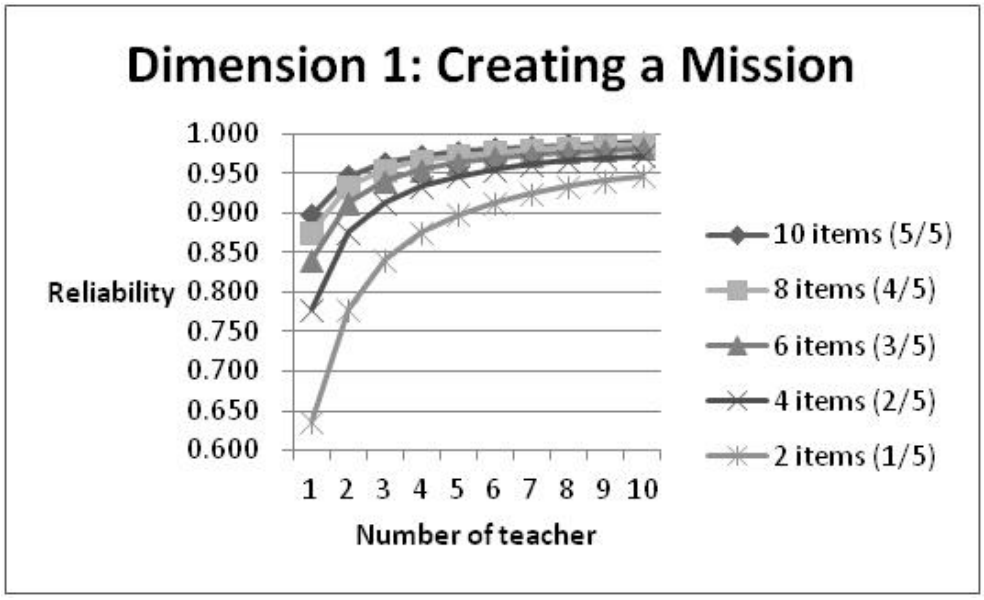


Figure 4.3. Effect of sample size on reliability for Creating a Mission

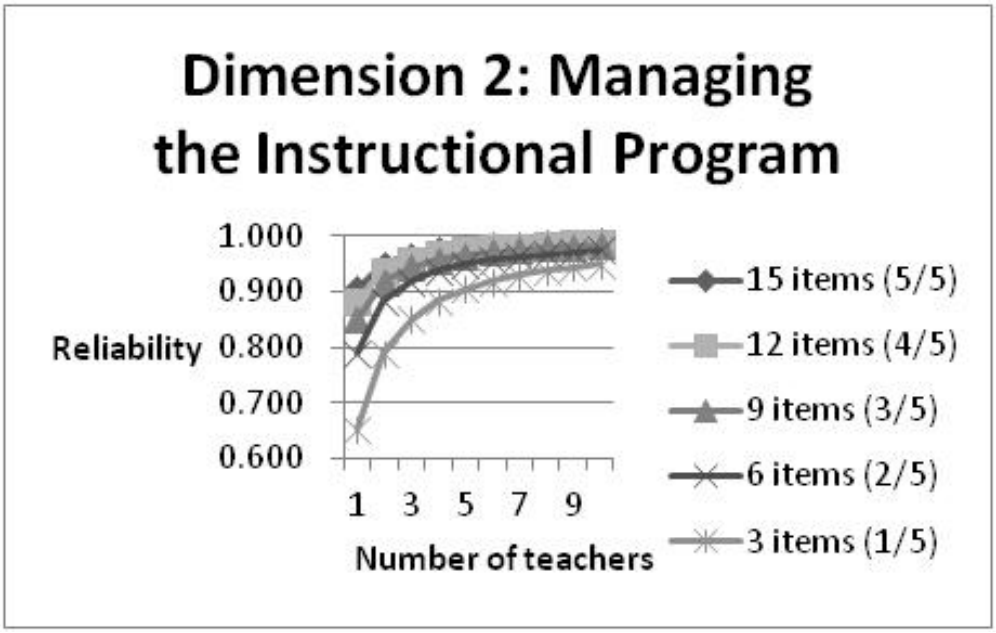


Figure 4.4. Effect of sample size on reliability for Managing the Instructional program

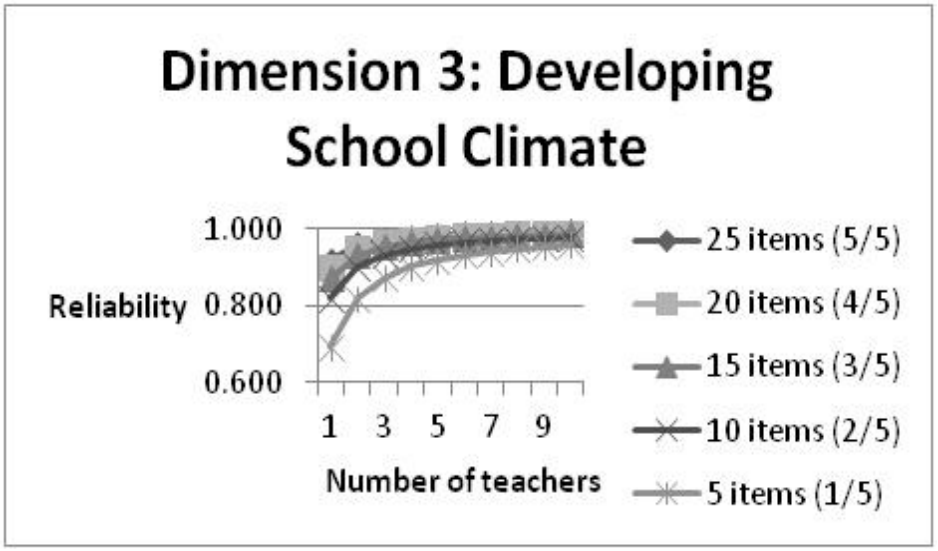


Figure 4.5. Effect of sample size on reliability for the School Learning Climate Dimension

These figures identify the teacher sample size required for each reliability standard and the number of items. If you plan to use fewer than the standard number of items, these figures can help you to decide on the number of teachers and items soon for reaching the reliability standard that you desire. For example, see Figure 4.3, if you desire the reliability of Defining School dimension to reach 0.9, you need to sample at least two teachers and use at least five items in this dimension.

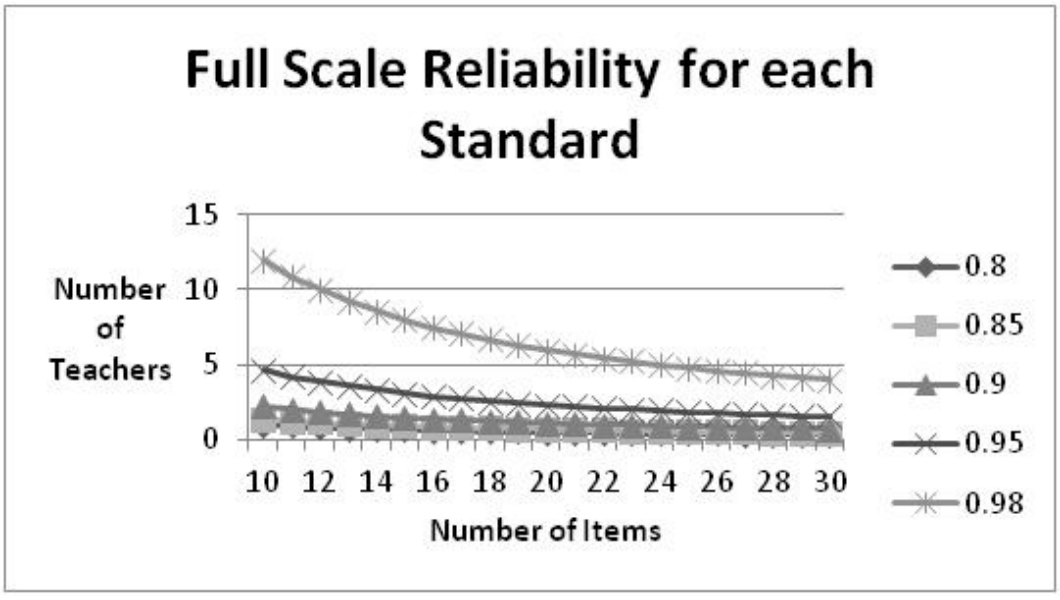


Figure 4.6 Sample size requirements for full scale reliability at various standards

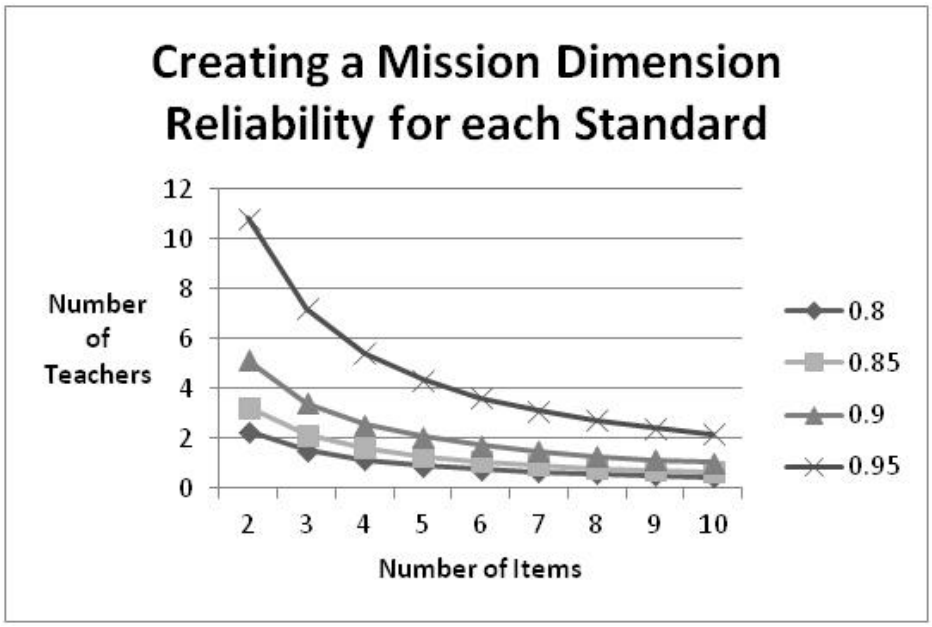


Figure 4.7 Sample size requirements for Mission Dimension needed to achieve different standards of reliability

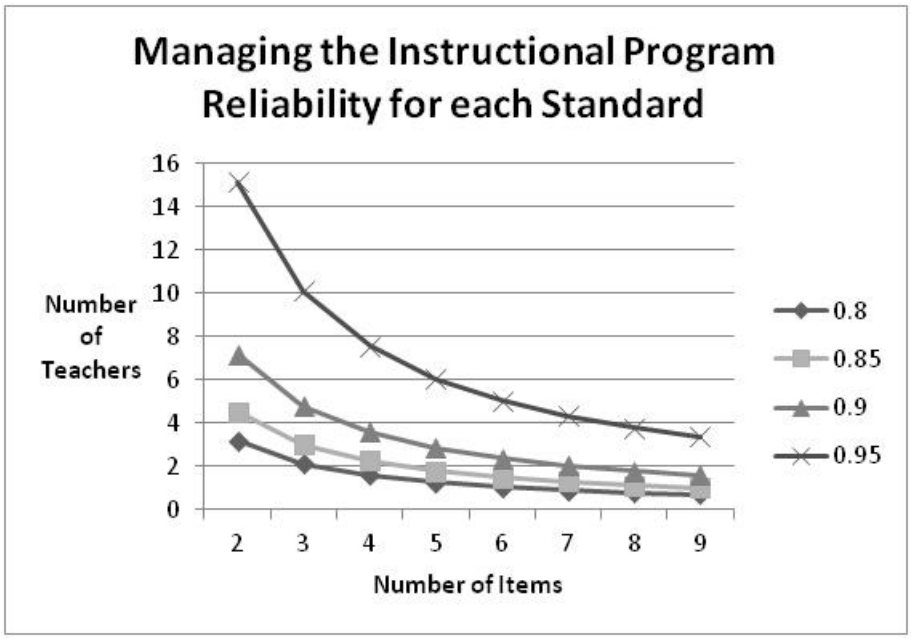


Figure 4.8. Sample size requirements for Managing the Instructional Program needed to achieve different standards of reliability

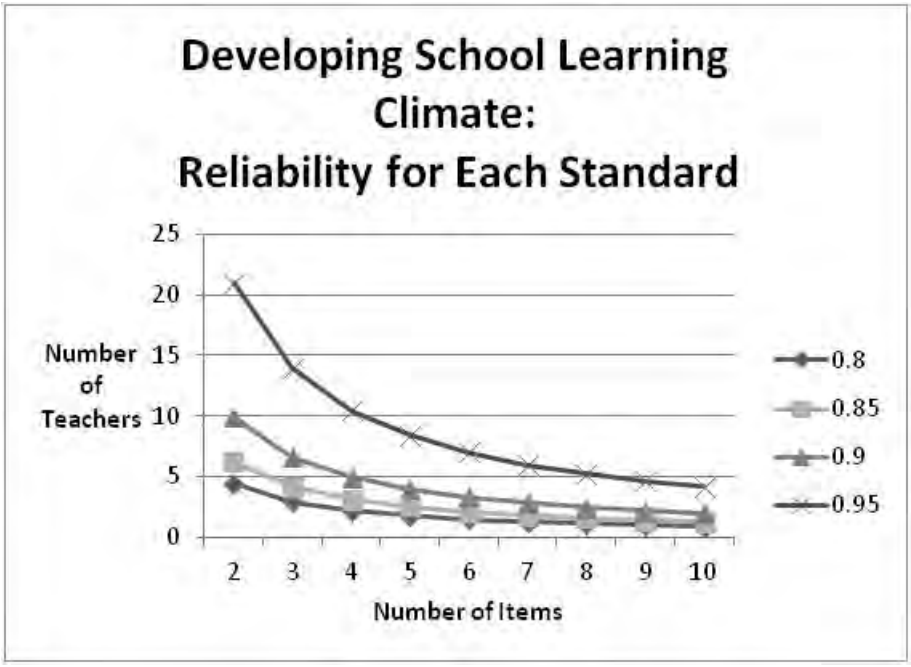


Figure 4.9 Sample size requirements for Managing School Learning Climate needed to achieve different standards of reliability

When you use this session to predict the reliability, there are some points you should notice. First, the sampling of teachers and items are random sampling. For example, we assume the sampled teachers from each school is random, and simultaneously we randomly sample 30 items from full 50 items, rather than based on any rule for sampling such as item contents and teacher's seniority. Secondly, this predictive value is the expected value of infinite sampling. It means the sampling error will cause the different results of reliability from each sample, and ρ'_{p} is the expected value of those reliabilities. For example, although we calculate the $\rho'_{\mathrm{p}} = 0.97$ to predict the reliability while sampling 3 teachers and 30 items, we may obtain lower or higher value of Gen reliability after collecting the real data. That is probably because the sampled teachers were all older fortuitously or the sampled items all tended to be agreed. Therefore, we should understand this predictive value just helps us to decide the sample design, but it can't be used as a reliability estimate for reporting in your own research.

Conclusion

This chapter sought to provide a comprehensive assessment of the reliability of the PIMRS instrument. The research questions revolved around the measurement properties of different PIMRS forms and scales, as well as its use in different contexts and for different purposes. In this final section, we first present a summary and interpretation of the results. Then we discuss the limitations and implications of the findings.

Summary and Interpretation

Meta-analyses of reliability results were conducted separately for the Principal and Teacher Forms of the PIMRS. In each case, we provided analyses for the whole scale as well as its component sub-scales. The pattern of results was quite consistent with Gay's (1992) observation that even in

highly reliable instruments, sub-scales based on fewer items tend to yield lower reliability than longer scales.

We concluded that the Principal Form of the PIMRS demonstrated moderately high reliability. Alpha coefficients exceeded 0.90 both for the whole scale and three dimension-level subscales, and 0.80 for the 10 function-level sub-scales. Although the number of studies and sample sizes varied considerably, we conclude that there was no substantial variation in the pattern of the results for the Principal Form of the PIMRS across school levels or between the two contexts included in this study, namely the USA and East Asia. Meta-analysis revealed that the Teacher Form demonstrated a consistently higher level of reliability for all three levels of scale measurement (i.e., >0.90) and across the measured organizational and cultural contexts.

Based upon the standards of reliability discussed in the second section of the chapter, we conclude that the Principal Form of the PIMRS can be used reliably for the purposes of either research or principal needs assessment. Moreover, the instrument appears to yield reliable for all three levels of measurement: the whole scale, three instructional leadership dimensions or the 10 functions. Principal self-report data are not typically employed as the basis for personnel decisions. Therefore, we conclude the reliability of the scale meets or exceeds the standard applied to research and needs assessment instruments.

The results further indicate that the PIMRS Teacher Form meets a standard of reliability required for use in personnel assessment as well as in research. We found that the instrument's reliability is replicable with primary, middle and secondary school principals, particularly in the USA. We do not consider the results for East Asia conclusive due to the limited coverage and small number of studies included in the sample. Thus, although we conclude that both forms of the scale meet high standards of reliability, we are more confident in the North American results as a result of the large number of studies and consistency in results.

Limitations and Implications

This chapter reporting on the reliability of the PIMRS raises a number of implications with relevance for both researchers and practitioners. While the data suggest that researchers should feel confident in using the PIMRS for collecting data on principal instructional leadership, our analyses found slightly different patterns of results for the two forms of the scales. In addition, it was clear that reliability for the full scale and three dimensions was higher than for the 10 leadership functions. Users will wish to take note of these differences in order to determine their relevance in relation to the purpose for which the data will be employed.

We were fortunate in being able to access data from 43 independent studies in order to conduct this meta-analysis of PIMRS reliability findings. While this represents a substantial database, it was but one third of the full set of PIMRS studies. Thus we note that although researchers were assiduous in reporting reliability estimates obtained from the early validation studies (e.g., Hallinger, 1983; Jones, 1987; Leitner, 1990; Taraseina, 1993), only about one third of documented users included a reliability analysis in their own reports. In our view, inclusion of reliability analysis should not be considered an option, especially when the instrument is being used in a setting that differs in meaningful ways from the original validation site(s). For example, we earlier noted that the PIMRS has been employed in 22 different countries. Yet, to date we were only able to obtain relevant reliability estimates for seven of those countries. We consider this an easily remedied, though important oversight. Since most of these studies were master and doctoral

dissertations, we recommend that supervisors be more stringent in making this a requirement for future studies.

In Chapter Two we included a quotation from Bridges (1982) who had implied that progress in our field would remain stunted in the absence of stronger conceptual frameworks and more robust research instruments. We noted that the PIMRS was developed in direct response to the need for research instruments that could contribute to a program of research on how leadership impacts learning. Scholars have concluded that important progress has been made on this issue over the ensuing three decades, and that the PIMRS has played a small part in this effort (e.g., Hallinger, 2011b; Hallinger & Heck, 1996; Leithwood et al., 2006; Louis et al., 2010; Robinson et al., 2008). Nonetheless, two important lines of inquiry await further elaboration. The first concerns continued exploration of the 'paths' through which leadership impacts learning. The second centers on unpacking the manner by which the school's context creates conditions in which leadership is exercised (Hallinger, 2011b; Hallinger & Heck, 1996; Leithwood, Patten, & Jantzi, 2010; Louis et al., 2010).

The data reported in this chapter indicate that both forms of the PIMRS instrument meet standards of reliability consistent with their respective purposes. As suggested in the immediately preceding paragraphs, however, corresponding analyses of the instrument's validity will shed further light on important issues of validity. This is the topic we take up in the next Chapter.

Chapter 5

Validation of the PIMRS

Background of the Concept of Validity

An appraisal instrument, in order to be useful, must provide data that not only are accurate and consistent (i.e., reliable) but that also measure the construct as conceptualized by the researcher (Lang & Heiss, 1998). In the words of Latham and Wexley, "A valid measure should yield consistent (reliable) data about what it is concerned with regardless of the time of day, week, or month the measures are taken, and regardless of who takes the measure" (Latham & Wexley, 1981, p. 65).

It is possible for an instrument to be reliable, but not valid. For example, a bathroom scale may consistently display five pounds higher than the actual weight of the person if repeated weightings are taken under the same conditions. This scale will yield consistent measurements, but the inference about how much one weighs will be faulty.

As noted earlier, data collected from a measurement instrument are typically used to make inferences about the relationships among the constructs under investigation. Proper use of data collected by an instrument requires that the user of the instrument be able to justify both the use of the instrument and the inferences drawn from the scores it yields.

An instrument must measure that which it purports to measure, or as Nunnally (1978) states, "strictly speaking, one validates not a measuring instrument but rather some use to which the instrument is put" (p. 87). Three broad uses of test scores may be distinguished which correspond roughly to three major categories of validity evidence (Cronbach, 1971). The uses of test scores are described briefly here and the approaches to their validation are examined in more detail below.

One use is to select or predict in which case, criterion-related validity is important. Using test scores to describe is a second use and in this instance, content validity is most relevant. The third use is using test scores to explain, which is the predominant interest in research. In these situations, construct validity is most appropriate. It is apparent then, that an investigator validates not the instrument itself, but a particular intended interpretation of the scores it yields. As with reliability, the type of validation procedure utilized depends on that intended interpretation.

Taxonomy of Validity

Validity analyses are commonly classified into construct-related, criterion-related, and content-related validation. However, these three categories are not distinct forms of validity. To provide a more communicative vocabulary for types of validation, Lissitz and Samuelsen (2007) suggested a systematic structural view of categorization of validity, which includes two primary investigative focuses and two primary perspectives. The categorization of primary investigative focuses considers the source of evidence. One source is internal evidence (internal validity), which involves the analysis of the test and the procedure of test and item development. The other source is external evidence (external validity), which involves the relationship between the test and other measures or criteria variables. The two primary perspectives are theoretical (conceptual) perspective and practical (empirical) perspective to inference of test score. The overview of the taxonomy of validation referred by Lissitz and Samuelsen is showed in Figure 5.1.

		Perspective	
		Theoretical	Practical
Investigative Focus	Internal	Latent Process	Content and Reliability
	External	Nomological Network	Utility and Impact

*Figure 5.1.*Taxonomy of test evaluation procedures

Internal validity includes content validity, reliability and latent process. Content validity and reliability belong to the practical perspective and latent process belongs to the theoretical perspective. In the theoretical perspective, a strict procedure of developing the test can be an evidence for evaluating the test. Examination of pattern of intercorrelation of items, factor analysis, and Rasch analysis also belong to this category of validity. In the practical perspective for internal validity, some evidences including the documentation of match between items and blueprint, and examination of item characteristics from test data are to be collected.

The four-building-block approach to test development (Wilson, 2005), describes explicit approaches and standards to develop a valid measurement. It can be used to assess theoretical internal validity. The four building blocks are *construct map*, *item design*, *outcome space,* and *measurement model*. In a construct map, the test developers classify the construct into a few ordinal categories and identify typical behaviors (item responses) to describe each category. Items are created to measure persons in each category according to the construct map. For each item, outcome space is defined and scoring rubric is developed. Finally, after test data are collected, Rasch measurement models are applied to produce measures and give feedback to the test development cycle. In Rasch models, a single latent trait is posted to account for the correlation among items. When item responses conform to the Rasch model, the derived Rasch measures are at the interval scale. In reality, items may not be clearly written and persons (subjects) may not respond to items reasonably. Thus, it is important in Rasch analysis to identify items or persons with poor fit.

External validity includes test utility and the impact of application. Predictive and concurrent validity (Cronbach & Meehl, 1955) are external validity with the practical and theoretical perspective, respectively. Predictive validity describes the relationship between the test measures and the criterion performance in the future. This evidence involves the utility or impact of the test. For example, score of a personality test is used to predict work performance in the future. Concurrent validity describes the relationship between the test measures and other measures that have been previously validated. The multitrait-multimethod matrix (Campbell & Fiske, 1959) is an example of external validity with the theoretical perspective. It considers both the theoretical relationship among tests and the effect of data collecting methods.

In this chapter, we document the validity of PIMRS via the four categories of validation showed in Figure 5.1. Firstly, the 4-building-block approach to the development of PIMRS was established. Wilson (2005) suggested tests developed via the approach usually have good psychometric properties. Then, we examined the PIRMS on (a) the theoretical internal validity using subscale inter-correlations and Rasch analysis; (b) the practical internal validity using content validity, school document analysis and differential item functioning; and (c) the theoretical external

validity using criterion-related validity (concurrent validity) and multitrait-multimethod analysis. The practical external validity is given in the external validity section.

Methods and Procedures

Datasets

Two kinds of data sets were collected: principal data and teacher data. The principle data were collected from 13 studies between 2008 and 2012, consisting of three from East Asia and nine from the USA (see Table 5.1). Five studies focused on secondary school principals, four on primary school principals, and three had collected data from principals across school levels. The sample sizes in these studies ranged from 6 to 195 principals, with a mean of 50 principals. There were altogether 649 principals.

The teacher data set were obtained from 13 independent PIMRS studies conducted between 2008 and 2012. In these 13 studies (see Table 5.2), respondents had completed the PIMRS Teacher Standard Form. The sample sizes of teacher respondents in the 13 studies ranged from 95 to 1,610, with a mean of 336 teachers. This represented a total of 4,370 teachers rating 651 principals.

Table 5.1 Data Source for principal self-report studies

Author	Year	Nation	N	School Level
Carr	2011	USA	6	Primary
Carson	2011	USA	77	Secondary
Gjelaj	2010	USA	36	Primary
Goldring	2012	USA	58	Primary/Secondary
Greb	2011	USA	31	Primary
Long	2008	USA	67	Secondary
Lyons	2010	USA	15	Secondary
Munro	2009	USA	35	Primary
Shafeeu	2011	Mald	10	Primary/Secondary
Todd	2006	USA	122	Secondary
Wang	2011	Chin	23	Secondary
Wong	2010	Malay	195	Primary/Secondary

For the purpose of assessing external validity, secondary data were collected from two doctoral studies, conducted by Dale (2010), and Greb (2011). Dale (2010) studied the effects of principals' transformational leadership and instructional leadership on student achievement. He used the Leadership Practice Inventory (LPI) to assess the transformational leadership ot a sample of 57 principals located in the Eastern Shore of Maryland (USA).

The Greb (2011) study was conducted in public elementary schools in the United States exploring the effect of transformational leadership, transactional leadership in conjunction with instructional leadership on student achievement. Data were collected from 31 principals. Greb's study assessed the three principal leadership styles through two scales: the PIMRS and the Multifactor Leadership Questionnaire (MLQ).

Table 5.2. Data Sources for teacher ratings of principals

Author	Year	Country	Teachers	No.of Schools	School Level
Adam	2012	USA	128	9	Secondary
Carson	2011	USA	95	77	Secondary
Dale	2011	USA	177	36	Primary
Fancera	2009	USA	580	100	Secondary
Goldring	2012	USA	1610	58	Primary/Secondary
Greb	2011	USA	107	31	Primary
Lyons	2010	USA	176	15	Secondary
Prema	2011	Malay	105	14	Primary
Shafeeu	2011	Malay	201	10	Primary/Secondary
Shatzer	2009	USA	280	37	Primary
Wang	2011	China	156	23	Secondary
Long	2008	USA	586	69	Secondary
Fulton	2009	USA	169	No school code	Secondary

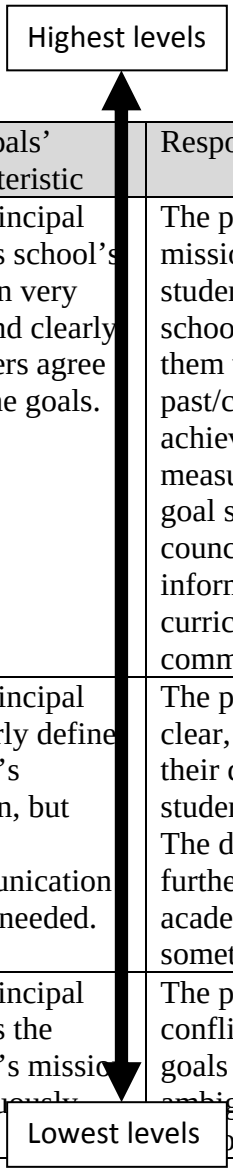
Results: Internal Validity of the PIMRS

In this section we report the results of various analyses of validity that have been conducted with the PIMRS. These include the original validation results as well as secondary analyses designed to offer insights into the internal and external validity of the PIMRS.

Internal Validity: Four Building Blocks

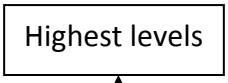
Construct Map

Figures 5.2 to 5.4 show the construct maps for the three dimensions of "*Defining the School's Mission*", "*Managing the Instructional Program*" and "*Promoting a Positive School Learning Climate*" in PIMRS, respectively. Each dimension has three levels: basic, proficient, and advanced. Each level has a specific definition and corresponding concrete behaviors.



Level label	Principals' characteristic	Responses to items
Advanced	The principal defines school's mission very well and clearly Teachers agree with the goals.	The principal works with staff to ensure the school has a clear mission and the mission is focused on academic progress of students. The Principal ensures that the importance of the school's goals is understood by discussing and reviewing them with staff. The goals could incorporate data on past/current student performance and staff responsibilities for achieving the goals. Performance goals could be expressed in measurable terms. Both formal communication channels (e.g., goal statements, staff bulletins, articles in the principal or site council newsletter, the school handbook, assemblies) and informal ones (e.g., parent conferences, teacher conferences, curricular meetings, other discussions with staff) are used to communicate the school's primary purpose.
Proficient	The principal regularly define school's mission, but more communication is still needed.	The principal, in collaboration with staff, sets the school a clear, academic goal that staff support and incorporate into their daily practice. Sometimes the principal uses data on student performance to frame the school's academic goals. The description of goal could be understood but it needs to be further explained. The principal regularly discusses the academic goal with teachers. However, the academic progress sometimes is not so clear or high.
Basic	The principal defines the school's missio ambiguously	The principal sets the school's mission, but it sometimes conflicts with academic or non-academic goals. The academic goals are not agreed by many teachers and the description is ambiguous and abstract. Some teachers do not know the ol's academic progress.

Figure 5.2. Construct map for Defining the School's Mission



Level label	Principals' characteristic	Responses to items
Advanced	The principal supervises and evaluates instruction, coordinates the curriculum and monitors studen progress very efficiently.	The principal effectively monitors students' progress. The principal provides sufficient instructional support to teachers and monitors classroom instruction through numerous informal classroom visits. The principal has a high degree of curricular coordination.
Proficient	The principal supervises, evaluates, and coordinates the curriculum well	The principal coordinates the classroom objectives and school objectives with teachers and evaluates classroom instruction. The principal sometimes reviews student work products. The principal regularly points out specific strengths and weakness in teacher's instructional practices in conferences or written evaluations. The principal often reviews the curricular materials and assesses the overlap between the school's curricular objectives and the school's achievement tests. The principal uses tests and other performance measure to assess progress toward school goals
Basic	The principal evaluates intuitively, coordinates not very clearly, an seldom monitor student progress	The principal sometimes fail to ensure the classroom priorities of teachers are consistent with the goals of the school. The principal may point out the strengths or weaknesses in teacher's instructional practices in an informal way. The principal coordinates the curriculum but no one is clearly responsible for coordinating the curriculum across grade levels. The principal seldom review the curricular materials or discuss student progress with individual teachers. The ipal assesses progress toward school goals by personal vation. Some teachers do not know the school's performance.

Figure 5.3. Construct map for Managing the Instructional Program

Highest levels

Level label	Principals' characteristic	Responses to items
Advanced	The principal protects instructional time, maintains high visibility, provides incentives for teachers, promotes professional development, and provides incentives for learning.	The principal models values and practices that create a climate and supports the continuous improvement of teaching and learning. The principal creates an "academic press" through the development of high standards and expectations and a culture that fosters and rewards continuous learning and improvement. The principal provides teachers with blocks of uninterrupted work time. The principal usually takes time to talk informally with students and teachers during breaks. The principal always makes the best use of both formal and informal ways of providing teachers with praise when it is deserved. The principal obtains the participation of the whole staff in important inservice activities and ensures those activities are consistent with the school's goals. The principal provides frequent opportunities for students to be rewarded and recognized for their academic achievement and improvement.
Proficient	The principal knows how to promote a positive school learning climate and sometime implements it.	The principals sometimes encourages teachers to use instructional time for teaching and practicing new skills and concepts. The principal understands that protecting instructional time is important. The principal does not allow that students are called to the office during instructional time. The principal sometimes attends extra- and co-curricular activities. The principal reinforces teachers' superior performance in staff meetings, newsletters, or memos. The principal sometimes supports the use in the classroom of skills acquired during inservice training.
Basic	The principal has some difficulty in promoting a positive school learning climate	The principal knows that protecting instructional time is important but sometimes fails to avoid that students are called to the office during instructional time. The principal does not spend time to visit classrooms to discuss school issues with teachers and students. The principal gives only higher salary to reward teachers for their efforts or performance. The ipal does not set aside time at faculty meeting for ers to share ideas or information from inservice activities. Sometimes the principal fails to consider whether teachers in their recognition and reward of student contributions to and accomplishment in class.

Lowest levels

Figure 5.4. Construct map for Promoting a Positive School Learning Climate

Item Design

Items design is the means by which instrument construction seeks to measure the theoretical construct in the real world situation. The development of items for inclusion in the PIMRS was reported in Chapter Three.

Outcome Space

The outcome space in the PIMRS is a five-point rating scale. Its scoring is straightforward: "1 = almost never; 2 = seldom; 3 = sometimes; 4 = frequently; 5 = almost always.

Measurement Model

A measurement model transforms categorical item responses that measure the construct. Because all items in the PIMRS use the same rating scale, the Rasch rating scale model (Andrich, 1978) was adopted. In the model, the probability of person n endorsing score j on item i (P_{nij}) is determined by the person's ability θ_n, item difficulty δ_i, and threshold parameter τ_j (relative to δ_i) for each score j. The relationship among these is:

$$P_{nij} = \frac{\exp(\sum_{j=1}^{J}(\theta_n - \delta_i - \tau_j))}{1 + \sum_{j=1}^{J}\exp(\sum_{j=1}^{J}(\theta_n - \delta_i - \tau_j))}, \quad (1)$$

where $J + 1$ is the number of categories in the item (in the PIRMS, $J = 4$). In the PIMRS, persons are principals; θ represents the rating given by a teacher or the principal herself, the higher the value of θ, the higher the rating; δ represents the item's threshold, the higher the value of δ, the more difficult it is for a principal or a teacher to obtain a high score on that item.

In addition to the three-building-blocks approach, studies have further tested the PIMRS for face validity and content validity (Hallinger et al., 1994; Howe, 1995; Jones, 1987; Leitner, 1994; O'Day, 1984). In these validation studies, five criteria were used:

1. *Content validity* — practical perspective of internal validity. Items making up each subscale of the instrument must be relevant to the critical requirements of the job; each item assigned to a subscale must achieve a minimum average agreement of .80 among a group of raters.
2. *School document support* — practical perspective of internal validity. An analysis of school documents related to the instructional management behavior of the principals should yield profiles of the principals' instructional management performance similar to those obtained from teachers on the questionnaire.
3. *Subscale inter-correlation* — theoretical perspective of internal validity. Groups of items within a subscale must inter-correlate more strongly with each other than with other subscales.
4. *Rasch analysis* — theoretical perspective of internal validity. The relationship between item responses and the construct is explicitly specified (e.g., Equation 1). Items or persons with poor fit to the model's expectation are identified. When data conform to the model, good measurement properties (e.g., specific objectivity and interval scales) are obtained.

5. *Differential item function* — practical perspective of internal validity. Items may function differently for different groups of persons. This is referred to as differential item functioning (DIF; Embretson & Reise, 2000) or measurement variance across groups of persons. Here, we were interested whether the PIRMS functioned differently between primary and secondly school levels.

Content Validity

Content validity refers to the degree to which the items which make up the subscales of the instrument are appropriate measure of various job functions. The procedures used to assess the content validity of the instrument followed those outlined by Latham and Wexley (1981, pp. 62-63). Content validity was determined by having persons knowledgeable in particular field, (i.e., instructional management) assign the potential items from a randomly ordered list into the functional categories. Latham and Wexley (1981) suggest that potential items should achieve at least eighty percent agreement among the raters in order to be considered a valid measure of a given functional category. The items which meet this standard within each functional category may be used as items in the instrument.

In this study, four professionals familiar with the instructional management functions of school principals (three principals and one vice principal), who had not been involved in the generation of the job behaviors, were enlisted to assist in the content validation of the instrument, They were each given a randomly ordered list of the potential items and a sheet of paper with eleven columns headed by the names of the functional categories (e.g., framing the school's goals or monitoring student progress). They were then asked to assign each item to the category in which they felt the item belonged. If an item did not fit in any of the categories, it was left unassigned.

After this process was completed, eighty-one items remained within the eleven functional categories. These items were reviewed with the participating superintendent and ten of the items were discarded in order to decrease the number of items in certain categories and the length of the questionnaire overall. The eleven categories and their assigned items, seventy-one in total, formed the rating instrument.

Table 5.3. Content validation: Average agreement on items among judges

Subscale	**Number of items**	**Average Agreement**
Frames Goals	6	91%
Communicates Goals	6	96%
Supervision/Evaluation	11	80%
Curricular Coordination	7	80%
Monitors Progress	8	88%
Protects Time	5	85%
Incentives for Teachers	4	100%
Professional Development	10	80%
Academic Standards	5	95%
Incentives for Learning	4	94%

School Document Analysis

The final test of the instrument's validity was accomplished through a comparison of the collected by the instrument with information related to the principals' instructional management behavior contained in school documents. The document analysis serves as an independent check on the perceptions collected by the rating instrument. The documents were used either to confirm or to call into question the validity of the questionnaire data on selected subscales.

Several types of school documents were collected and analyzed: school goal statements, school handbooks, principal newsletters, staff bulletin, staff meeting agendas and minutes, principal's written evaluations of teachers, and school site council minutes. The availability of these documents varied from school to school; consequently, in a few cases it was difficult to use the documents to validate the questionnaire data.

First, the documents were skimmed in order to determine how the data contained in them related to the various subscales. Subscales were selected for inclusion in the questionnaire — document comparison if the documents contained sufficient information related to the subscale's area of measurement. Sufficient documentary data existed for six of the eleven subscales: framing goals, communicating goals, supervision and evaluation of instructions, monitoring student progress, encouraging professional development for teachers, and providing incentives for student learning. It should be noted that even in the case of these subscales, there was not always a one to one correspondence between the items comprising the questionnaire subscales and the information obtained from the documents. For example, informal processes for communicating the schools goals to teachers are not picked up in the documents, although they are part of the subscale appraisal criteria.

The documents for each school were analyzed by the author on a subscale by subscale basis without prior knowledge of the school questionnaire ratings, thus minimizing rater bias. All documents from a school were scanned for information related to the each of the selected subscales. Each behavior or activity which corresponded to an item on the rating instrument was recorded along with the name and date of the source document. After all of the documentary data from a school had been scanned for a given subscale and the relevant behaviors or practices had been recorded, the principal was rated on that subscale. A one to five scale similar to that used in the rating instrument was utilized for this analysis.

After each principal had been rated on the subscales included in the document analysis, these ratings were compared with the appraisals derived from, the teacher questionnaire. First, each set of ratings (i.e., documentary and questionnaire) was put in rank order; then they were grouped into top and bottom thirds. These top and bottom groups of principals on the two sets of ratings were then compared in order to determine the extent to which the rankings based upon the teacher ratings of the principals matched those resulting from the document analysis. The degree of fit between the two sets of ratings varied across the subscales. The results of the rank order comparison is discussed briefly for each of the subscales.

Framing the School's Goals

This subscale was well covered by the available documentary data. The principals ranked one and two by the documentary analysis. The principal ranked third appears to have been overrated by his/her teachers in light of the documentary analysis. The instrument did not appear to discriminate

as well at the bottom third, based upon the document analysis. Only one of the lowest rated principals by the questionnaire was ranked in the bottom third by the document analysis.

Communicating the School's Goals

The school documents were only able to check the principal's use of formal settings and channels for communicating the school's goals. The instrument's ratings of the principals were confirmed for principals in both the top and bottom thirds.

Monitoring Student Progress

The documentary evidence on this subscale was fairly strong. The agreement between the ratings derived from the instrument and those from the documents is mixed for the top principals on this subscale, but is quite strong for those in the bottom third.

Supervision and Evaluation

There is strong documentary evidence on this subscale on both the principals' performance in supervising classroom instruction and on evaluating teachers. The agreement between the ratings and the documents is perfect for the principals ranked in the top three. At the bottom of the ratings, the documents were less supportive. Only one of the three principals ranked in the lower third on the ratings was also ranked in this group on the document analysis.

Professional Development

Although documents relevant to this subscale were obtained from the schools, they did not show very much variation in the behavior of the principals in this job function. This may be explained by the fact that most of the staff development in this district has been provided by or through the district office. This finding by the document analysis is congruent with the relatively low level of between school variance detected by the instrument.

Incentives for Learning

This subscale has very strong documentary evidence with which to compare the ratings. There is perfect agreement between the two sets of ratings for principals in the top third. At the bottom end there is also strong, though not perfect agreement.

The document analysis provides an independent check on the va1idity of the principals' performance ratings obtained with the instructional management rating scales. Although the strength of the validation varied across the six selected subscales, the document analysis generally supported the construct validity of those subscales.

The application of these methods to development of the PIMRS yielded an instrument that met common standards of internal validity in the original validation study conducted by Hallinger (1983). Studies by Jones (1987), Howe (1995), and Taraseina (1993) largely replicated the original validation study's results at the secondary school level in studies conducted in Canada, USA, and Thailand respectively. Readers interested in the specific results can consult the original research reports (Hallinger, 1983; Hallinger & Murphy, 1985; Howe, 1995; Jones, 1987; Taraseina, 1993; Wotany, 1999).

Subscale Inter-correlations

This criterion provides an assessment of the degree to which the persons being evaluated possess the quality or construct (i.e., instructional management) presumed to be reflected in the

performance instrument (Latham & Wexley, 1981 p.68). It compares the inter-correlation between each pair of subscales with each subscale's reliability coefficient. The purpose of such a comparison is to examine the extent to which the subscales seem to be measuring different aspects of the principals' behavior. Latham and Wexley clarify this issue:

To show construct validity of the appraisals, there should be agreement among knowledgeable observers of the employee's performance on each criterion. However, how employees are evaluated on one criterion (e.g., technical competence) should not correlate highly with low they are evaluated on another criterion (e.g., interpersonal ski1l). A high correlation among the different criteria is traditionally interpreted as evidence of halo error. That is, it is presumed that the raters are making one overall global rating without taking into account how each employee is really doing on the different aspects or dimensions of the job. The assumption underlying this argument is that it is unrea1istic to think that everyone who is outstanding on one criterion measure is equally good on all aspects of a job. A performance appraisal system with construct validity should reflect these strengths and weaknesses (1981, p. 69).

In theory the inter-correlations among subscales should be 1ow. This provides further confirmation to the test of discriminant va1idity that the subscales are measuring discrete job functions. In addition, the inter-correlation between subscales measuring different job functions should lower than the subscale reliability coefficients. This indicates that items within each subscale correlate more strongly with each other than with groups of items in other subscales; that is, items forming a subscale linked empirically as well as conceptually.

The subscale reliability coefficients were larger than the inter-correlation coefficients in all cases. This supports the earlier evidence suggesting that the items grouped conceptually as subscales belong together and are measuring different job functions.

It should be noted, however, that the inter-correlations among several the subscales are quite high; that is, many are above .60 (in both teacher and principal data). This suggests that despite the higher within subscale correlations, several of the job functions are closely related. In addition, all of the inter-correlation coefficients are statistically significant at the .01 level, indicating that the corre1ation were not likely to have resulted from chance. This result is not surprising given the relatively narrow job area (i.e., instructional management) being appraised. For example, one would expect closely related job functions such a framing and communicating the schools goals to be highly correlated. The fact that the inter-correlation between these two subscales is quite high (.85) lends support to the conceptualization of the subscales. Overall this test that the subscales are measuring different components of instructional management.

Rasch Analysis of Principal Data

The aforementioned content validity, subscale inter-correlation and school document analysis adopt the practical perspective towards assessment of the scale's internal validity. In contrast, Rasch analysis adopts a theoretical perspective on measuring the scale's internal validity.

The principal data and teacher data were analyzed separately. Even though the cognitive procedure of principals' self-report was different from that of teachers rating principals, we expected that the framework of instructional leadership would be adequate to both forms. One of the advantages of Rasch analysis is that it builds the hypothetical unidimensional line as construct map which locates items and persons according to their difficulty and ability measures. The second advantage is that each item can be tested the quality by item fit index (mean square error; MNSQ). In DIF analysis,

item invariance across different groups of persons (here, teachers or principals at primary or secondary school levels) is evaluated.

Item Fit

The dataset has 649 principals, including 329 principals in primary school and 320 principals in secondary school. For each of the three dimensions, a Rasch analysis was conducted separately. Table 5.4 shows the results of Rasch analysis on the dimension of *Defining the School's Mission.*

The first column "Sample size" represents the number of valid responses for each item. The "item difficulty" is the parameter of an item, with possible values ranging from negative infinity to positive infinity. The larger the value, the more difficult to receive a high score on the item. The "SE" is the standard error of the item parameter. The "Outfit MNSQ" indicates the item fit. When an item fits the Rasch model's expectation, the Outfit MNSQ has an expected value of unity. The farther away the value to unity, the worse the item fit. Normally, a value within the range of 0.6 to 1.4 would be considered as acceptable fit (Wright, Linacre, Gustafen & Martin-Lof, 1994), and a value within the range of 0.8 to 1.2 as good fit. The "item-test correlation" is the correlation between the item score and the total test score. A value above 0.5 would be considered as good. Generally, the item fit was good or cacheable and the item-test correlation was good, suggesting a good internal validity.

Table 5.4. Item fit statistics for "*Defining the School's Mission***"**

Item Label	Sample size	Function Label	Item Difficulty	SE	Outfit MNSQ	Good of fit	Item-Test Correlation
I_FSG_05	648	Frame	-0.5	0.07	0.76	Acceptable	0.59
I_FSG_02	632	Frame	0.09	0.06	0.81	Good	0.63
II_CSG_07	648	Comm	-0.45	0.07	0.83	Good	0.60
I_FSG_01	648	Frame	-0.57	0.07	0.86	Good	0.59
I_FSG_04	645	Frame	-0.97	0.07	0.95	Good	0.56
II_CSG_08	647	Comm	-0.13	0.06	0.97	Good	0.62
I_FSG_03	647	Frame	0.1	0.06	1.10	Good	0.63
II_CSG_10	450	Comm	1.6	0.06	1.17	Good	0.75
II_CSG_06	647	Comm	-0.15	0.06	1.19	Good	0.62
II_CSG_09	648	Comm	0.99	0.06	1.40	Acceptable	0.69

Table 5.5 shows the results of Rasch analysis on the dimension of "*Managing the Instructional Program*". The outfit MNSQ statistics suggested a good or acceptable fit and the item-test correlations were good. Thus, it can be concluded the data fit the Rasch model's expectation well.

Table 5.5. Item fit statistics for "*Managing the Instructional Program*"

Item Label	Sample size	Function Label	Item Difficulty	SE	Outfit MNSQ	Good of fit	Item-Test Correlation
IV_SEI_18	647	Supervise	0.16	0.06	0.71	Acceptable	0.64
III_CC_11	645	Coordinate	-0.31	0.06	0.79	Acceptable	0.61
V_MSP_22	647	Monitor	0.12	0.06	0.79	Acceptable	0.64
IV_SEI_19	645	Supervise	0.36	0.06	0.81	Good	0.65
V_MSP_23	645	Monitor	-0.24	0.06	0.88	Good	0.62
III_CC_14	647	Coordinate	-0.28	0.06	0.92	Good	0.62
III_CC_15	645	Coordinate	-0.27	0.06	0.96	Good	0.62
IV_SEI_16	645	Supervise	-0.42	0.06	0.98	Good	0.61
V_MSP_21	645	Monitor	0.41	0.06	1.01	Good	0.66
III_CC_13	645	Coordinate	-0.12	0.06	1.02	Good	0.63
IV_SEI_17	647	Supervise	-0.61	0.06	1.09	Good	0.59
III_CC_12	646	Coordinate	0.38	0.06	1.18	Good	0.65
IV_SEI_20	646	Supervise	0.31	0.06	1.28	Acceptable	0.65
V_MSP_24	647	Monitor	-0.2	0.06	1.37	Acceptable	0.62
V_MSP_25	449	Monitor	0.72	0.07	1.40	Acceptable	0.70

Table 5.6 shows the results for the dimension of "*Promoting a Positive School Learning Climate*". Items VII_PIT_34 (Cover classes for teachers until a late or substitute teacher arrives) and VII_PIT_35 (Tutor students or provide direct instruction to classes) did not have an acceptable fit. It appeared that these two items measured slightly different construct from the other items measured. Further studies are needed to look into possible reasons and revise their wordings. Given their MNSQ values were not very far away from unity, we decided to keep them in the scale

Table 5.6. Item Fit Statistic for **"***Promoting a Positive School Learning Climate***"**

Item Label	Sample size	Function Label	Item Difficulty	SE	Outfit MNSQ	Good of fit	Item-Test Correlation
IX_PPD_42	645	Prof Dev	-0.52	0.05	0.66	Acceptable	0.44
VIII_PIL_39	646	Inc Lnr	0.35	0.04	0.70	Acceptable	0.52
VIII_PIL_36	648	Inc Lnr	-0.10	0.05	0.73	Acceptable	0.48
IX_PPD_41	647	Prof Dev	-0.69	0.06	0.75	Acceptable	0.42
X_MHV_50	394	Visible	-0.19	0.07	0.81	Good	0.49
X_MHV_49	647	Visible	0.42	0.04	0.82	Good	0.53
IX_PPD_44	648	Prof Dev	-0.54	0.05	0.84	Good	0.44
IX_PPD_43	588	Prof Dev	-0.42	0.06	0.85	Good	0.45
IX_PPD_45	645	Prof Dev	0.07	0.05	0.87	Good	0.50
VI_PIT_30	589	Protect	-0.10	0.05	0.88	Good	0.48
X_MHV_48	646	Visible	0.59	0.04	0.90	Good	0.54
VII_PIT_32	648	Inc Tchr	0.13	0.05	0.92	Good	0.50
VIII_PIL_38	648	Inc Lnr	0.73	0.04	0.95	Good	0.55
VIII_PIL_37	648	Inc Lnr	-0.17	0.05	0.97	Good	0.47
VI_PIT_29	646	Protect	-0.50	0.05	0.98	Good	0.44
VIII_PIL_40	591	Inc Lnr	0.34	0.05	1.01	Good	0.52
X_MHV_46	646	Visible	-0.52	0.05	1.07	Good	0.44
X_MHV_47	647	Visible	0.12	0.05	1.10	Good	0.50
VII_PIT_31	647	Inc Tchr	-0.22	0.05	1.10	Good	0.47
VI_PIT_27	646	Protect	0.34	0.04	1.11	Good	0.52
VII_PIT_33	644	Inc Tchr	-0.18	0.05	1.18	Good	0.47
VI_PIT_26	644	Protect	-0.1	0.05	1.33	Acceptable	0.48
VI_PIT_28	645	Protect	0.04	0.05	1.38	Acceptable	0.49
VII_PIT_34	646	Inc Tchr	0.30	0.05	**1.48**	**Poor**	0.52
VII_PIT_35	648	Inc Tchr	0.82	0.04	**2.09**	**Poor**	0.56

Differential Item Function (DIF)

It is of great importance to check whether the items functioned differently between primary school principals and secondary school principals. If an item had statistically different difficulties for the two groups of principals, this item was deemed to exhibit DIF. Because the sample sizes of the two

groups were relatively large, even a small amount of difference would be statistically significant. We thus classified a DIF as substantial when the difference in the item difficulties between groups was greater than or equal to 0.5 logits. The ConQuest software (Wu, Adams & Wilson, 1997) was adopted to implement the DIF analysis.

In the “Defining the School's Mission” dimension, only item 4 had a substantial DIF (0.56 logits). In this item (Use data on student performance when developing the school's academic goals), the primary school principals had a larger difficulty in receiving a high score than that for the secondary school principals. This might be because student performance was less important for setting academic goals in primary school than in secondary schools.

In “Managing the Instructional Program” dimension, only item 24 exhibited a substantial DIF (0.68 logits). In this item (Inform teachers of the school's performance results in written form (e.g., in a memo or newsletter), primary school principals had a larger difficulty in receiving a high score than that for the secondary school principals. This might be because school’s performance results were less important in primary schools than in secondary schools.

In “Promoting a Positive School Learning Climate” dimension, items 33, 35, 45, and 46 had a DIF of 0.56, 0.76, 0.51 and 0.58 logits, respectively. In items 33 (Attend/participate in extra- and co-curricular activities) and 46 (Recognize students who do superior work with formal rewards such as an honor roll or mention in the principal's newsletter), primary school principals had a larger difficulty in receiving a high score than for the secondary school principals.

On the other hand, in items 35 (Tutor students or provide direct instruction to classes) and 45 (Set aside time at faculty meetings for teachers to share ideas or information from inservice activities), secondary school principals had a larger difficulty in receiving a high score than that for the primary school principals. This might be because it was easier for primary school principals to provide direct instruction to classes and to meet teachers than secondary school principals.

Correlation matrices for primary and secondary schools

Table 5.7 shows the correlation matrices of the three dimensions for primary and secondary schools. Generally, the correlations were very similar between these two school levels. Indicating the internal structure of the three dimensions was fairly invariant between school levels. Table 5.7 also shows the variances of the three dimensions for the two school levels. It seemed that secondary school principals had a larger variation on the three dimensions than primary school principals.

Table 5.7. Correlations between dimensions and variances of the three dimensions by primary schools and secondary schools

Primary				**Secondary**			
	D1	D2	D3		D1	D2	D3
D1	0.79			D1	1.34		
D2	0.89	1.66		D2	0.82	1.81	
D3	0.75	0.84	0.59	D3	0.72	0.84	0.71

Note. D1 is "Defining the School's Mission"; D2 is "Managing the Instructional Program"; D3 is "Promoting a Positive School Learning Climate."

Rasch Analysis of Teacher Form

For the dimension of "Defining the School's Mission," the fit statistics were located in the range of 0.6 to 1.4 for all items, suggesting a good fit. For the dimension of "Managing the Instructional Program," item15 (Point out specific weaknesses in teacher instructional practices in post-observation feedback (e.g., in conferences or written evaluations)) had a poor fit of (Outfit MNSQ=1.45), suggesting this item measured a slightly different concept from the other items in the same dimension. For the dimension of "Promoting a Positive School Learning Climate," items 34 (Cover classes for teachers until a late or substitute teacher arrives), 26 (Limit interruptions of instructional time by public address announcements) and 28 (Ensure that tardy and truant students suffer specific consequences for missing instructional time) had a poor fit of (Outfit MNSQ=1.55, 1.47 and 1.44), respectively, suggesting these three items measured slightly different concepts from that measured by the other items in the same dimension.

The teacher form has the same number of items as the principal form. To evaluate a principal, normally many teachers will be solicited. Given that a large number of teachers evaluate a principal, there is no need to have a lengthy teacher from. In Chapter Six, we documented how a shorter teacher form was created. The reliability and validity remained very high in the short teacher form.

Results External Validity

As noted earlier, we employed several tests to examine the external validity of the PIMRS. External validity seeks to understand if the instrument is measuring the construct of instructional leadership as conceptualized in theory and in practice.

Concurrent Validity

Due to data availability, we were not able to evaluate predictive validity. Only concurrent validity was provided, in which two relevant instruments, the Leadership Practice Inventory (LPI) (Kouzes &Posner, 1995, 2002) and the Multifactor Leadership Questionnaire (MLQ) (Bass, 1985; Bass, Avolio, Jung, & Berson, 2003), were compared. PIMRS assesses instructional leadership; LPI assesses transformational leadership, and MLQ assesses both transformational leadership and transactional leadership.

LPI was developed through a triangulation of qualitative and quantitative research methods and studies. It covers five areas of transformational leadership: Modeling the Way, Inspiring Shared

Vision, Challenging the Process, Enabling Others to Act, and Encouraging the Heart. The reliabilities for the LPI subscales are all above .60 (Kouzes & Posner, 2002). MLQ was first published by Bass (1985) with 63 items and later revised by Avolio et al. (2003) into a 45-item short form. This instrument was developed based on the full range leadership model, which included four subscales measuring transformational leadership: idealized influence; inspirational motivation; intellectual stimulation; individualized consideration; and two subscales measuring transactional leadership: contingent reward; and management-by-exception. The validity and reliability of MLQ have also been well tested (Shatzer, 2009).

Instructional leadership focuses on school mission and school climate, together with coordination of the curriculum. Interactions between different levels within the school are encouraged. School leaders with high instructional leadership emphasize a shared school mission and keen to promote a positive learning community. Transformational leadership aims to achieve organizational goals through promoting the integration of the members' motivations with the shared vision of the organization. School leaders with high transformational leadership are proactive. Transactional leadership, on the other hand, considers the exchange/negotiation between school leaders and their teachers based on the performance of the teachers. School leaders with high transactional leadership often make response to the teachers only when the teachers perform badly or the possible risk comes. They are action-oriented and primarily passive. Based on the definitions of instructional leadership, transformational leadership, and transactional leadership, we expected to find a moderately high correlation between instructional leadership and transformational leadership, and a low correlation between instructional leadership and transactional leadership.

In comparison, instructional leadership, as measured by PIMRS, highlights management of curriculum and instruction through evaluating teachers and student progress by including the dimension of Managing the Instructional Program. In contrast, transformational leadership, as measured by LPI and MLQ, emphasizes greatly on individuals (Individualized consideration).

The three dimensions of instructional leadership, as measured by PIMRS, were expected to be differently correlated with the dimensions of transactional leadership, as measured by MLQ. Specifically, the concept of Management-by-Exception in MLQ is not heavily covered in PIMRS, so that a nil or low correlation between them were expected. The concept of Contingent Reward in MLQ was at least partially covered in the dimensions of Managing the Instructional Program and Promoting Positive School Learning Climate, so that the correlations were expected to be higher than that between Management-by-Exception and the two dimensions of PIMRS.

Table 5.9 lists the expected correlations of the three dimensions of PIMRS and the subscales in LPI and MLQ. A correlation above 0.5 was defined as "high", between 0.3 and 0.5 as "moderate", between 0.1 and 0.3 as "low", between -0.1 and 0.1 as "nil" (Cohen, 1992). For example, the correlation between Defining the School's Mission and Modeling the Way was expected to be high and positive.

The datasets provided by Dale (2010) and Greb (2010) were analyzed. The participants in Dale's study included 57 principals from primary schools located in six counties on the Eastern Shore of Maryland. They were requested to take LPI and PIMRS concurrently. In Greb's study, 31 principals working in primary schools were collected in Wisconsin. They included 14 male and 17 female. Greb's study assessed the three principal leadership styles through two scales: PIMRS and MLQ.

The correlation between the subscales of PIMRS and LPI are presented in the upper panel of Table 5.9. Among these 20 correlations, only two were not consistent with the expectations, which were the correlation between Inspire Vision (LPI) and Managing Instruction (PIMRS), r = .49; and between Encourage Heart (LPI) and Managing Instruction (PIMRS), r = 60. The former correlation was expected to be low, but turned out to be moderate; and the latter was expected to be low but turned out to be high. The consistency rate was as high as 90%.

The lower panel of Table 5.9 lists the correlations between the subscales of PIMRS and MLQ. Among the 24 correlations, 14 were consistent with the expectations, resulting in a consistency rate of 58%, which can be considered satisfactory. The two largest discrepancies were between Individual Consideration and Managing Instruction (r = -.23), which was expected to be moderate positive, and between Contingent Reward and School's Mission (r = .77), which was expected to be nil. There may be several reasons. The sample size of Dale study is 57, and is 31 in Greb's study. Another possible explanation refers to the measurement development issue. Although they claim to be measuring the same concept, there might be differences between what the two scales are truly measuring.

In general, the consistence rates were high, especially for the correlations between instructional leadership and transformational leadership. Thus, it can be concluded that PIMRS has a good concurrent validity.

Table 5.9. Observed and expected Correlations between PIMRS and LPI and between PIMRS and MLQ based on the data of Dale (2010) and Greb (2010)

Insatrument/Subscales	**PIMRS (Total Score)**	**School's Mission**	**Managing Instruction**	**School Climate**
LPI: Transformational Modeling Way	0.62 (H)	0.52 (H)	0.44 (M)	0.66 (H)
Inspire Vision	0.68 (H)	0.56 (H)	0.49 (L)*	0.73 (H)
Challenge Process	0.75 (H)	0.58 (H)	0.50 (H)	0.84 (H)
Enable Act	0.80 (H)	0.63 (H)	0.61 (H)	0.86 (H)
Encourage Heart	0.76 (H)	0.62 (H)	0.60 (L)*	0.78 (H)
MLQ: Transformational				
Idealized influence	0.38 (M)	0.43 (M)	0.16 (L)	0.45 (H)*
Inspirational motivation	0.51 (M)	0.54 (H)	0.31 (L)*	0.53 (H)
Intellectual stimulation	0.77 (M)*	0.31 (M)	0.68 (L)*	0.97 (H)
Individual consideration	0.13 (L)	-0.05 (N)	-0.23 (M)*	0.63 (M)*
MLQ: Transactional				
Contingent reward	0.77 (M)	0.77 (N)*	0.750 (H)	0.49 (M)

Manage-by-exception	0.09 (L)	-0.02 (N)*	0.047 (N)	0.19 (L)

*inconsistent with the expectation; labels in the parentheses are the expected correlations: H = high; M = moderate; L = low, N = nil.

Multitrait-Multimethod

The multitrait-multimethod matrix (MTMM) method (Campbell & Fiske, 1959) was adopted to evaluate the convergent and divergent validity of PIMRS. There were three "traits" –instructional leadership measured by PIMRS, transformational leadership measured by MLQ, and transactional leadership measured by MLQ – and two "methods" – principal self-evaluation and teacher ratings of principals.

The analyzed dataset also were provided by Greb(2010). Greb measured 31 principals' instructional leadership and transformational leadership by self-evaluation of PIMRS and MLQ. Concurrently, Greb also collected 107 teachers working with these 31 principals to rate their principal by PIMRS and MLQ. The participants all were from the primary schools in Wisconsin.

Table 5.10 showed is the correlation matrix among the three traits and two methods. The diagonals were the reliabilities and the off-diagonals were correlations. For principal data, the reliability was the Cronbach's alpha, whereas for teacher data, the reliability was the Gen reliability (see Chapter Four). It appeared that the reliability for PIMRS and transformational leadership was higher than that for transactional leadership.

Convergent validity was the degree to which traits that should be related theoretically are related empirically; whereas discriminant validity was the degree to which trait that should not be related theoretically are not related empirically. In theory, the correlation between PIMRS and transformational leadership should be higher than that between PIMRS and transactional leadership. As shown in Table 5.10, the correlation was .41 and .35, respectively, for the principal data; and was .70 and .07, respectively, for the teacher data. It seemed that the PIRMS teacher form had a higher convergent validity and discriminant validity than the PIRMS principal form.

Table 5.10 also shows the correlations between methods on the same trait (monotrait-heteromethod). The correlation was .21 for PIMRS, .07 for transformational leadership, and -.01 for transactional leadership. Although these correlations were very low, they were reasonable and consistent with the literature. The systematic difference between self-evaluation and others ratings was investigated (Murphy & Deshon, 2000). It is possible that principals may evaluate themselves according to whether they can achieve the behavior described in the test. Teacher ratings might emphasize observed practices or even the effects of those practices. Regardless of why the differences occur, the low correlation between teacher and principal ratings on the instrument suggest the suitability of employing multiple assessments of principal practice.

Table 5.10. Multitrait-multimethod matrix among PIMRS, transformational leadership, and transactional leadership

Method！		**Principal**			**Teacher**		
！	Instrument	PIMRS	Trans-formation	Trans-action	PIMRS	Trans-formation	Trans-action
Principal	PIMRS	(0.84)	！	！	！	！	！
	Transformational	**.41***	(0.83)	！	！	！	！
	Transactional	**0.35**	**0.08**	(0.60)	！	！	！
Teacher	PIMRS	*0.21*	-0.08	0.04	(0.93)	！	！
	Transformational	0.31	*0.08*	0.02	**.695****	(0.94)	！
	Transactional	0.06	-0.09	*-0.01*	**0.07**	**0.24**	(0.62)

Note: ** p<0.01, * p<0.05. Numbers in parentheses are reliability; numbers with boldface are monotrait-heteromethod correlations; and numbers with underline and italic are heterotrait-monomethod correlations.

Conclusion

In this chapter, we adopted the taxonomy of validation proposed by Lissitz and Samuelsen (2007). This offers a comprehensive and rigorous perspective on testing instrument validity. The taxonomy incorporated multiple measures of internal and external validity.

The internal validity of PIRMS was assessed in multiple ways. Although the PIMRS was developed in 1982, its development was still consistent with current validation methods such as the four-building blocks procedure (Wilson, 2005). Content validity and school documentation analysis justified the appropriateness of item content, and subscale inter-correlation analysis demonstrated a good internal structure. Rasch analysis further demonstrated that most items within each subscale had a good fit, such that the assumption of unidimensionality was supported. Finally, very few items were found to exhibit DIF across school levels, there by supporting a conclusion of model invariance.

The external validity of PIMRS was evaluated by concurrent validity and multitrait-multimethod methods. The empirical relationship between PIMRS and transformational leadership and transactional leadership supported a good concurrent validity of PIRMS. Although the principal data had a different pattern of correlation with external measures from that of the teacher data, both datasets showed that PIRMS had a higher correlation with transformational leadership than that with transactional leadership, which matched the theoretical expectation. In short, all the validation procedures demonstrate a high validity in PIMRS.

Chapter 6

Developing a PIMRS Teacher Short Form

Given the widespread continuing use of the PIMRS, the authors recently updated information on the instrument's reliability (Hallinger, 2013) and validity (Hallinger, 2013). In conjunction with this research and development (R & D) effort, the authors also engaged in instrument revision aimed at increasing the efficiency and effectiveness of the PIMRS as a tool for data collection. One of these revisions entailed creating a shortened version of the PIMRS Teacher Form. Although the standard form of the PIMRS is not overly long by survey standards (i.e., 50 items), in leadership research instruments are often used along with complementary scales that measure moderating and mediating constructs (Baron & Kenny, 1986; Hallinger & Heck, 1996; Leithwood et al., 2006). A shorter instrument would increase efficiency in data collection, thereby reducing an impediment (i.e., time) to collecting data from teachers. A shorter survey could also improve the instrument's effectiveness, by increasing quality of teacher responses (Gay, 1992). These represent useful goals as long as the shortened version of the instrument continued to meet high standards of reliability, validity, and utility.

This chapter reports on the procedures and results of this R & D project aimed at developing a PIMRS Teacher Short Form. More specifically, we describe the steps undertaken in instrument revision, and report results with respect to reliability, validity, and utility of the PIMRS Short Form. The paper begins with a review of background information on the PIMRS instrument, proceeds to discuss the methods employed in instrument revision, and then presents the results. In the concluding section, we discuss the implications of this revision of the PIMRS for researchers, staff developers, and evaluators.

Method

In this section we outline the procedures used to reduce the length of the PIMRS Teacher Standard Form. This R & D process entailed the use of secondary data (see Hallinger, Wang & Chen, 2013b). Therefore, prior to discussing the steps in instrument development, we briefly discuss the data that were used for the analyses reported in this study.

Data Sources

In order to reduce the length of the PIMRS Teacher Standard Form, we had two choices. We could either collect new data or reanalyze data collected in previous studies. In either case, we would require 'item-level data' in order to conduct the necessary analyses. We were fortunate to be able to locate and obtain data collected in 13 independent PIMRS studies conducted between 2008 and 2012 (see Table 1).

In these 13 studies, respondents had completed the PIMRS Teacher Standard Form. The sample size of teacher respondents in the 13 studies ranged from 95 to 1,610, with a mean of 336 teachers per study. This represented a total sample of 4,370 teachers rating 651 principals. Respondents

included teachers in both primary and secondary schools. This combined dataset was employed in the tests used to assess the reliability and internal validity of the PIMRS Teacher Short Form.

Insert Table 1 about here

Data Analysis

Any effort to reduce the length of an established instrument must be evaluated against several criteria. Reduction in the number of items should minimize response bias while maintaining high reliability, validity, and comparability with results obtained from the longer form of the instrument. These three criteria were applied in the current study.

An important decision concerned the 'levels' of scale measurement that we would seek to support in the shorter form of the instrument. As noted earlier, the PIMRS Teacher Standard Form can yield 'scale scores' for the whole instrument, its three dimensions, 10 instructional leadership functions. We decided for reasons of utility and measurement quality that instrument revision should seek to support measurement of the full scale and the three instructional leadership dimensions (i.e., Defining School Mission, Managing the Instructional Program, Developing a Positive School Learning Climate). In terms of the primary uses of the PIMRS (i.e., research, principal evaluation), obtaining a profile based on the three instructional leadership dimensions was deemed sufficient. Moreover, it would have been difficult to measure the 10 functional subscales with a sufficiently high level of data quality using substantially fewer items. That is, the PIMRS Teacher Short Form should yield stable scores that represent the same latent trait measured by the Standard Form.

Rasch analysis (Rasch, 1960) is widely used in psychological measurement. Rasch analysis can provide a confirmation for measurement models, thereby offering insight into the internal validity of an instrument. More specifically, through Rasch analysis we can check the model-data fit for each item and then select items of higher quality for our questionnaire from the existing pool of items. Data output from Rasch analysis also offers an indication of whether items are relatively easier or more difficult for respondents. This allows the instrument developer to create an instrument that will differentiate the responses of persons (i.e., teachers in this study). The item selection strategy used in developing a PIMRS Teacher Short Form entailed the use of three indices derived through Rasch analysis. Those can be obtained by using WINSTEPS software (Linacre, 2005).

The first index used in assessing the quality of items is item difficulty. In the Rasch rating scale model (Andrich, 1978), the probability of person *n* endorsing score *j* at item *i* is divided into the person's ability θ_n, item difficulty δ_i, and step parameter τ_j (relative to δ_i) for each score *j*. The relationship among these is written as:

$$P_{nij} = \frac{\exp(\theta_n - \delta_i - \tau_j)}{1 + \sum_{j=1}^{J} \exp(\theta_n - \delta_i - \tau_j)}, \qquad \text{(Equation 1)}$$

where $J + 1$ is the number of categories in the item. In this study, persons are teachers; θ represents the rating given by a teacher to his or her principal, the higher the value of θ, the higher the rating; δ represents the item's threshold, the higher the value of δ, the more difficult it is for a principal to receive a high score on that item. In other words, high item difficulty means that the teachers less frequently award a higher score to their principal on those items. An optimal test design includes items that represent a full range of difficulty. This means that the instrument is capable of assessing every level of principal task performance. Therefore, the first strategy is to ensure that all levels of difficulty are maintained when the number of items is reduced.

Item difficulty can be identified by examining the pattern of actual scores among a sample of principals. A Wright Map (Wright & Masters, 1982) displays the distribution of item difficulty in relation to the distribution of teachers' ratings of their respective principals along a vertical line from the highest difficulty at the top to the lowest at the bottom. The distribution of teachers' ratings of their principals is shown along the left hand side of the line and the distribution of item difficulty on the right hand side. The mean item difficulty is located at the zero point on the vertical line.

Using the Wright Map, we can clearly identify the distribution of items along these two parameters. The map also profiles the number of items and teachers located on each level of the two parameters. This information provides insight into how the item distribution changes according to the inclusion of different 'sets' of items. The optimal result is achieved when both the item and rater (i.e., teachers) means and the variance and shape of the distributions are similar.

The second index used in assessing item quality is the item fit statistic 'outfit mean square' (outfit MNSQ). In instrument construction, we propose that items located in the same dimension or sub-scale are assessing the same latent trait (construct). This MNSQ measures the fit of the observed data to the expectation of the Rasch model. Wright and colleagues (Wright, Linacre, Gustafen & Martin-Lof, 1994) recommended that scale items demonstrate an acceptable fit if their OUTFIT MNSQ falls within the range of 0.6 to 1.4, and a very good fit if they fall in a range from 0.8 to 1.2. With this in mind, the second strategy employed in item selection is to eliminate items with a MNSQ outside the range from 0.6 to 1.4 first. The next step involves inspection of items in order to maximize the number that fall within the 0.8~1.2 range.

The third index used in item selection was the item-test correlation. This statistic, commonly used in test development, is the correlation between the item score and the total score of the corresponding scale dimension. Items with a low item-test correlation (e.g., < 0.2) are generally eliminated from the scale (Linacre, 2005). Our goal in development of the PIMRS Teacher Short Form was twofold: (a) for all items to have an item-test correlation above 0.2, and (b) for a majority of items to yield an item-test correlation higher than 0.5.

At the same time that we viewed the statistical results with respect to specific items, it was also necessary to examine the 'content distribution' of items. Although the PIMRS Teacher Short Form would not yield function-level scores, it should continue to maintain a representative selection of

items drawn from the 10 function-level subscales. Therefore an additional step aimed at maintaining the content validity of the three dimensions was incorporated into our procedures. At the point of scanning the data on item fit and difficulty, we also attended to the distribution of items across function-level subscales and used this as an additional criterion in item selection.

After calibrating the teachers' abilities through Rasch analysis, we examined the structure of the conceptual framework through confirmatory factor analysis (CFA). CFA assesses the extent to which the data fit to the PIMRS conceptual framework (e.g., tin this case, he three dimensions that comprise the PIMRS). CFA has become a standard approach for examining the internal validity of tests. In this study, we used several fit indices to determine how well the data fit to the conceptual framework, including goodness of fit index, Tucker Lewis index, root mean square error of approximation, and standardized root mean square residual. CFA was applied to the 22 items comprising the PIMRS Teacher Short Form. As noted earlier, the sample consisted of 4,341 teachers rating 651 principals.

Following these procedures focusing on item selection and internal validity, we retested the resulting scale to ensure that the PIMRS Teacher Short Form would still meet a high standard of reliability. We used the Gen Theory test of reliability discussed earlier in the paper for this purpose. The reliability test was applied at two levels to produce a total instrument reliability coefficient as well as coefficients for each of the three dimensions. Since the revised instrument could potentially be used for multiple purposes (i.e., research, needs assessment, personnel evaluation), we wished to achieve reliability coefficients above .90 if possible.

The last criterion used to assess the measurement properties of the PIMRS Teacher Short Form was comparability between results obtained from the PIMRS Teacher Standard Form and the Short Form. Two analyses were conducted. The first tested the correlation in raw scores between the Standard Form and Short Form on the whole scale as well as the three dimensions. A correlation that approaches 1.0 would indicate a high level of comparability.

The second test of comparability analyzed differences in Rasch measures between the two forms on each dimension. A zero difference would again indicate a high level of comparability. More specifically, let $\hat{\theta}_{n,\text{standard}}$ and $\hat{\theta}_{n,\text{short}}$ be the Rasch measures for teacher *n* on the standard and short forms, respectively; and $SE(\hat{\theta}_{n,\text{standard}})$ and $SE(\hat{\theta}_{n,\text{short}})$ be their standard errors, respectively. Under the null hypothesis of no difference between forms, the following statistic should follow approximately the standard normal distribution:

$$\frac{\hat{\theta}_{n,\text{standard}} - \hat{\theta}_{n,\text{short}}}{\sqrt{\text{SE}(\hat{\theta}_{n,\text{standard}})^2 + \text{SE}(\hat{\theta}_{n,\text{short}})^2}} \sim Z\,. \qquad \text{(Equation 2)}$$

When the .05 nominal level was used for significance, it was expected on average approximately 5% of teachers would have a statistically significant difference.

Results

In this section of the paper we describe results of our effort to produce a PIMRS Teacher Short Form. As noted above, the steps entailed an item-level analysis for each of the three dimensions that comprise the PIMRS.

Item Selection

According to the MNSQ, most of the 10 items comprising the first dimension, Defining the School's Mission, had a good fit. Items in bold font in Table 2 (i.e., items 08, 09, 10, 03, 07) were likely candidates for item reduction based on lower quality fit. The item-test correlation of the remaining items was above 0.5. Considering optimal content coverage for this dimension, we selected item No. 08 instead of No. 02 for inclusion in the Short Form.

Insert Table 2 about here

The Wright Map shown in Figure 1 shows the distribution of these 10 items based on their difficulty. The bold items indicate likely candidates for item reduction for the Short Form. We noted that the distribution of person ability (i.e., the teacher rating) appears to be somewhat higher than item difficulty. Although not optimal, this could be due to the fact most principals engage proactively in mission-building and goal-setting in their schools. This data trend did not cause us to change our item selection.

Insert Figure 1 about here

Table 3 shows the MNSQ for the second dimension, Managing the Instructional Program. According to the approach described above, we observed that in boldface type (i.e., items 15, 25, 24, 18, 19, 22, 13, 14) were likely candidates for item reduction. All items had MNSQ between 0.8 and 1.2, and item-test correlations exceeding 0.5.

Insert Table 3 about here

The Wright Map in Figure 2 shows the distribution of item difficulty for these 15 items. The items in bold font were likely candidates for reduction based on criterion of item difficulty. After eliminating these items, the remaining set is comprised of a good distribution of items in terms of difficulty.

Insert Figure 2 about here

In the third dimension, Developing a Positive School Learning Climate, the bold items (i.e., items 34, 28, 26, 46, 35, 27, 47, 50, 42, 41, 39, 43, 32, 36 in Table 4) appeared to be candidates for elimination. We also eliminated item 30 after reviewing the distribution of item content coverage. After this process of item reduction, a total of 10 items remained in this dimension of the prototype

PIMRS Teacher Short Form. This set of 10 items demonstrated high item-test correlations and adequately covered all five of the leadership functions within this dimension.

Insert Table 4 about here

Figure 3 displays the Wright Map for the third dimension. Based on the MNSQ index, the two items with the highest difficulty (items No. 34 and 35) demonstrate a marginally good fit (see Figure 3). Therefore, taking these multiple criteria into consideration, we decided to eliminate these two items in order to maintain better psychometric integrity for this dimension.

Insert Figure 3 about here

These steps yielded a prototype of the PIMRS Teacher Short Form that consisted of 22 items measuring the three dimensions of the PIMRS framework. This included five items in Defining the School's Mission, seven items in Managing the Instructional Program, and 10 items in Developing a Positive School Learning Climate. While the data analysis procedures described above indicated that the items 'fit' with the three conceptual dimensions of the scale, it remained to verify that the instrument continued to meet desirable standards of reliability and comparability.

Reliability Results

Next we ran the Gen Theory test of internal consistency to determine if the prototype instrument was reliable for the total scale (i.e., 22 items) as well as the three dimension-level constructs. The reliability results were .943 for the whole scale, .935 for the first dimension, .901 for the second dimension, and .912 for the third dimension. These findings confirm that the PIMRS Teacher Short Form meets or exceeds the reliability standards for instrument used in research, needs assessment, and personnel evaluation.

Validity Results

Tests of the validity of the prototype PIMRS Teacher Short Form were to internal validity. We employed CFA to assess the fit between the sample data and the conceptual framework at the dimension level. All of the factor loadings were higher than 0.7. The fit indices were as follows: goodness of fit index = 0.965, root mean square error of approximation = 0.088. Together these indicate a good fit between the data and the conceptual framework. The pattern of results further suggest that the dimensions measure related but different conceptual constructs subsumed under instructional leadership.

In the PIMRS Teacher Standard Form, the correlation among the three dimensions were .91 (between dimensions 1 and 2), .83 (between dimensions 1 and 3), and .91 (between dimensions 2 and 3). In the short form, they were .90, .81, and .89, respectively. These very small differences in correlation between the two forms indicate that the factor structure had remained stable after the elimination of items.

Comparability of Results

The correlation of raw scores between the Standard Form and Short Form were .94 for the first dimension, .97 for the second dimension, .97 for the third dimension, and .99 for the whole PIMRS scale. This very high correlation between the two forms is evidence of high comparability. We then used Equation 2 described in the Methods section to check the difference in the Rasch measures between the Standard Form and the Short Form. We found that, on average, scores obtained from 3.7%, 2.1% and 3.6% of teachers had a statistically significant difference at the .05 nominal level on the three dimensions, respectively. These percentages fell within an expected range (i.e., within 5%). In short, both analyses supported a high level of comparability between forms.

Because the test lengths of the PIMRS Teacher Standard Form and Short Form are very different, raw scores obtained from the two forms cannot not be compared directly. To resolve this problem, we employed a set of test equating techniques (Kolen & Brennan, 2004) in order to create a conversion table for raw scores from the two test. Using this table, raw scores from the two forms can be equated in cases where users wished to compare scores obtained from the two different forms. For example, for the first dimension, a raw score of 10 in the short form was approximately equivalent to a raw score of 18 in the standard form (not tabled). Details of the procedures used in developing the conversion table as well as the table itself are available in the PIMRS Technical Report (Hallinger et al., 2013b).

Discussion

Reviews of research conducted over the past half-century have consistently reported a strong preference among scholars studying educational leadership and management scholars for employing surveys as the method of choice (Bridges, 1982; Campbell& Faber, 1961; Haller, 1979; Hallinger & Heck, 1996). Given their prevalence, the importance of high quality instrumentation cannot be overstated as a foundation for high quality programmatic research (Bridges, 1982; Hallinger & Heck, 1996; Leithwood et al., 2006). Moreover, in the accountability-oriented, global context of 21st century schooling (Leithwood, 2001), school systems that undertake more systematic approaches to principal evaluation require instruments that meet high measurement standards (Goldring, Cravens, Murphy& Elliot, 2009).

This study reported the results of a research and development project aimed at creating a shortened version of the Teacher Form of the *Principals Instructional Management Rating Scale* (Hallinger, 1983). Although the PIMRS has a long track record of use in empirical research on leadership for learning (Hallinger et al., 2013a; Hallinger & Heck, 1996; Leithwood & Jantzi, 2005; Robinson et al., 2008), a recent review of PIMRS studies (Hallinger, 2011) found that many researchers have chosen to rely solely upon the Principal Form for data collection despite the stronger validity of the PIMRS Teacher Form. Since many principals do not wish to burden teachers with long surveys, the authors wished to see if it was possible to create a Short Form of the PIMRS that was capable of yielding comparable data at a similar level of data quality.

This R & D project reduced the PIMRS Teacher Standard Form from 50 to 22 items. The resulting PIMRS Teacher Short Form is capable of producing a full-scale score as well as scores for three dimensions of instructional leadership: Defining a School Mission, Managing the Instructional Program, Developing a Positive School Learning Climate. Using a Gen Theory test, our results confirmed that the PIMRS Teacher Short Form yields data that meet high standards of reliability (i.e., above 0.90) for the three dimensions as well as the full scale. Rasch analysis and factor analysis (i.e., CFA) further confirmed that the Short Form continued to maintain high levels of internal validity. Finally, we found that raw scores from these two forms of the PIMRS were highly correlated for the three dimensions(0.94. 0.97, 0.97) and the full scale (0.99)). We noted that forms of the same instrument comprised of different numbers of items never yield exactly the same scores. Therefore, we also developed a conversion table that equates raw scores of the PIMRS Teacher Short Form to raw scores of the Standard Form on the full scale and three dimensions.[iv]

These tests confirm that the PIMRS Teacher Short Form meets our previously stated criteria of reliability, validity and comparability. The resulting instrument reduces the time needed for teachers to complete the scale by more than half, to about 10 minutes. Future researchers who use the PIMRS in combination with other instruments can be confident that the PIMRS Teacher Short Form is a more efficient yet equally effective instrument for data collection when compared with the Standard Form.

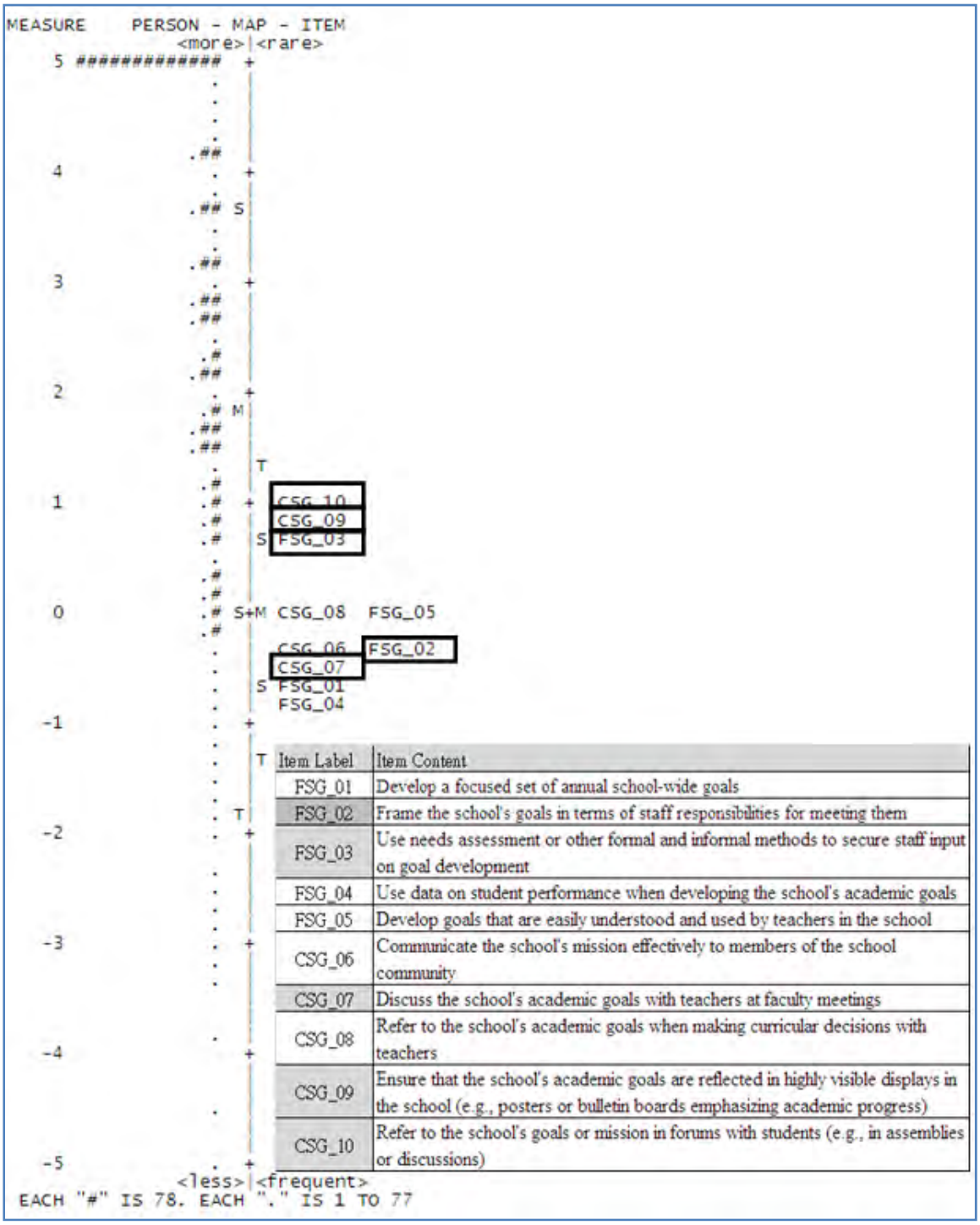
At the same time, we wish to take note of several limitations of this study. First, as discussed earlier, this R & D effort was limited to development of a shortened version of the PIMRS Teacher Form. The PIMRS Principal and Supervisor Forms are only available in the Standard 50 item version of the instrument. Measurements that rely on single raters typically require a larger item pool in order to achieve a high level of reliability (Gay, 1992). Based upon this principle, as well as the results of our earlier reliability study (Hallinger & Wang, 2013a), we have no intention to undertake development of a short form of the PIMRS Principal or Supervisor instruments.

A second limitation flows from our decision not to maintain the capacity of the PIMRS Teacher Short Form to yield data on the 10 instructional leadership functions that are measured in the Standard Form. Users for whom detailed information on the 10 leadership functions is deemed critical would, therefore, still wish to use the Standard Form. For example, in situations where detailed teacher feedback to principals is used for developmental purposes, users may wish to continue using the Standard Form.–

Finally, this instrument revision study did not extend to the measurement of external validity of the PIMRS Teacher Short Form. Our tests of validity were restricted to features of the scale's internal validity (e.g., content and construct validity). Further establishing the external validity of the Short Form remains a target for future research.

With these limitations in mind, we consider this effort to refine the capabilities of a widely-used research tool successful. Development of the PIMRS Teacher Short Form contributes another useful tool to the battery of tests used in the study and evaluation of school leadership. We hope

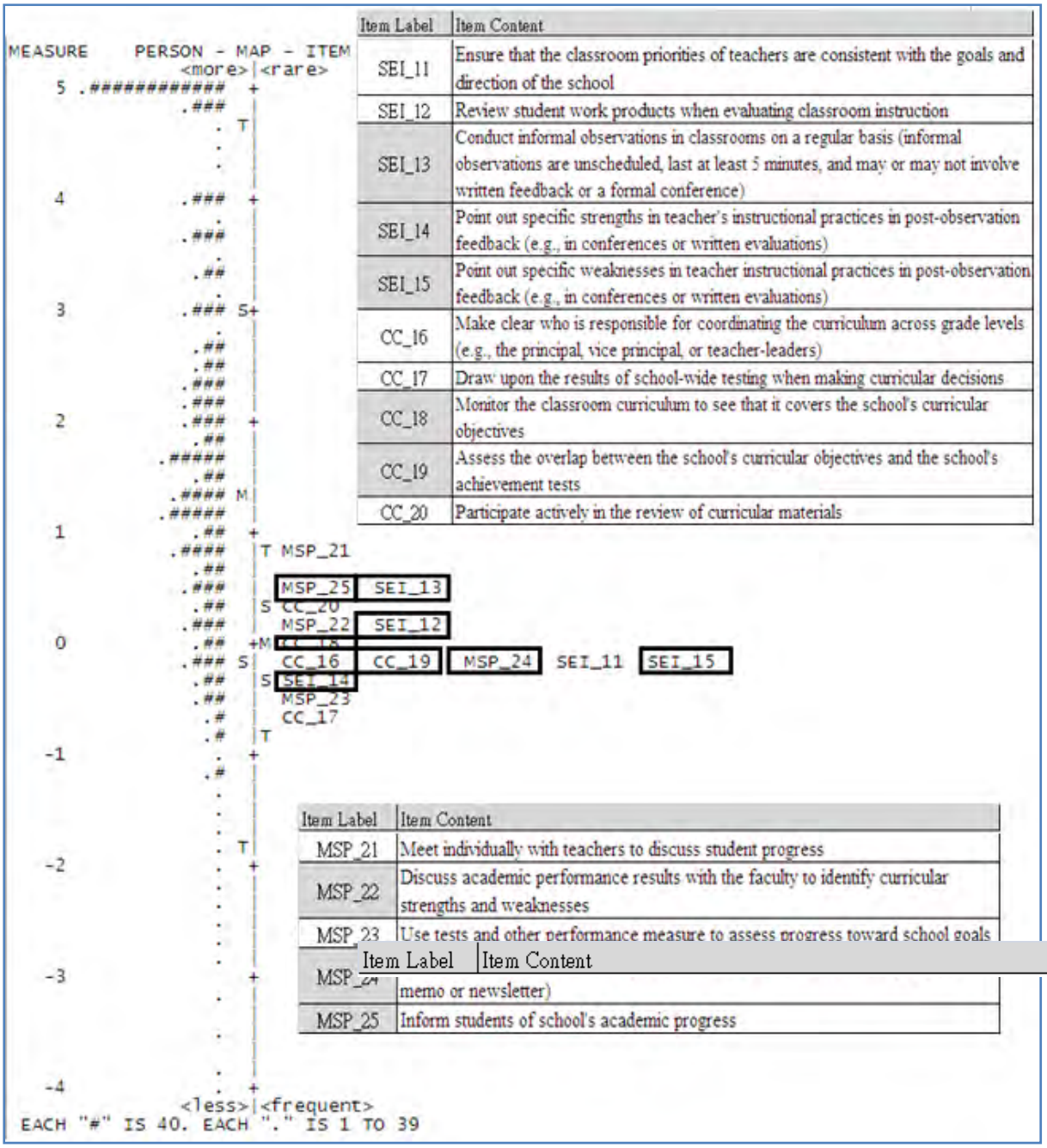
that the result of this R&D project will enable both researchers and practitioners to use the PIMRS more flexibly and efficiently in the future.



Item Label	Item Content
FSG_01	Develop a focused set of annual school-wide goals
FSG_02	Frame the school's goals in terms of staff responsibilities for meeting them
FSG_03	Use needs assessment or other formal and informal methods to secure staff input on goal development
FSG_04	Use data on student performance when developing the school's academic goals
FSG_05	Develop goals that are easily understood and used by teachers in the school
CSG_06	Communicate the school's mission effectively to members of the school community
CSG_07	Discuss the school's academic goals with teachers at faculty meetings
CSG_08	Refer to the school's academic goals when making curricular decisions with teachers
CSG_09	Ensure that the school's academic goals are reflected in highly visible displays in the school (e.g., posters or bulletin boards emphasizing academic progress)
CSG_10	Refer to the school's goals or mission in forums with students (e.g., in assemblies or discussions)

Figure 1. Item Person Map for Defining the School's Mission

Note: Item Label includes the name of Function and item number. FSG=Frame the School's Goals, CSG=Communicate the School's Goals

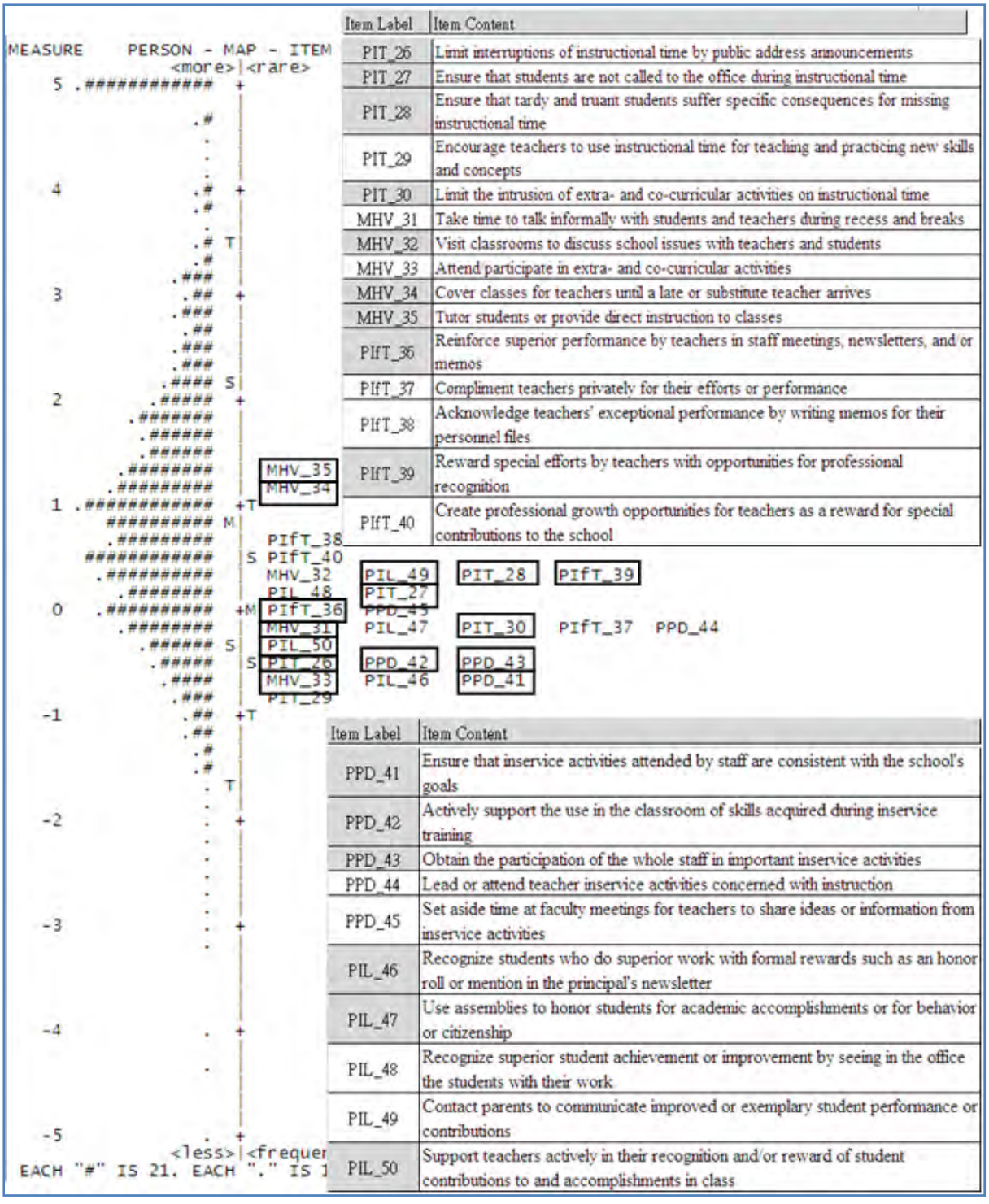


Item Label	Item Content
SEI_11	Ensure that the classroom priorities of teachers are consistent with the goals and direction of the school
SEI_12	Review student work products when evaluating classroom instruction
SEI_13	Conduct informal observations in classrooms on a regular basis (informal observations are unscheduled, last at least 5 minutes, and may or may not involve written feedback or a formal conference)
SEI_14	Point out specific strengths in teacher's instructional practices in post-observation feedback (e.g., in conferences or written evaluations)
SEI_15	Point out specific weaknesses in teacher instructional practices in post-observation feedback (e.g., in conferences or written evaluations)
CC_16	Make clear who is responsible for coordinating the curriculum across grade levels (e.g., the principal, vice principal, or teacher-leaders)
CC_17	Draw upon the results of school-wide testing when making curricular decisions
CC_18	Monitor the classroom curriculum to see that it covers the school's curricular objectives
CC_19	Assess the overlap between the school's curricular objectives and the school's achievement tests
CC_20	Participate actively in the review of curricular materials

Item Label	Item Content
MSP_21	Meet individually with teachers to discuss student progress
MSP_22	Discuss academic performance results with the faculty to identify curricular strengths and weaknesses
MSP_23	Use tests and other performance measure to assess progress toward school goals
MSP_24	memo or newsletter)
MSP_25	Inform students of school's academic progress

Figure 2. Item Person Map for "Managing the Instructional Program"

Note: Item Label includes the name of Function and item number: SEI= Supervise and Evaluate Instruction, CC= Coordinate the Curriculum, MSP= Monitor Student Progress



Item Label	Item Content
PIT_26	Limit interruptions of instructional time by public address announcements
PIT_27	Ensure that students are not called to the office during instructional time
PIT_28	Ensure that tardy and truant students suffer specific consequences for missing instructional time
PIT_29	Encourage teachers to use instructional time for teaching and practicing new skills and concepts
PIT_30	Limit the intrusion of extra- and co-curricular activities on instructional time
MHV_31	Take time to talk informally with students and teachers during recess and breaks
MHV_32	Visit classrooms to discuss school issues with teachers and students
MHV_33	Attend/participate in extra- and co-curricular activities
MHV_34	Cover classes for teachers until a late or substitute teacher arrives
MHV_35	Tutor students or provide direct instruction to classes
PIfT_36	Reinforce superior performance by teachers in staff meetings, newsletters, and/or memos
PIfT_37	Compliment teachers privately for their efforts or performance
PIfT_38	Acknowledge teachers' exceptional performance by writing memos for their personnel files
PIfT_39	Reward special efforts by teachers with opportunities for professional recognition
PIfT_40	Create professional growth opportunities for teachers as a reward for special contributions to the school

Item Label	Item Content
PPD_41	Ensure that inservice activities attended by staff are consistent with the school's goals
PPD_42	Actively support the use in the classroom of skills acquired during inservice training
PPD_43	Obtain the participation of the whole staff in important inservice activities
PPD_44	Lead or attend teacher inservice activities concerned with instruction
PPD_45	Set aside time at faculty meetings for teachers to share ideas or information from inservice activities
PIL_46	Recognize students who do superior work with formal rewards such as an honor roll or mention in the principal's newsletter
PIL_47	Use assemblies to honor students for academic accomplishments or for behavior or citizenship
PIL_48	Recognize superior student achievement or improvement by seeing in the office the students with their work
PIL_49	Contact parents to communicate improved or exemplary student performance or contributions
PIL_50	Support teachers actively in their recognition and/or reward of student contributions to and accomplishments in class

Figure 3. Item Person Map for Promoting a Positive School Learning Climate

Note: Item Label includes the name of Function and item number. PIT= Protect Instructional Time, MHV= Maintain High Visibility, PIfT=Provides Incentives for Teachers, PPD= Promote Professional Development, PIL= Provide Incentives for Learning

Chapter 7

Approaches to Assessing Principal Practice

We begin by noting that it is possible to gather data on the principals' instructional leadership practice through a variety of means including direct observation, interviews, and questionnaires (Goldring, Porter, Murphy, Elliot, & Cravens, 2009; Hallinger & Murphy, 1985a, 1987). Each method has advantages and disadvantages in terms of efficiency (i.e., time and effort) and effectiveness (i.e., quality of information).

Systematic use of direct observation was pioneered in principal professional development in the United States as long ago as the 1980s. Researchers at the Far West Lab in San Francisco developed the Peer-Assisted Leadership program or PAL (Barnett & Long, 1986). This professional leader learning program grew out of the research process employed in early observational studies instructional leadership carried out at the Far West Lab by Bossert and his colleagues (Dwyer, 1986; Dwyer et al., 1983). In this program principals use a combination of semi-structured observation and reflective interviewing to gather information and provide non-judgmental feedback to peers. This process was geared towards stimulating the principal to reflect on personal patterns of instructional leadership practice and link these to important goals, as well as to create a more intentional awareness among principals of their approach to instructional leadership.

The trend towards using peer coaching and mentoring in leader learning programs has gathered force in recent years. Thus, in many parts of the world today, peer coaching and mentoring have been introduced into training and development programs for school leaders (e.g., Browne-Ferrigno, & Muth, 2004; Ehrich, Hansford, & Tennent, 2004; Goldring, 2010; Walker, Chan, Cheung, Chan, Wong, & Dimmock, 2002). These programs train principals in formal coaching skills that they can then use with peers as means of stimulating on-the-job learning and development.

While observation offers possibly the most direct means of obtaining data on leadership practice, it is time-consuming. Multiple observations are needed to generate valid results. Observational data are not easily synthesized to provide a picture of performance across individuals. For these reasons, we view direct observation as a useful, but supplementary method of generating data on principal practice. In addition, we suggest observation of the principal may be more suitably employed for the purpose of professional development than for personnel evaluation.

Interviews with stakeholders can be employed to generate a picture of the principal's instructional leadership. Again, however, interviews are time consuming and of limited validity when used as the sole method of assessment. Respondents may be reluctant to make direct statements during interviews concerning the practices of their superordinate. Concerns over confidentiality may also inhibit the validity of responses.

Questionnaires represent a commonly used means of generating perceptual data. They are efficient, since it generally takes less time to complete and score a questionnaire than to conduct a single observation. Although questionnaires rely on the staff perceptions rather than observed behavior, numerous studies have found that they can provide reliable, valid data on managerial behavior (Latham & Wexlev 1981).

Issues arise concerning who should complete questionnaires that provide data on principal practice (e.g., the principal, teachers, supervisors, parents). This is important in that the assumption behind a behavioral questionnaire is that it is provides a 'perceptual sampling' of the principals' behavior. Therefore, respondents must have had sufficient opportunities to 'observe' the principal in practice if they are to provide valid data in response to questionnaire items (Hallinger & Murphy, 1985a, 1987; Latham & Wexley, 1981).

According to Duke and Stiggins (1985), the type of data needed in assessing leadership practice varies with the purpose of the assessment. Where assessments are used for personnel evaluation and other accountability-oriented purposes, the data must meet specific legal and professional standards of reliability and validity (Latham & Wexlev 1981). Few principal evaluation systems even approach such standards, and the procedures used seldom meet the criteria administrators must apply to the evaluation of teachers (Goldring, Cravens, Murphy, & Elliot, 2009). Where assessments are used only for professional improvement, there can be greater leeway in the nature of the data used (Duke & Stiggins, 1985). Within the scope of this chapter, we focus primarily on employing the *PIMRS* for the latter purpose.

The *PIMRS* can be administered to a principal as a self-assessment instrument as well as to supervisors and teachers to provide a broader picture of the principal's leadership. The choice of appraisers depends on the purposes of the assessment. When professional improvement is the sole concern, self-report data from the principal is acceptable, at least as a starting point. However, greater care must be exercised when collecting data as part of the evaluation process.

Using the PIMRS in Principal Evaluation

Principal self-assessment using the PIMRS provides useful comparative results, but taken alone, may not provide a valid picture of principal instructional leadership. Some principals tend to overestimate their role behavior, while others underestimate the degree of leadership exercised in this domain. We have found, in particular, that new principals and some highly effective principals tend to under-rate their performance.

Therefore, when used as part of a principal evaluation system, it is essential that the PIMRS be administered to the teaching staff of the principal's school. The reason for this is that only the teacher scores have demonstrated validity and reliability (see attached paper, Assessing the Instructional Management Behavior of Principals). You may choose to have the entire teaching staff of a school complete the PIMRS, or just a random sample in large schools. Note that only teachers who have worked with the administrator for a full year should be asked to complete the PIMRS.

Our experience with the PIMRS also suggests that district office supervisors should approach their assessment of the principal's instructional leadership with caution. A superintendent or assistant superintendent can complete the PIMRS arid share the results with the principal, but the supervisor must acknowledge the limited validity of the results. That is, the supervisor's perceptions are but one set of perceptions of the principal's leadership. The validity of the PIMRS is based upon the assumption that the respondent has observed the principal's leadership behavior in a reasonably large sample of situations. Without a sufficient sample of observations, the respondent may provide invalid results. This is analogous to teacher evaluation systems in which the principal assesses a teacher's performance on the basis of an insufficient number of classroom observations.

Research suggests that few central office administrators spend the requisite time in schools to able to make valid assessments of the specific behaviors measured by the PIMRS.

Thus, we advise principals to obtain the perceptions of multiple audiences in order to produce an accurate instructional leadership behavior profile. In addition, it is suggested that archival information such as school goal statements, faculty meeting agendas, principal newsletters, and teacher evaluation reports be used to comp1et~ I the picture of principal leadership behavior and to check the accuracy of perceptions collected through use of the PIMRS.

The Use of the PIMRS in Professional Development

When used as part of a professional development program, the principal choose to use the PIMRS with greater leeway in terms of the role groups surveyed with the instrument. Although from a measurement standpoint the same reliability and validity concerns exist, from the perspective of improving practice these technical issues are less important. It is the principal alone who is making use of the results and the most useful results for professional development purposes are those that the principal is ready to believe and act upon. Therefore, the principal may choose to complete a self-assessment for use in goal-setting, without additional feedback from/teachers or supervisors. Alternatively, the principal may choose to have the faculty and/or a supervisor provide feedback as well. These data can be contrasted with his/her self-perception along the 10 instructional leadership functions.

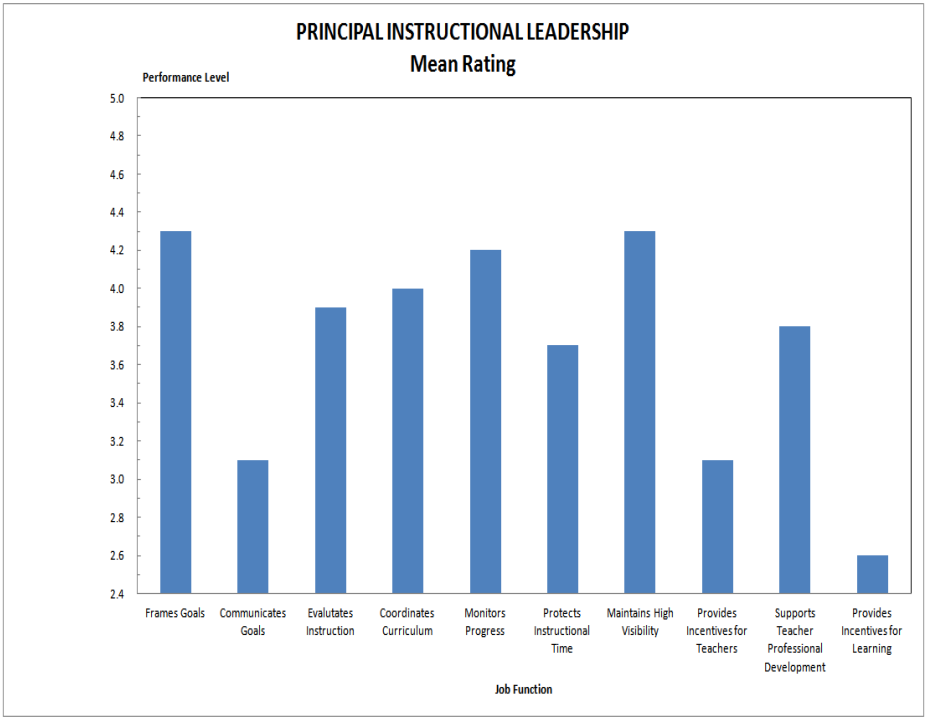


Figure 7.2. Rating of an Individual Principal with the PIMRS

A sample self-report profile of an individual principal on the 10 instructional leadership functions was shown above in Figure 7.2. There are several features to keep in mind when interpreting this particular profile. First, the analysis focuses on functional responsibilities rather than the broader dimensions (e.g., Creating a Mission) or specific items. From the perspectives of both

measurement and practical experience, this often offers a useful level of detail for users, whether they are principals or system-level personnel. Second, we note that the scale is based on a level of frequency of demonstrating the specified behaviors. While in moderately high ratings woe preferred, a rating of 5 may not reflect the optimal quality of performance. Therefore, the fact that a principal has not achieved a profile of 'full marks' should not be interpreted as a 'personal deficit'.[v] Thus, a principal would typically approach the interpretation of this profile in terms of an identification of 'relative strengths' in the 10 functional areas of instructional leadership.

In order to maximize the validity of the PIMRS assessment results, we do suggest that both teachers and the principal complete the scale. This not only yields more valid information, but also additional interpretive perspectives for principals to consider. A sample profile comparing the perceptions of teachers and a principal is shown in Figure 7.8.

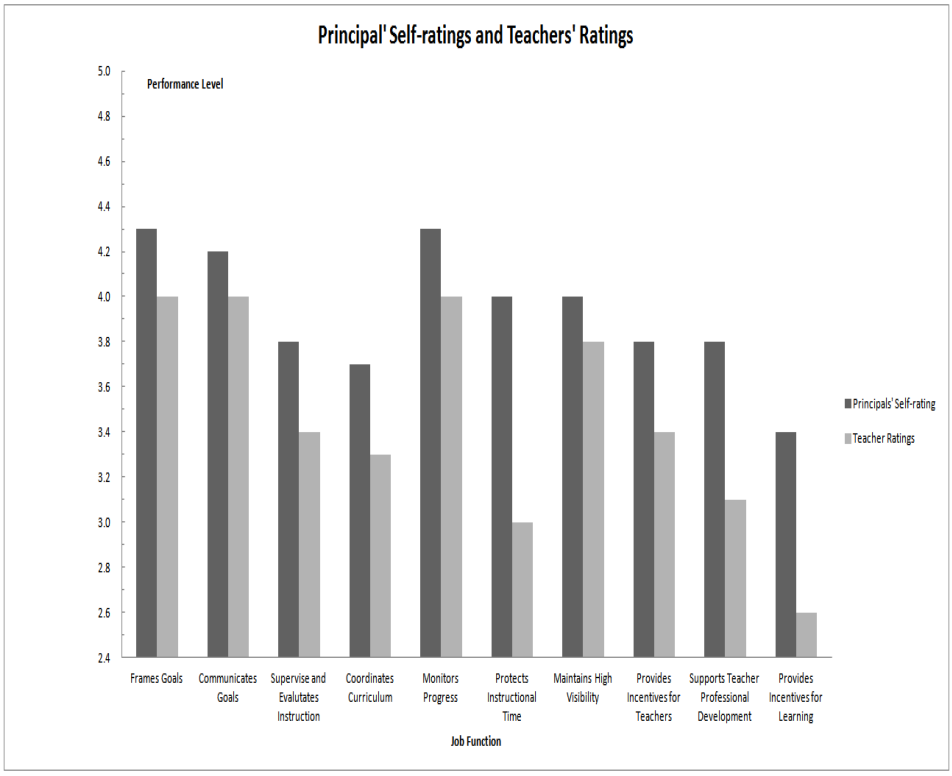


Figure 7.8. Comparing Teacher and Principal Perceptions on Instructional Leadership Functions

This comparative profile of the principal's instructional leadership offers an opportunity for the principal to compare his/her self-perceptions with those of the teaching staff and/or a supervisor(s). Thus, in Figure 7.3, the principal would not only identify areas relative strength with respect to the 10 instructional leadership functions, but also differences between self-perceptions and those of one's colleagues. If we accept coaching as a process of data-gathering, feedback and self-reflection (Barnett & Long, 1986; Goldring, 2010), then these differences in perception can be employed as a stimulus for reflection, goal-setting, action steps, and further data gathering. Notably, the use of the profile also offers a data-driven approach to assessing change over time in the principal's practice in specific areas within this key domain of the role.

System administrators can also aggregate data obtained across a number of schools in order to plan identify system-wide strengths and weaknesses. This information can be used for the purposes of planning staff development for principals, recruitment and selection of new principals and middle level leaders, succession planning, and revision of system policies. For example, Figure 8.4 offers a different data displays for 10 principals from the same school system, this time focusing analysis on the three instructional leadership dimensions.

Interpretation of Figure 7.4 would focus on mean performance as well as variability of performance across the principals on the three dimensions. For example, we can see that the principals as group as a whole appear considerably stronger in terms of establishing a clear mission than in managing the instructional program or developing the school learning climate. This is reflected in a stronger overall mean performance (not tabled) as well as the fact that this dimension was strongest for every principal, regardless of their personal mean rating.

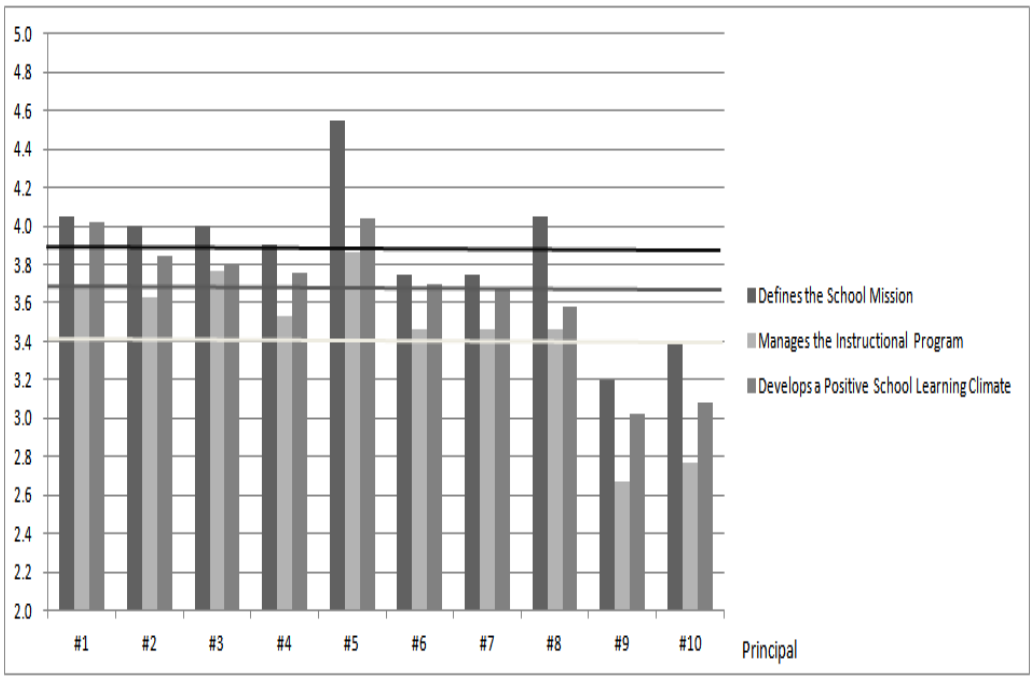


Figure 7.4 Comparison of Instructional Leadership Profiles Among 10 Principals

The profile also highlights differences between principals. Thus we can see that #5 appears to stand out as an instructional leader, while principals #9 and 10 appear to be relatively weaker. Based on this profile, one cold subsequently drill down to examine these performance trends in terms of the 10 leadership functions, as well as individual performance profiles. These illustrative profiles are offered in order to indicate the direction that school systems have taken in employing these data, both to stimulate individual principal learning and development as well as for planning system-wide training and policy revisions.

Chapter Eight

The Role of the PIMRS in Future Research on Leadership and Learning

This chapter draws upon findings from a comprehensive review of PIMRS. We highlight some of the key conclusions that emerged from the review concerning predominant research topics and methods that have engaged the attention of scholars over the past several decades. Then we highlight implications of these findings. The chapter aims to provide a succinct and pointed summary of recommendations for future research that employs the PIMRS.

Introduction

In 2011, Hallinger (2011a) undertook a systematic review of research conducted with the PIMRS over the preceding 30 years. Hallinger's review examined 135 PIMRS studies conducted between 1983 and 2011. This review was undertaken with a primary purpose of contributing to our understanding of methodological progress in studying principal instructional leadership. The review found that even as fads and fashions in leadership have waxed and waned, scholarly interest in instructional leadership has remained surprisingly consistent and strong. Indeed, the data reviewed in the review found that instructional leadership has become firmly entrenched in the firmament of professional practice and gained currency as a focal construct in the eyes of scholars. One can conclude that instructional leadership is even more relevant in 2010 than thirty years ago.

Moreover, evidence offered in this book indicates that PIMRS instrument provides reliable and valid data on instructional leadership. Thus, we conclude that the PIMRS can play a *potentially* useful role in empirical research on instructional leadership. We highlight the word 'potentially' because a survey instrument is only a research tool. The successful application of this tool is linked to the nature of the research question that is posed by the researcher in concert with the use of suitable conceptual models, research designs, and methods. In this chapter we summarize findings from the above-mentioned review of research and outline directions for future research with respect to topics, conceptual and methods.

Summary of Findings

Concerning the use of conceptual models to provide perspective studies of leadership, we noted that the frequency of use of direct effects models (see Figure 2.1) was consistently strong throughout this 27 year period. Unfortunately, researchers employing direct effects have tended to do so in a largely a-theoretical fashion. For example, with few exceptions (e.g., Cunningham, 2004), direct effects studies that examined the relationship between principal gender and instructional leadership did so without testing a theoretical explanation for why gender differences would impact leadership. Therefore, after three decades and 30+ PIMRS studies of principal gender, we know that female principals consistently receive higher ratings on the PIMRS than males, but are no closer to a theoretically grounded explanation for why this is the case.

Second, although we found increased used of mediated-effects frameworks during the final period of the review, the relatively small total number of mediated effects studies does not qualify this as evidence of major progress. Because leadership is enacted in complex organizational settings, past reviewers have recommended that scholars employ conceptual models that are capable of portraying these multivariate relationships. Antecedent/direct effects studies (Model A-1 in Figure 2.1) highlight factors that influence the exercise of leadership. However, researchers employing this model have too often failed to link those relationships to the impact of leadership. Model B studies highlight leadership effects on school conditions, but often without employing sufficiently comprehensive perspectives. Thus, we join other reviewers (e.g., Bossert et al., 1982; Bridges, 1982; Haller, 1979; Hallinger & Heck, 1996a; Murphy et al., 1983) in recommending that researchers employ frameworks that contextualize the enactment of leadership. Studies that incorporate personal antecedent variables should explore their effects not only on instructional leadership, but also on school-level conditions and/or school outcomes.

Demonstrable improvements were noted in the pattern of use of statistical methods within the 135 PIMRS studies. Indeed, they suggested a clear pattern of improvement over time within the three-decade period of the current review. This conclusion holds when the data were compared against findings from prior eras as well as over time within the three decades covered in this review. At the same time, however, this mostly entailed a reduced reliance on the use of descriptive statistics and an increased use of bivariate tests without controls. Improvements did not extend to a more frequent use of statistical methods able to shed light on multivariate relationships concerned with school context, personal characteristics of leaders, instructional leadership, and student learning.

Citation analysis suggested that this body of studies has not yet yielded an impact on knowledge accumulation, even among those who are toiling in precisely the same field of inquiry. This finding is startling, both for the magnitude (or lack thereof) of the result and the clarity of interpretation. That said, we should note that these results closely mirror those reported by Bridges (1982), even though his study included both dissertations and published studies.[vi] Indeed, Bridges' conclusion continues to ring true almost 30 years later: '[T]hese results point to a lack of interconnectedness among the research studies on school administrators. This lack of systematic knowledge-building seems to be a pervasive characteristic of research in the field of educational administration" (Bridges, 1982, p. 24).

Directions for Future Research with the PIMRS

As discussed in earlier chapters, the PIMRS was developed in response to an expressed need for research-informed instruments capable of contributing to a program of research on learner-centered leadership in education (Bossert et al., 1982; Bridges, 1982; Hallinger & Murphy, 1985). Our analysis of subsequent research conducted with the PIMRS yields implications for researchers who choose to employ this instrument in the future. We close this chapter and the book by highlight some issues that appear most relevant to improving the quality of future research.

Topical Foci for Future Research

With respect to desirable foci for research, we suggest that scholars who use the PIMRS more squarely accept the challenge of investigating the linkages between instructional leadership and

school-level variables that mediate effects on teacher effectiveness, and student learning. Numerous scholars have noted the need to shed light on the "black box" which contains the processes through which leadership contributes to the improvement capacity of schools to create a positive impact on student learning (Hallinger & Heck, 1996a; Heck & Hallinger, 2005; Leithwood et al., in press). We noted increased interest in studying this issue over the last decade, and wish to encourage it further using comprehensive conceptual models, multivariate statistics, as well as through mixed method studies.

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

As suggested above, we believe that impact should be studied in terms of student learning. However, worthy research may also target other intermediate and distal variables such as teacher collective efficacy, satisfaction and commitment, school health, organizational learning, teacher change, and student engagement. Useful models for conducting empirical research on the relationship between school leadership and these variables using comprehensive models and robust statistical methods exist in the literature (e.g., see Hallinger, Bickman & Davis, 1996; Heck & Hallinger, 2009; Heck, Larson & Marcoulides, 1990; Leitner, 1994; Leithwood & Jantzi, 2000; Marks & Printy, 2003). Doctoral students are encouraged to draw upon these models and methods, rather than simply citing findings from these studies.

Conceptual and Methodological Implications

Our discussion of conceptual models incorporates two levels of conceptual analysis. First, we briefly review how leadership itself is conceptualized. This discussion highlights the conceptualization of instructional leadership in relation to other leadership models in studies of leadership and learning. The second level of conceptual analysis considers how leadership is framed in relation to other variables. For example, the Bossert et al. (1982) model illustrated in Chapter Two framed instructional leadership as influencing student learning through other mediating variables. Here we examine issues that bear upon the researcher's explicit definition of an over-arching conceptual model for the study.

Conceptualizing School Leadership

As noted in Chapter One, during the 1990s a paradigm war pitted instructional leadership against transformational leadership. Leithwood's (1994; Leithwood & Jantzi, 1999, 2000, 2005) conceptualization of transformational school leadership was framed a potentially more powerful model of describing how school leaders achieve effects on organizational outcomes. This led to a perception that instructional leadership offered a less powerful explanation of leadership and learning. Over time, however, this 'debate' has largely been resolved through programmatic empirical studies conducted by a variety of researchers (Day et al., 2010; Hallinger & Heck, 2010, 2012; Heck & Hallinger, 2009, 2011; Marks & Printy, 2003; Robinson et al., 2008; Witziers et al., 2003).

The most 'direct' testing of the two models was conducted by Robinson and colleagues (2008) in a meta-analytic study that compared the effects of instructional and transformational leadership on student learning. After quantitatively integrating the findings of numerous studies, the authors concluded that instructional leadership offered a more potent explanation of the means by which leaders impact learning in schools. Although the methodology of this highly cited study has attracted criticisms (e.g., Scheerens, 2012), the findings from this review are broadly accepted.

Other analytic reviews of the related literature by Hallinger (2003) and Leithwood and colleagues (2006) have emphasized the overlapping contributions of these theoretical models. Thus, we observe that recent conceptual models of leadership and learning often include aspects of both instructional and transformational leadership (e.g., see Day et al., 2010; Goldring, Porter, Murphy, Elliot & Cravens, 2009; Hallinger & Heck, 2011; Leithwood et al., 2010, 2012).

Another conceptual issue that deserves mention concerns the source of instructional leadership. Over the past decade there has been a concerted global press towards distributing leadership in schools among a broader set of key stakeholders, especially teachers (Barth, 1990; Gronn, 2002; Lambert, 2002, 2003; Spillane, 2006). Paradoxically, the latest thinking suggests that the drive to develop distributed leadership in schools neither diminishes nor comes at the expense of the principal's responsibilities for leadership. Indeed, scholars and policymakers alike assert that principal leadership remains a key driver for change and source of support for building leadership capacity among others (e.g., Childs-Bowen, Moller, & Scrivner, 2000; Gewirtz, 2003; Lambert, 2002, 2003; Murphy, 2009; Stricherz, 2001). As Mayrowetz and colleagues observe: "[P]rincipals occupy the critical space in the teacher leadership equation and center stage in the work redesign required to bring distributed leadership to life in schools."

Nonetheless, conceptualizing instructional leadership as a 'shared responsibility' has implications for measurement. As noted in this volume, the PIMRS has been used primarily to assess instructional leadership performed by the principal. Yet in contexts where these responsibilities are structurally distributed, the results of the PIMRS may only capture a partial picture of instructional leadership in practice. To date the PIMRS has not been adapted for assessing shared instructional leadership. Although it is possible to address this by changing the questions in the scale from a focus on the principal to a focus on collective effort (e.g., see Heck & Hallinger, 2009), the instrument has yet to be formally validated for this usage.

Modeling the Relationship Between Leadership and Learning

Another relevant conceptual perspective concerns how the researcher conceptualizes te relationship between leadership and learning. In a series reviews conducted during the mid-1990s, Hallinger and Heck (1996a, 1996b,1998) asserted that the use of bi-variate or direct effects models in the study of school leadership effects represented a 'dry well'. They further elaborated on why this was the case, identifying theoretical, design and methodological reasons for this result. Consequently, they argued for a discontinuation of such studies on the grounds that they represented a misallocation of effort with little likelihood of contributing to knowledge in the future. Other scholars have reprised this argument in various publications over the succeeding years (e.g., Leithwood et al., 2006, 2010, 2012; Robinson et al., 2008; Witziers et al., 2003).

Hallinger's (2011a) review conducted 15 years hence found that researchers have generally heeded this admonition. The review identified a trends towards the use of mediated effects studies of leadership and learning. As noted earlier in this book, a recent study by Hallinger and Heck (2011) explicitly tested direct, indirect and reciprocal effects models. We concluded unequivocally that the latter two conceptual models yielded more robust results of this relationship.

At the same time, however, we would be remiss if we did not note one dissenting line of scholarly inquiry. In recent years, selected scholars (e.g., Nettles & Herrington, 2007; Silva et al., 2011) who have continued to explore the efficacy of the direct effects model. They argue that the high stakes context for school leadership in the USA *demands* that school leaders take direct actions to impact student learning. They suggest that creating the mission, managing instruction and shaping the school climate are insufficient in schools that face government takeovers due to poor performance in student learning outcomes.

With this rationale in mind, they have continued to employ a direct effects conceptualization of principal leadership (e.g., Silva et al., 2011). As one example of this approach, they studied the implementation of direct effects model of instructional leadership using an experimental design. The intervention that was studied consisted of the principal coaching student at risk of failing to increase their focus and motivation. The findings from their study remain preliminary and unsubstantiated through additional research. Therefore, although we continue to believe that mediated effects models have replaced the direct effects model on theoretical and practical grounds, exploration of the direct effects model does continue.

In their 1996 review, Hallinger and Heck concluded that the challenge of elaborating and testing the 'paths' through which leadership impacts learning represented the most compelling challenge in this domain of research. Despite the increased efforts of scholars to address this challenge, in our opinion this continues to describe the key direction of research in this field. Indeed, other scholars have also made this argument (e.g., Leithwood et al., 2006, 2010, 2012; Marks & Printy, 2003; Mulford & Silins, 2010; Robinson et al., 2008 ; Witziers et al., 2003). We wish to draw readers' attention, in particular, to Leithwood's recent attempts to define and test these paths through which leadership works to influence student learning (see Leithwood et al., 2010, 2012).

We wish to note an additional implication of our discussion concerning direct effects and mediated effects models. Along with other reviewers (e.g., Bossert et al., 1982; Bridges, 1982; Leithwood et al., 2006, 2010; Witziers et al., 2003), we have been unremittingly critical of using direct effects conceptualizations in the study of school leadership and learning. We must, however, note that this critique applies to studies in which the distal dependent variable (i.e., outcome) is student learning. Our rationale lies fundamentally in the observation that with the exception of a one-room schoolhouse, principals do not teach students. Teachers do. Therefore, instructional leadership must be geared simultaneously to create an environment in the school that motivates teachers, develops teacher capacity, and motivates students to achieve. While principals may take some 'direct actions' that impact students, at the end of the day the impact of these will necessarily be limited when compared to the indirect actions (Hallinger & Heck, 1996a, 1998; Kleine-Kracht, 1993; Leithwood et al., 2006, 2010, 2012; Robinson et al., 2008; Witziers et al., 2003)

What do the above comments imply in terms of using 'direct effects conceptual models' in future research on leadership and learning? First, as indicated, we assert that the most robust research

findings will accrue from studies that apply multivariate mediated for reciprocal effects models (e.g., models B, B-1, C in Figure 2.3). However, scholars have noted since the 1970s note that most research in our field is conducted by doctoral students (Bridges, 1982; Haller, 1979; Hallinger, 2011a). Doctoral students must fashion research studies that take into account various limitations that impact their choice of research conceptualizations and designs. For example, although we recommend longitudinal designs as a preferred means of assessing the causal effects of leadership, the time duration requirements would typically exceed the scope of a doctoral study. Similarly, studying the effects of leadership on student achievement requires access to data that meet quite requirements (see Hallinger, 2011a; Hallinger & Heck, 2012) that go beyond what most doctoral students could collect within the context of their dissertation research. Compromises on the quality of student achievement data often mean that students spend significant amounts of time collecting and analyzing massive amounts of data that cannot be used to reliably predict relationships. The result is that numerous studies are conducted on leadership and learning without contributing significantly to knowledge (Bridges, 1982; Campbell, 1979; Haller, 1979; Hallinger, 2011a; Heck & Hallinger, 2005; Leithwood, 2006a; Murphy, Vriesenga & Storey, 2007).

We suggest two potential solutions to this design dilemma that face doctoral researchers in particular. The first is, where possible, to conduct doctoral studies in the context of larger funded projects conducted by faculty members in the department or an affiliated research center. This solution ensures that students will have access to data that enables them to apply more sophisticated conceptual models and statistical methods. We recognize that this may run counter to norms in some institutions that require students to collect their own data in doctoral dissertations. However, we strongly suggest that the quality test of a doctoral study should be the potential to contribute to knowledge. It should not be based on the procedural steps in conducting a study that cannot contribute to knowledge because of weak data.

An alternative solution is to conduct studies that examine the relationship between leadership and variables that are directly related within models of leadership and learning (e.g., see Figure 2.1) In this manner scholars can make more limited by useful contributions. Such studies could focus, for example, on the relationship between antecedents of instructional leadership (e.g., gender, experience, expertise) and instructional leadership. Or they could examine the impact of leadership on variables that mediate its effects on learning (e.g., instructional organization, staff capacity, school learning climate, teacher commitment, teacher satisfaction). The data requirements of these types of studies are less likely to exceed the reach of doctoral students and who are working without the benefit of funding and/or research infrastructure support.

Appendix A: List of PIMRS Studies

Appendix A: List of PIMRS Studies

1. Abdullah, A.G.K., & Wahab, M.A. (2007). *The impact of principal's instructional leadership behavior on PPSMI teachers' teaching practices.* Paper presented at the 5th ASEAN Symposium on Educational Management and Leadership (ASEMAL 5), Kuala Lumpur, Malaysia.
2. Abdullah, J.B., & Kassim, J. (2011). Instructional leadership and attitude towards organizational change among secondary schools principal in Pahang, Malaysia. *Procedia - Social and Behavioral Sciences, 15,* 3304-3309.
3. Abdullah, J.B., & Kassim, J. (2012). Promoting learning environment among the Islamic school principals in the state of Pahang, Malaysia. *Multicultural Education & Technology Journal, 6*(2),100-105.
4. Adams, D.S. (2002). *A multiple-site case study of instructional leadership in low-performing, high-improvement high schools*. Unpublished doctoral dissertation, Azusa Pacific University, Azusa, CA.
5. Adam, P. (2012). *The effect of principal instructional leadership characteristics on the academic growth of lower socio-economic students*. Unpublished doctoral dissertation, University of Kansas, Lawrence, KS.
6. Adkins, d. (1990). *The Relationship between visionary leadership and instructional leadership behavior of secondary school principals: Regression analysis and hermeneutic interpretation.* Unpublished doctoral dissertation, West Virginia University. Eric Document No. 9121851.
7. Ali, Y. (2012). *Effectiveness of principal instructional leadership in preparatory schools in South Wollo Zone*. Unpublished doctoral dissertation, Addis Ababa University, Addis Ababa, Ethiopia.
8. Anderson, J. (2006). *An analysis of the relationship of high school principals perceived instructional leadership management behaviors to school size and to student achievement.* Unpublished doctoral dissertation, Texas A&M University, Commerce, TX.
9. Anderson, Y. (2010). *Developing a portrait of an elementary school principal desiring to produce an effective school.* Unpublished doctoral dissertation, Bowie State University, Bowie, MD.
10. Apolonia, C. (1998). *Relationship of the instructional leadership behavior of administrators and morale as perceived by teachers in Philippine Cultural High School, Caloocan Campus, School Year 1997-1998.* Unpublished master thesis, Ateneo de Maila U., Manila, Philippines.
11. Aste, M. (2009). *The relationship between principals instructional focus and academic achievement of high poverty students.* Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA.

12. Atkinson, R. (2013). *An assessment of the perceived instructional leadership behaviors of assistant principals. Unpublished doctoral dissertation,* Virginia Commonwealth University, Richmond, VA.

13. Augustine, L. (1998). *Principal learners: An examination of a self-designed, foundation-funded learning experience for principals as a model of professional development.* Unpublished doctoral dissertation, Columbia University, Teachers College, NY. Eric Document No. 9909406.

14. Aviles, I.J.T. (2009). *Retos, estrateggias y liderazgo instructional de directores novels en escuelas primarias*. Unpublished doctoral dissertation, Universidad Autónoma de Yucatán , Mérida de Yucatán, Mexico.

15. Babcock, C. (1991). *Perceived instructional leadership behavior: Gender differences, female principals, male principals*. Unpublished doctoral dissertation, Western Michigan University, Royal Oak, MI.

16. Balsamo, M. A. (2004). *Assessing principal practices in a standards-based environment and examining the association between principal practices and student achievement.* Unpublished doctoral dissertation, University of Rochester, Rochester, NY.

17. Benoit, J. (1990). *Relationships between principal and teacher perceptions of principal instructional management and student achievement in selected Texas school districts with an emphasis on an index of effectiveness*. Unpublished doctoral dissertation, New Mexico State University, Las Cruces, NM.

18. Bravo, B. (1991). *A Study of the perceived characteristics for effective leadership of monolingual and bilingual supervisors of special education programs.* Unpublished doctoral dissertation, Columbia University, Teachers College, New York, NY.

19. Brooks, B. (2011). *Teachers' perceptions of instructional management in elementary schools*. Unpublished doctoral dissertation, Capella University, Newberry SC.

20. Brown, G.T.L., & Chai, C. (2012). Assessing instructional leadership: A longitudinal study of new principals. *Journal of Educational Administration, 50*(6), 753 – 772.

21. Brown, J. (1991). *Leader behavior and school effectiveness*. Unpublished doctoral dissertation, University of Texas, Austin, TX.

22. Brown, R. (1991). *Critical attributes of instructional Leadership in nationally recognized secondary-level blue ribbon schools.* Unpublished doctoral dissertation, East Texas State University, Commerce, TX. Eric Document No. 9202160.

23. Burwell, G. (1988). *Elementary Teachers' Perceptions of Effective Principal Behaviors.* Unpublished doctoral dissertation, University of Akron, Akron OH.

24. Buzek, E. S. (2004). *The relationship between instructional leadership behaviors of middle school principals in Texas and student achievement.* Unpublished doctoral dissertation, University of Houston, Houston, TX.

25. Calvert, K. (2013). *Administrator leadership and content knowledge: Effects on literacy achievement on male students in grades four through eight.* Unpublished doctoral dissertation, University of Arkansas, Little Rock, AR.

26. Campbell, D.C. (1999). *The principal's role in changing school culture and implementing Title I reform.* Unpublished doctoral dissertation, South Carolina State University, Orangeburg, SC.

27. Cantu, M. (1994). *A study of principal instructional leadership behaviors manifested in successful and nonsuccessful urban elementary schools.* Unpublished doctoral dissertation, University of Texas, Austin. ERIC Document No. 9428472.

28. Carr, T. (2011). *An examination of leadership styles in implementing instructional technology: A case study to examine the elementary school principal perspective.* Unpublished doctoral dissertation, North Central University, Minneapolis, MN.

29. Carson, C. (2012). *The relationship between principal instructional leadership and student achievement in Missouri middle schools*. Unpublished paper. Saint Louis University, St. Louis, MO.

30. Chan, Y.C. (1992). *A study of principals' instructional leadership in Hong Kong secondary schools*. Unpublished master thesis, Chinese University of Hong Kong, Hong Kong.

31. Chan, Y.C., & Cheng, Y.C. (1993). A study of principals' instructional leadership in Hong Kong secondary schools. *Educational Research Journal. 8*, 56-67.

32. Chappelear, T.C. (2011). *The relationship of teachers' perceptions of high school principal's monitoring student progress and student achievement in southeastern Ohio.* Unpublished doctoral dissertation, West Virginia University, Morgantown, WV.

33. Chen, M.(2010). *An instructional leadership model for senior high schools: the student perspective.* Unpublished doctoral dissertation, East China Normal University, Shanghai, China.

34. Cheatham, T. (2010). *The relationship between principals' leadership behaviors and student academic performance in four classifications of schools.* Unpublished doctoral dissertation, Mercer University, Macon GA.

35. Chi, L.C. (1997). *Principal and teacher perceptions of principal instructional management behavior in Taiwan, Republic of China*. Unpublished doctoral dissertation, University of South Dakota, Bismarck, SD.

36. Clabo, B.T. (2010). *The high school principal as instructional leader: An explanatory, mixed methods case study examining principal leadership within the context of rural secondary schools*. Unpublished doctoral dissertation, University of Tennessee, Knoxville, TN.

37. Clark, I. (2009). *An analysis of the relationship between K-5 elementary school teachers' perceptions of principal instructional leadership and their science teaching efficacy*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis-St. Paul, MN.

38. Collins, F. C. (1998). *An investigation of perceived differences in instructional leadership and school climate of African American and Caucasian female principals in Ohio's urban public schools.* Unpublished doctoral dissertation, Bowling Green State University, Bowling Green, OH.

39. Coltharp, G. (1989). *The relationship between leadership behaviors and instructional leader effectiveness of building level administrators*. Unpublished doctoral dissertation, Kansas State University, Manhattan, KS.

40. Corkill, R. (1994). *Instructional leadership behaviors perceived as important by teachers and principals for academic achievement in selected South Texas Elementary Schools*. Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

41. Courtney, M. (1987). *The relationship between the perceived degree of instructional leadership and principal management style and teacher stress*. Unpublished doctoral dissertation, Teacher's College, Columbia University, New York, NY.

42. Cunningham, J. (2004). *The effect of gender-role identity of female principals on in Texas on teacher perceptions of instructional leadership behaviors*. Unpublished doctoral dissertation, Texas A&M University, Corpus Christi, TX.

43. Dale, A. (2010). *A causal study examining how instructional leadership, transformational leadership, and the mediating effects of teacher self-efficacy influence the math achievement of third through fifth grade students as measured by the Maryland school assessment*. Unpublished doctoral dissertation, University of Maryland, Eastern Shore, MD.

44. Dale, A., & Phillips, R. (2011). *Influences of instructional leadership, transformational leadership and the mediating effects of self-efficacy on students achievement.* Paper presented at the 6th International Conference of the American Institute of Higher Education, Charleston, SC.

45. Daud, K., & Basiron, M.N. (2011) *Persepsi guru-guru terhadap kepimpinan pengajaran guru penolong kanan pentadbiran semasa guru besar mengikuti Program Khas Pensiswazahan Guru Besar (PKPGB) Di sebuah sekolah kebangsaan di daerah Johor Bahru.* unspecified . 1-9.

46. Dawson, D. A. (2010). *Leadership practices that enhance reading achievement for African American males: A case study*. University of South Carolina). *ProQuest Dissertations and Theses,* , 186. Retrieved from http://search.proquest.com/docview/824025494?accountid=11441. (824025494).

47. Delano, K. (1985*). An analysis of ratings by teachers of elementary school principals with different classroom teaching experiences.* Unpublished doctoral dissertation, Auburn University, Montgomery, AL. Eric Document No. 8604734.

48. Dennis, C. (2009). *The relationship between principals' self-perceptions and teachers' perceptions of high school principals' instructional leadership behaviors in South Carolina schools*. Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

49. Dickerson, D. (1999). *Teachers' perceptions of the role of the principal as instructional leader in the context of school reform in Northwest Georgia secondary schools.* Unpublished doctoral dissertation, University of Alabama.

50. Diego, J. (2013). *Evaluation of a program designed to enhance the instructional leadership competencies of principals.* Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

51. 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 Southern Mississippi, Hattiesburg, MS.

52. Dunn, A. (2010). *A quantitative study of the perceptions of school leadership preparation in the State of Georgia*. Unpublished doctoral dissertation, Walden University, Minneapolis, MN.

53. DuPont, J. P. (2009). *Teacher perceptions of the influence of principal instructional leadership on school culture: A case study of the American embassy school in New Delhi, India.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis-St. Paul, MN.

54. Duryea, M. (1988). *The relationship between perceived principal instructional leadership behavior and locus of control.* Unpublished doctoral dissertation, University of Houston, Houston, TX. ERIC Document No. 8908543.

55. Edwards, C. A. (2007). *An analysis of the relationship of superintendent instructional leadership behaviors and district performance outcomes.* Unpublished doctoral dissertation, Tarleton State University, Stephenville, TX

56. Fancera, S. F. (2009). *Instructional leadership influence on collective efficacy and school achievement.* Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

57. Fancera, S., & Bliss, J. (2012). Instructional leadership influence on collective teacher efficacy to improve school achievement. *Leadership & Policy in Schools, 10*(3), 349-370.

58. Freshour, D. (1990*). Emphasis on and training for instructional leadership among secondary administrators in Northern Rocky Mountain States.* Unpublished doctoral dissertation, University of Montana. Eric Document No. 9029107.

59. Fulton, T. (2009). *High school principal instructional leadership behavior in high and low need and high and low achievement schools*. Unpublished doctoral dissertation, Dowling College, Oakdale, NY.

60. Gallon, S. (1998). *A study of teachers' perceptions of the principal as an instructional leader: A comparative analysis of elementary, middle and senior high school principals*. Unpublished doctoral dissertation, Florida International University, Miami, FL.

61. Garcia. L. (1999). *Principal instructional leadership in majority-minority secondary schools.* Unpublished doctoral dissertation, University of Texas, Austin, TX.

62. Gaziel, H., Cohen-Azaria, Y., & Ifanti, A. The antecedents of primary school principals' management and leadership behaviors. *Journal of Educational and Research, 2*(2), 143-153.

63. Geiselman, S. (2004). *Predicting elementary school student achievement: The impact of principal gender and principal leadership skills*. Unpublished doctoral dissertation, University of Louisville, Louisville, KY.

64. Gerrell, N. S. (2005). *The relationship between principals' instructional leadership skills and the academic achievement of high-poverty students.* Union University, Jackson, TN.

65. Gibson, C. A. (2005). *A comparison of self-reported instructional leadership management behaviors of elementary principals of high-performing and low-performing schools in Texas.* Unpublished doctoral dissertation, University of Houston, Houston, TX.

66. Gjelaj Merturi, E. (2010). *The perceptions of principal-based leadership practices on student reading achievement.* Unpublished doctoral dissertation, University of Phoenix, Phoenix, AZ.

67. Goldring, E.B., Cravens, X., Murphy, J., & Porter, A. (2012). *The convergent and divergent validity of the Vanderbilt Assessment of Leadership in Education™ (VAL-ED): Instructional leadership and emotional intelligence*. Paper presented at the annual meeting of the Association for Education Finance and Policy in Boston, MA.

68. Grande, M. (2012). *Comparison of principals' leadership practices by methods of professional development.* Unpublished doctoral dissertation, Walden University, Minneapolis, MN.

69. Greb, W. (2011). *Principal leadership and student achievement: What is the effect of transformational leadership in conjunction with instructional leadership on student achievement?* Unpublished doctoral dissertation, Marian University, Indianapolis, IN.

70. Grier, D. (1988*). Instructional Leadership Behaviors and Their Impact on Selected TELS Exemplary Elementary Schools.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

71. Griffin, M. (1993). *Instructional leadership behaviors of Catholic secondary school principals.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT. Eric Document No. 9327348.

72. Groff, M. L. (2002). *Elementary instructional leadership: Factors that affect performance.* Unpublished doctoral dissertation, Saint Louis University, St. Louis, MO.

73. Haack, M. (1991). *An investigation of the instructional management behaviors of principals in mid-sized public high schools in Iowa*. Unpublished doctoral dissertation, University of Northern Iowa, Cedar Falls, IO.

74. Haasl, W. (1989). *Multiple perspectives of role definition of small school principals and their influence on effective instructional management*. Unpublished doctoral dissertation, University of Wisconsin, Madison, WI.

75. Haggard, R. L. (2008). *The instructional leadership practices of principals in K--8 schools and their impact on student learning outcomes.* Unpublished doctoral dissertation, California State University, Fresno, CA.

76. Hallinger, P. (1983). *Assessing the instructional management behavior of principals.* Unpublished doctoral dissertation, Stanford University, Stanford, CA. Eric Document No. 8320806.

77. Hallinger, P. (2011). A review of three decades of doctoral studies using the *Principal Instructional Management Rating Scale*: A lens on methodological progress in educational leadership. *Educational Administration Quarterly, 47*(2) 271-306.

78. Hallinger, P. & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal, 86*(2), 217-248.

79. Hallinger, P., Taraseina, P., & Miller, J. (1994). Assessing the instructional leadership of secondary school principals in Thailand. *School Effectiveness and School Improvement, 5*(4), 321-348.

80. Hallinger, P., & Wang, W.C., & Chen, W.C. (In press). Assessing the measurement properties of the *Principal Instructional Management Rating Scale*. A meta-analysis of reliability studies. *Educational Administration Quarterly.*

81. Hao, T.N., & Wu, B. (2011). Much ado about many things: Principle functions analysis and evaluation of primary principals' instructional leadership in Vietnam. *International Journal of Innovative Management, Information & Production, 3*(2), 61-73.

82. Harris, E. (2002). *The relationship between principals' instructional leadership skills and the academic achievement of high-poverty students.* Unpublished doctoral dissertation, University of South Carolina.

83. Hart, J. (2006). *A cross-cultural study of principals' leadership behaviors according to teachers' perceptions*. Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA.

84. Henderson, N. (2007). *Teacher and principal perceptions of effective instructional leadership: An exploration of guiding practice and personal beliefs*. Unpublished doctoral dissertation, University of Texas, San Antonio, TX.

85. Holyfield, B. (2010). *Teaches' perceptions of the instructional leadership role of the principal in implementing effective instructional practices for the Georgia high school graduation test*. Unpublished doctoral dissertation, Capella University, Newberry, SC.

86. Horton, T, (2013). *The relationship between teachers' sense of efficacy and perceptions of principal instructional leadership behaviors in high poverty schools.* Unpublished doctoral dissertation, University of Texas-Arlington, Arlington, TX.

87. Holyfield, B. (2010*). Teacher perceptions of the instructional leadership roleof the rincipal in implementing effective instructional leadership practices toprepare students for the Georgia high school graduation test*. Unpublished doctoral dissertation, Capella University, Newberry, SC.

88. Howard-Schwind, M. (2010). *Instructional leadership responsibilities of assistant principals in large Texas high schools.* Unpublished doctoral dissertation, University of North Texas, Denton, TX.

89. Howe, W. (1995). *Instructional leadership in Catholic elementary schools: An analysis of personal, organizational, and environmental correlates*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

90. Howell, V. (1989). *Analyses by Gender and by School Enrollment, Ratings of Elementary School Principals' Instructional Leadership Behaviors.* Unpublished doctoral dissertation, Tennessee State University, Nashville, TN.

91. Hung, D.K.M., & Ponnusamy, P. (2010). Instructional leadership and schools effectiveness. *Knowledge Management, Information Systems, E-Learning, and Sustainability Research Communications in Computer and Information Science, 111*, 401-406.

92. Hunter, C. (1994). *Los Angeles Unified School District middle school principals' instructional leadership behaviors and academic achievement.* Unpublished doctoral dissertation, Peppardine University, Los Angeles, CA. ERIC Document No. 9511556.

93. Jaeger, M. G. N. (2001). *Principal selection decisions made by teachers on school councils: The influence of person characteristics, job candidate experience, and teacher job assignment.* Unpublished doctoral dissertation, University of Louisville, Louisville, MO.

94. Jennings, T.C. (2013). *School principals' personality types and instructional leadership practices in selected distinguished and non-distinguished Title One elementary Schools in South Carolina.* Unpublished doctoral dissertation, South Carolina State University, Orangeburg, S.C.

95. Johnson, A. B. (2005). *The relationship between leadership behaviors of elementary principals from various school contexts and student achievement.* Unpublished doctoral dissertation, Argosy University, Schaumburg, IL.

96. Johnson, D. M. (2006). *Principal instructional leadership and academic performance in Tennessee high schools.* Unpublished doctoral dissertation, University of Tennessee, Knoxville, TN.

97. Jones, P. (1987). *The relationship between principal behavior and student achievement in Canadian secondary schools*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

98. Kassim, J.M., & Abdullah, J.B. (2011). Promoting learning environment and attitude towards change among secondary school principals in Pahang Malaysia: Teachers' perceptions. *Procedia - Social and Behavioral Sciences*, *28*, 45–49

99. Keith, S. (1989). *Teacher efficacy and the relationship between elementary principals instructional leadership and self-perception of efficacy.* Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA. ERIC Document No. 9002836.

100. Kelly, C. (1986). *The importance of an administrator's organizational perspective in developing styles of effective instructional leadership.* Unpublished master thesis, Simon Fraser University, Vancouver, Canada.

101. Kincaid, D. (2006). *Selected factors affecting instructional leadership skills of public, private, and charter elementary school principals in Texas.* Texas A&M University. Unpublished doctoral dissertation, Commerce, TX.

102. Kennedy, E. (1993). *Principal leadership functions and teacher effectiveness.* Unpublished doctoral dissertation, Fordham University, New York, NY. Eric Document No. 9328415.

103. Knezek, E. J. (2001). *Supervision as a selected instructional leadership behavior of elementary principals and student achievement in reading.* Unpublished doctoral dissertation, University of Texas, Unpublished doctoral dissertation, Austin, TX.

104. Kroeze, D. (1992). *District context and its impact on the instructional leadership role of the elementary school principal.* Unpublished doctoral dissertation, University of Chicago, Chicago, IL. Eric Document No. T-31945.

105. Krug, F. (1986). *The relationship between the instructional management behavior of elementary school principals and student achievement.* Unpublished doctoral dissertation, University of San Francisco, San Francisco, CA.

106. Lee, M.S., Walker, A., & Chui, Y.K. (2012). Contrasting effects of instructional leadership practices on student learning in a high accountability context. *Journal of Educational Administration, 50*(5), 586-611.

107. Lehl, A. (1989). *A comparison study of the instructional knowledge and the instructional leadership of elementary principals in Nebraska.* Unpublished doctoral dissertation, University of Nebraska, Lincoln NE. Eric Document No. 9004690.

108. Leitner, D. (1990). *The effects of principal instructional management behavior on student learning: An organizational perspective*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

109. Leitner, D. (1994). Do principals affect student outcomes: An organizational perspective. *School Effectiveness and School Improvement,* 5(3), 219-238.

110. Long, C. L. (2008). *Instructional leadership: Perceptions of Mississippi career and technical education administrators and teachers.* Unpublished doctoral dissertation, Mississippi State University, Jackson, MS.

111. Lord, C. T. (2001). *Instructional leadership teams and school climate: A descriptive study of leadership behavior and indicators of climate in secondary schools.* University of Connecticut. Unpublished doctoral dissertation, Storrs, CT.

112. Lubbers, M. (1988). *An investigation to determine if a relationship exists between teacher efficacy, principal behaviors and student achievement*. Unpublished doctoral dissertation, Michigan State University, East Lansing, Michigan.

113. Lyons, B. J. (2010). *Principal instructional leadership behavior, as perceived by teachers and principals, at New York State recognized and non-recognized middle schools.* Unpublished doctoral dissertation, Seton Hall University, South Orange, NJ.

114. Maciel, R. G. (2005). *Do principals make a difference? an analysis of leadership behaviors of elementary principals in effective schools.* Unpublished doctoral dissertation, University of Texas - Pan American, Edinburg, TX.

115. MacNeil, A. (1992). *Principal instructional management and its relation to teacher job satisfaction*. Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

116. Mallory, B. (2003). *The relationship of teachers perceptions of instructional leadership and student achievement measured by state testing in ten North Carolina high schools*. Unpublished doctoral dissertation, East Carolina State University, Greenville, NC.

117. Mann, K. (1988*). Conditions affecting the instructional leadership behavior of elementary school principals in a rural county school district.* Unpublished doctoral dissertation, University of Maryland, College Park, MD. Eric Document No. 8912376.

118. Marshall, A. K. (. (2005). *Instructional leadership: Perceptions of middle school principals and teachers.* University of Texas - Pan American, Texas A&M University, Commerce, TX.

119. Maslyk, J. (2012*). A qualitative study of blue ribbon elementary school principals: Perspectives on promoting student achievement*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Indiana, PA.

120. McCabe, B. (1993). *Antecedents to principal instructional management behavior*. Unpublished doctoral dissertation, University of California, Santa Barbara, Santa Barbara, CA. Eric Document No. 9419092.

121. McCarthy, M. (2009). *Teachers' perceptions of high school principals' leadership behaviors using the Principal Instructional Management Rating Scale (PIMRS) and the relationship to the AYP (adequate yearly progress) status in high poverty suburban school districts located in Southeastern Pennsylvania.* Unpublished doctoral dissertation, Saint Joseph's University, Philadelphia, PA.

122. McCier, M. H. (2003). *A study of high school principals' instructional leadership behaviors as perceived by teachers in urban comprehensive and magnet high schools.* Wayne State University, Detroit. MI.

123. McDonald, J. (2012). *Instructional leadership and student achievement: The role of Catholic identity in supporting instructional leadership*. Unpublished doctoral dissertation, Drexel University, Philadelphia, PA.

124. Mcilavain, S. (1986). *Improving the instructional leadership of elementary principals through in-service training.* Unpublished doctoral dissertation, University of Kansas, Lawrence, KS. Eric Document No. 8619866.

125. Mctaggart, M. (1991). *Using the Principal Instructional Management Rating Scale to assess instructional leadership among school administrators.* Unpublished doctoral dissertation, University of South Dakota, Bismarck, SD. Eric Document No. 9137432.

126. Meek, J. (1999). *Relationship between principal instructional leadership and student achievement outcomes in North Carolina elementary schools.* Unpublished doctoral dissertation, North Carolina State University, Raleigh, N.C.

127. Mercer, S. J. (2004). *The relationship of teachers' perceptions of principals' instructional leadership skills and school performance in four high-poverty South Carolina middle schools.* Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

128. Meyer, T. (1990). *Perceived leader authenticity as an effective indicator of perceived instructional leadership behavior in middle level principals*. Unpublished doctoral dissertation, Andrews University, Berrien Springs, MI.

129. Michelo, J. (2013). *An evaluation study of the Diploma in Education Management offered at National In-Service Teachers' College (NISTCOL) – Chalimbana: A case study of selected schools in Lusaka Urban District of Zambia*. Unpublished doctoral dissertation, Zambian Open University, Lusaka, Zambia.

130. Miller, L. J. (1991). *Perceptions of Alabama middle school teachers and principals regarding the principal's instructional leadership role and competencies as developed by effective school studies. .* Unpublished doctoral dissertation, University of Alabama, Montgomery, AL

131. Minus, E. (2010*). Leading in the middle: Leadership behaviors of middle level principals that promote student achievement*. Unpublished doctoral dissertation, George Washington University, Washington D.C.

132. Moffit, J. (2007). *What works? Principal leadership behaviors that positively impact student achievement in elementary schools*. Unpublished doctoral dissertation, Georgia Southern University, Statesboro, GA.

133. Moore, Q. (2003). *Teachers' perceptions of principals' leadership skills in selected South Carolina secondary schools*. Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

134. Morrow, C. (2010). *An analysis of high school principals' technology use pertaining to instructional leadership impacting student achievement.* Unpublished doctoral dissertation, Tarleton State University, Stephenville, TX.

135. Munroe, M. D. (2009). *Correlation of emotional intelligence and instructional leadership behaviors.* Unpublished doctoral dissertation, University of Phoenix, Phoenix, AZ.

136. Neuenswander, W. (1986). *Achievement directed administration.* Unpublished doctoral dissertation, University of Kansas, Lawrence, KS. Eric Document No. 8619867.

137. Newton, R. M. (1998). *Teacher reactions to school council member job descriptions: The effects of council job attributes, principal council role, and teacher job assignment level.* University of Louisville. Unpublished doctoral dissertation, Louisville, MO.

138. Nogay, K. (1995). *The relationship of superordinate and subordinate gender to the perceptions of leadership behaviors of female secondary principals*. Unpublished doctoral dissertation, Youngstown State University, Youngstown, OH.

139. Nogay, K., & Beebe, R. (1995). Gender and perceptions: Females as secondary principals. *Journal of School Leadership, 18*(6), 583-602. ERIC Document EJ888569.

140. Noor, M., & Audryanah, A. (2007). Kepemimpinan pengajaran dan efikasi kendiri pengetua sekolah menengah dan hubungannya dengan pencapaian akademik sekolah. master thesis, Universiti Teknologi Malaysia, Faculty of Education.

141. Nyau, W.Y. (2010). *Instructional leadership in Malaysia.* Unpublished master thesis, Open University of Malaysia, Kuala Lumpur, Malaysia.

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

143. O'Donnell, R. (2002). *Middle-Level principals' instructional Leadership Behaviors and student achievement.* Unpublished doctoral dissertation, Lehigh University, Lehigh, PA.

144. O'Donnell, R., & White, G. (2005). Within the accountability era: Principals' instructional leadership behaviors and student achievement. *NASSP Bulletin, 89*, 645, 56-71.

145. Orange, J. (1990). *Instructional leadership of Kentucky school principals in elementary schools of various effectiveness.* Unpublished doctoral dissertation, University of Kentucky, Lexington, KY. Eric Document No. 9107123.

146. Otten, C. (1990). *A comparison of instructional leadership practices of the principals of 'AA' elementary schools in the state of Missouri.* Unpublished doctoral dissertation, Northeast Missouri State University. Eric Document No. 1341549.

147. Özdemir, M. (2012). The impact of principals' instructional leadership behaviors on teachers' attitudes toward work: Some evidence from Turkey. *International Journal of Education Administration and Policy Studies, 4*(6), 147-153.

148. Palmer, V. (2000). *Instructional leadership of principals in predominantly Hispanic schools.* Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA.

149. Parker, B. (1990*). The principal as an instructional leader in non-urban schools.* Unpublished doctoral dissertation, University of Missouri, Columbia, MO. Eric Document No. 9108370.

150. Pavan, B.N. & Reid, N. (1991). *Espoused theoretical frameworks and the leadership behaviors of principals in achieving urban elementary schools*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. ERIC Document ED337533.

151. Peariso, J. F. (2011). *A study of principals instructional leadership behaviors and beliefs of good pedagogical practice among effective California high schools serving socioeconomically disadvantaged and English learners.* Unpublished doctoral dissertation, Liberty University, Lynchburg, VA.

152. Ponnusamy, P., (2010). *The relationship of instructional leadership, teachers' organizational commitment and school effectiveness in Malaysia*. Unpublished master thesis, Universiti Sains Malaysia, Kuala Lumpur, Malaysia.

153. Poovatanikul, V. (1993). *Analyses of the perceptions of Thai principals and teachers as to the principal's role as instructional leader.* Unpublished doctoral dissertation, Southern Illinois University, Carbondale, IL.

154. Prater, M. E. (2004). *The relative impact of principal managerial, instructional, and transformational leadership on student achievement.* Unpublished doctoral dissertation, University of Missouri, Columbia, MO.

155. Pratley, D. (1992). *A study of middle level principals' involvement in instructional leadership.* Unpublished doctoral dissertation, Michigan State University, East Lansing, MI. Eric Document No. 9223238.

156. Ratchaneeladdajit, R. (1997). *Perceptions of Thai principals and teachers toward the principals' role as instructional leaders in private schools in Bangkok, Thailand.* Unpublished doctoral dissertation, Southern Illinois University, Carbondale, IL.

157. Reid, N. (1989). *Elementary school principals' behaviors in an urban school improvement program: a descriptive study of five principals*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

158. Rodrigues, L. (2012). *Instructional leadership of an elementary school principal in Portugal: A case study*. Unpublished doctoral dissertation, Universidade de Aveiro, Aveiro Portugal.

159. Rogers, J. (2005). *Improving student learning: Development of a resource guide for elementary school principals.* Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

160. Rose, M. (1991). *Perceptions of selected principals in South Carolina concerning their instructional management behaviors.* Unpublished doctoral dissertation, University of Southern Mississippi, Hattiesburg, MS. Eric Document No. 9206184.

161. Roudebush, D. (1996). *Teacher's perceptions of building principals as instructional leaders: A comparative study in year-round multi-track and conventional school settings.* Unpublished doctoral dissertation, University of Denver, Denver, CO.

162. Ruzicska, J. (1989). *The relationships among principals' sense of efficacy, instructional leadership, and student achievement.* Unpublished doctoral dissertation, University of San Francisco, San Francisco, CA. Eric Document No. 8913049.

163. Ryans, Y. (1989). *Perceptions about the elementary principal as an instructional leader.* Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA.

164. Saavedra, A. (1987). *Instructional management behaviors of secondary administrators.* Unpublished master thesis, Bukidnon State College, Malaybalay, Bukidnon, Philippines.

165. Salleh, M.J., & Hatta, M. (2011). *The relationship between' instructional leadership practices and students academic achievement of secondary schools in Banda Acheh, Indonesia*. Paper presented at the 2nd Regional Conference on Educational Leadership RCELAM, IAB - Institute of Aminuddin Baki, Kedah, Malaysia.

166. Salvador, A. (1999). *Relationship between the teachers' and administrators' perceived school organizational health and administrators' instructional leadership in Our Lady of Perpetual Succor College, High School Dept., 1998-99*. Unpublished master thesis, Ateneo de Manila University, Manila, Philippines.

167. San Nicolas, E.Q. (2003). *Instructional leadership in Guam's public elementary schools.* Unpublished doctoral dissertation, University of San Diego, San Diego, CA.

168. Sawyer, E. (1997). *The relationship between cognitive styles and instructional leadership. Unpublished doctoral dissertation*. Unpublished doctoral dissertation, University of Missouri, Columbia, MO.

169. Schoch, A. (1992). *The relationship between instructional leadership behavior, school effectiveness, school size, gender, race, and years of principalship experience in elementary schools in South Carolina.* Unpublished doctoral dissertation, University of South Carolina, Columbia, SC. Eric Document No. 9307987.

170. Shafeeu, I. (2011). *A study of principals' instructional leadership in schools of Gaafu Dhaalu Atoll, Maldives*. Unpublished master thesis, University of Malaya, Kuala Lumpur, Malaysia.

171. Shatzer, R. H. (2009). *A comparison study between instructional and transformational leadership theories: Effects on student achievement and teacher job satisfaction.* Unpublished doctoral dissertation, Brigham Young University, Provo, UT.

172. Sheppard, B. (1993). *A study of the relationship among instructional leadership behaviors of the school principal and selected school-level characteristics.* Unpublished doctoral dissertation, University of Ottawa, Ottawa, ON.

173. Sheppard, B. (1996). *Exploring the transformational nature of instructional leadership, The Alberta Journal of Educational Research, 42*(4), 325–344.

174. Sicina, E. (1996). *Perception of the principal as the instructional leader and teacher participation in decision-making.* Unpublished doctoral dissertation, Fordham University, New York, NY. ERIC Document No. 9708265.

175. Simpson, A. (1988). *Assessment of the instructional effectiveness of elementary principals.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, NE.

176. Singleton, R. (2006). *Elementary school principals' leadership practices: Differences among Mississippi schools.* Unpublished doctoral dissertation, Jackson State University, Jackson, MS.

177. Sinha, V.(2009). *Teachers' perceptions: Differences in principals' instructional leadership skills in high and lower performing high poverty South Carolina middle schools.* Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

178. Skiptunas, C. (1990). *Secondary principal leadership and school climate.* Unpublished doctoral dissertation, State University of New York at Buffalo, Buffalo, NY. Eric Document No. 9134629.

179. Smith, A. (2012). *Leadership that fosters a culture of high achievement of Afrian American students.* Unpublished doctoral dissertation, Georgia Southern University, Statesboro, GA.

180. Smith, S. (2007). *Principals' and teachers' perception of principals' instructional leadership*. Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

181. Staples, C. (2005). *A comparison of the leadership roles of public and private elementary school principals*. Unpublished doctoral dissertation, University of Central Florida Orlando, Florida

182. Stevens, L. (1996). *Instructional leadership: A single district study of the multiple perceptions of central office administrators, principals, and elementary teachers*. Unpublished doctoral dissertation, State University of New York at Buffalo, Buffalo, NY.

183. Stroud, D. (1989). *Factors influencing the instructional leadership behaviors of elementary school principals and the relationship to student achievement.* Unpublished doctoral dissertation, Peabody College for Teachers of Vanderbilt University, Nashville, TN. Eric Document No. 9017489.

184. Switzer, J. (2006). *Understanding administrative appraisal in international schools.* Unpublished doctoral dissertation, University of Bath, Bath, England.

185. Syarwan, A. (2012). *Instructional leadership practices of the principals of excellent schools in Aceh, Indonesia*. Unpublished doctoral dissertation, University of Malaysa, Kuala Lumpur, Malaysia.

186. Taff, B. (1997). *Teacher perceptions of principal role behavior and school effectiveness.* Unpublished doctoral dissertation, Auburn University, Montgomery, AL. Eric Document No. 9802477.

187. Tang, Tseng-Chung. (1997). *The relationships among transformational, transactional and instructional leadership behaviors of senior secondary school principals in Taiwan.* Unpublished doctoral dissertation, University of South Dakota, Bismarck, S.D.

188. Taraseina, P. (1993). *Assessing instructional leadership behavior of secondary school principals in Thailand*. Unpublished doctoral dissertation, Peabody College for Teachers of Vanderbilt University, Nashville, TN.

189. Todd, T. (2006). *Instructional leadership in high schools: The effects of principals, assistant principals and department heads on student achievement.* Unpublished doctoral dissertation, Florida Atlantic University, Boca Raton, FL.

190. Tomasetti, B. (2007). *Instructional leadership: Key perceptions in five central Pennsylvania elementary schools during times of legislated accountability.* Unpublished doctoral dissertation, Indiana University of Pennsylvania, Indiana, PA.

191. Trout, K. (1985). *The relationship between gender and selected characteristics associated with instructional leadership for senior high school principals.* Unpublished doctoral dissertation, University of South Carolina, Columbia, SC. Eric Document No. 8604286.

192. van de Grift, W., & Houteen, A. (1999). Educational leadership and pupil achievement in primary education. *School Effectiveness and School Improvement, 10*(4), 373-389.

193. van Pelt, J. (1993). *Instructional management and its relationship to the context of the school and characteristics of the principal.* Unpublished doctoral dissertation, Drake University, Des Moines, IO.

194. Vinson, T. (1997*). The performance of secondary principals as instructional leaders in Mississippi public schools as perceived by superintendents, principals, and classroom teachers*. Unpublished doctoral dissertation, University of Mississippi, Jackson, MS.

195. Wafir, N. (2011). *The role of the principal's leadership in contributing towards the professional development of teachers in schools.* Unpublished master thesis, University of Nottingham, Nottingham, UK.

196. Wang, S. (2011). *Leadership in Chinese junior secondary school: Discovering quantitative relationship between antecedent factors, instructional leadership, teacher perception and student academic achievement.* Unpublished master thesis, Newcastle University, Newcastle Upon Tyne, UK.

197. Wardlow, R. (2008). *Induction and support of new principals.* Unpublished doctoral dissertation, University of California, San Diego, San Diego, CA.

198. Waters, D. (2005). *Principal leadership behavior and 5th grade student achievement in high poverty schools*. Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA.

199. Watkins, P. (1992*). Instructional leadership in relation to classroom environment, student enrollment, removal, and completion within LDS released-time seminaries.* Unpublished doctoral dissertation, University of the Pacific, Stockton, CA. Eric Document No. 9303164.

200. Wells, G. (1993). *Instructional management behavior, time management, and selected background variables of elementary school principals in Connecticut's Urban school districts.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT. Eric Document No. 9333216.

201. Werner, M. (1991). *Relating California principals' instructional leadership behaviors to state distinguished recognition of their high- schools.* Unpublished doctoral dissertation, University of San Francisco, San Francisco, CA. Eric Document No. 9212686.

202. Williams, A. (1995*). An analysis of the instructional leadership behavior of suburban middle school principals regarding the level of alienation of eighth-grade male students.* Unpublished doctoral dissertation, Texas Southern University, Houston, TX. Eric Document No. 9721870.

203. Wilson, C. (2005). *Principal leadership, school climate and the distribution of leadership within the school community.* Unpublished doctoral dissertation, University of Montana, Missoula MT.

204. Winger, M. (1992). *Keeping on track: The instructional leadership of year-round elementary school principals.* Unpublished doctoral dissertation, University of California, Los Angeles, CA. Eric Document No. 9230816.

205. Wong, Y.N. (2010). *Instructional leadership in Malaysia*. Unpublished master thesis. Open University of Malaysia, Kuala Lumpur, Malaysia.

206. Wotany, J. (1999). *The training, selection, and hiring practices of Cameroonian school leaders and the evaluation of the present effectiveness of their instructional leadership behaviors*. Unpublished doctoral dissertation, U. of Missouri, Kansas City, MO.

207. Yang, C.S. (1996). *Instructional leadership behaviors of elementary school principals in Taiwan, Republic of China*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

208. Yates, P. H. (2000). *Instructional leadership behaviors of principals and the effects of a balanced beginning reading program in exemplary elementary schools in Northeast North Carolina.* Unpublished doctoral dissertation, Regent University, Virginia Beach, VA.

209. Yogere, L. (1996). *Instructional leadership behaviors of principals as perceived by themselves and the teachers of Assumption, San Lorenzo, during the school year 1995-1996*. Unpublished master thesis, Ateneo de Manila University, Manila, Philippines.

210. Zeanah, R. (1986). *A comparison of elementary school principals' instructional management roles in effective and ineffective schools.* Unpublished doctoral dissertation, Auburn University, Montgomery, AL.

Appendix B: Reference List

Adkinson, J.A. (1981). Women in school administration: A review of the research. *Review of Educational Research, 51*(3), 311-343.

Agusto, J. (2009). *Case studies of Ed.D. and Ph.D. dissertations in educational administration.* Unpublished doctoral dissertation, University of Kansas.

Andrews, R., & Soder, R. (1987). Principal instructional leadership and school achievement. *Educational Leadership, 44*, 9-11.

Andrich, D. A. (1978). A rating formulation for ordered response categories. *Psychometrika, 43,* 561–573.

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003) Context and leadership: An examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly, 14*(3), 261-295.

Archbald, D. (2008). Research versus problem solving for the educational doctoral thesis: Implications for form and function. *Educational Administration Quarterly*, *44*(5), 704-739.

Arikewuyo, M. O. (2007). Teachers' perception of principals' leadership capacities in Nigeria. *Academic Leadership Journal, 5*(3). Retrieved July 20th 2011 from http://www.academicleadership.org/pdf/ALJ_ISSN1533-7812_5_3_187.pdf

Atwater, L., Dionne, S., Avolio, B., Camobreco, J., & Lau, A. W. (1999). A longitudinal study of the leadership development process: Individual differences predicting leader effectiveness. *Human Relations, 52*, 1543-1562.

Bajunid, I. A. (1996). Preliminary explorations of indigenous perspectives of educational management: the evolving Malaysian experience, *Journal of Educational Administration, 34*(5), 50–73.

Bamburg, J. & Andrews, R. (1990). School goals, principals and achievement. *School Effectiveness and School Improvement, 2*(3), 175-191.

Barnett, B., & Long, C. (1986). Peer-assisted leadership: Principals learning from each other. *Phi Delta Kappan, 67*(9), 672-675.

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of personality and social psychology, 51*(6), 1173-1182.

Barth, R. (1986). On sheep and goats and school reform. *Phi Delta Kappan, 68*(4), 293-296.

Barth, R. (1980). *Run school run.* Cambridge, MA: Harvard University Press.

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

Barth, R., & Deal, T.E. (1982). *The principalship: Views from without and within.* Prepared for the national conference on the principalship convened by the National Institute of Education. ERIC Document 224176.

Bass, B. (2008). *Bass and Stogdill's handbook of leadership,* 3rd edition. New York: Free Press.

Bass, B. (1985). *Leadership and performance beyond expectations.* New York: Free Press.

Bass, B. M., & Avolio, B. J. (1995). *MLQ multifactor leadership questionnaire for research: Permission set.* Redwood City, CA: Mind Garden, Inc.

Bass, B., & Avolio, B. (1994). I*mproving organizational effectiveness through transformational leadership*. Newbury Park, CA: Sage.

Belchetz, D., & Leithwood, K. (2007). Successful leadership: Does context matter and if so, how? In C. Day & K. Leithwood (Eds.), *Successful principal leadership in times of change: An international perspective* (pp. 117-137). Dordrecht: Springer.

Bell, L., Bolam, R., & Cubillo, L. (2003). *A systematic review of the impact of school headteachers and principals on student outcomes*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Bennis, W. (2003). *On becoming a leader*. Boston: Perseus Publishing.

Berman, P. & McLaughlin, M. (l978). Implementation of educational innovations. *Educational Forum, 40,* 347-370.

Bishop, P., & Mulford, B. (1996). Empowerment in four primary schools: They don't really care. *International Journal of Educational Reform, 5*(2),193-204.

Biester, T., Kruse, J., Beyer, F., & Heller, B. (1984, April). *Effects of administrative leadership on student achievement*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Blank, R. (1987). The role of the principal as leader: Analysis of variation in leadership in urban high schools. *Journal of Educational Research, 82*(2), 69-80.

Blasé, J. (1987). Dimensions of effective school leadership: The teacher's perspective. *American Educational Research Journal, 24(*4), 589-610.

Blase, J., & Blasé, J. (1996). Principals' instructional leadership and teacher development: Teachers' perspectives. *Educational Administration Quarterly, 35*(3), 349-378.

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

Boyan, N. (1988). *Handbook of research in educational administration*. New York: Longman.

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

Brewer, D. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review, 12(*4), 281-292.

Bridges, E. (1967). A model for shared decision making in the school principalship. *Educational Administration Quarterly, 3*(1), 49-61.

Bridges, E. (1970). Administrative man: Origin or pawn in decision making? *Educational Administration Quarterly, 6*(1), 7-25.

Bridges, E. (1967). Instructional leadership: A concept re-examined. *Journal of Educational Administration, 5*(2), 136-147.

Bridges, E. (1982). Research on the school administrator: The state-of-the-art, 1967-1980. *Educational Administration Quarterly, 18*(3), 12-33.

Bridges, E. (1977). The nature of leadership. In L. Cunningham, W. Hack, & R. Nystrand (Eds.), *Educational administration: The developing decades* (pp. 202-230). Berkeley: McCutchan.

Briner, C., & Campbell, R.F. (1964). The science of administration. *Review of Educational Research, 34*(4), 485-492.

Brookover, W. & Lezotte, L. (1977). *Changes in school characteristics coincident with changes in student achievement*. East Lansing: Michigan State University Press.

Browne-Ferrigno, T., & Muth, R. (2004). Leadership mentoring in clinical practice: Role socialization, professional development, and capacity building. *Educational Administration Quarterly, 40*, 468-494.

Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. (2009). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Buchmann, C., & Dalton, B. (2002). Interpersonal influences and educational aspiration in 12 countries: The importance of institutional context. *Sociology of Education*, *75*(2), 99-122.

Burns, J. (1978). *Leadership*. New York: Harper and Row.

Bycio, P., & Allen, J. (1995). Further assessment of Bass's (1985) conceptualization of transactional and transformational leadership. *Journal of Applied Psychology, 80*(4), 468-478.

Bush, T., & Glover, D. (2003). *School leadership: Concepts and evidence*. Commissioned report. National College for School Leadership, Nottingham: UK.

Buttram, J.L., Mead, H., Loftus, D., & Wilson, J.O. (2006). Allocation of school leaders' time. Retrieved May 4th, 2010 from http://www.dasl.udel.edu/sites/dasl.udel.edu/files/users/user10/Allocation%20of%20School%20Leaders%27%20Time.pdf

Carnegie Foundation. (2010). Carnegie classifications. Carnegie Foundation for the Advancement of Teaching. http://classifications.carnegiefoundation.org/ descriptions/basic.php

Calsyn, R. J., Winter, J. P., & Burger, G. K. (2005). The relationship between social anxiety and social support in adolescents: a test of competing causal models. *Adolescence*. *40*(157), 103-13.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin company.

Campbell, R.F. (1979). A critique of the Educational Administration Quarterly. *Educational Administration Quarterly, 15*(3), 1-19.

Campbell, R.F., & Faber C. (1961). Administrative behavior: Theory and research. *Review of Educational Research, 31*(4), 353-367.

Cheng, K.M. (1995). The neglected dimension: Cultural comparison in educational administration. in K.C. Wong KC & K. M. Cheng (Eds.), *Educational leadership and change: An international perspective* (pp. 87–104). Hong Kong: Hong Kong University Press.

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), 299-317.

Cohen, E., & Miller, R. (1980). Coordination and control of instruction in schools. *Pacific Sociological Review, 4*, 446-473.

Cohen, D., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102,* 294-343.

Coleman, J., Campbell, E., Hobson, C., McPartland, F., Mood, A., Weinfeld, F. et al. (1966). *Equality of educational opportunity study*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education/National Center for Education Statistics.

Cook, T. & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings.* Boston: Houghton Mifflin Co.

Conger, J., & Pearce, C. (2003). A landscape of opportunities: Future research on shared leadership. In C. L. Pearce and J. Conger (eds.), *Shared leadership: Reframing the hows and whys of leadership* (pp. 285-303). Thousand Oaks, CA: Sage Publications.

Covey, S. R. (2004). *The seven habits of highly effective people.* New York: Free Press.

Crandall, D., Eiseman, J., & Louis, K.S. (1986). Strategic planning issues that bear on the success of school improvement efforts. *Educational Administration Quarterly, 22*(3), 21-53.

Cronbach, L. J. (1951). Coefficient *alpha* and the internal structure of tests. *Psychometrika, 16(3),* 297-334.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements.* New York: Wiley.

Crowther, F., Ferguson, M., & Hann, L. (2008). *Developing teacher leaders.* Thousand Oaks, CA: Corwin Press.

Cuban. L. (1988). *The managerial imperative and the practice of leadership in schools.* Albany, NY: SUNY Press.

Cuban, L. (1984). Transforming the frog into a prince: Effective schools research, policy, and practice at the district level. *Harvard Educational Review, 54*(2), 128-151.

Cunningham, J. (2004). *The effect of gender-role identity of female principals on in Texas on teacher perceptions of instructional leadership behaviors.* Unpublished doctoral dissertation, Texas A&M University - Corpus Christi, TX.

Darling-Hammond, L. (2000, January). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1), 1-49. Retrieved January 25, 2004, from http://epaa.asu.edu/epaa/v8n1/

Datnow, A., & Castellano, M. (2001). Managing and guiding school reform: Leadership in Success for All schools. *Educational Administration Quarterly, 37*(2), 219-249.

Davies, B., Ellison, L., & Bowring-Carr, C. (2005). *School leadership in the 21st century: developing a strategic approach*. London: Routlegde-Falmer.

Day, D., Gronn, P., & Salasc, S. (2006). Leadership in team-based organizations: On the threshold of a new era. *The Leadership Quarterly, 17(3*), 211-216.

Day, C., Harris, A., & Hadfield, M. (2001). Challenging the orthodoxy of effective school leadership. *International Journal of Leadership in Education, 4*(1), 39-56.

Day, C., Sammons, P., Leithwood, K., Hopkins, D., Harris, A., Gu, Q., & Brown, E. (2010). *Ten strong claims about successful school leadership*. Nottingham, UK: The National College for School Leadership.

Deal, T., & Peterson, K. (1999). *Shaping school culture: The heart of leadership*. San Francisco, Jossey Bass.

Dimmock, C. & Walker, A. (2000). Introduction justifying a cross-cultural comparative approach to school leadership and management. *School Leadership and Management, 20*(2), 137-141.

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy, 10*(1), 45-53.

Donaldson, G. A. (2001). *Cultivating leadership in schools: Connecting people, purpose, and practice*. New York: Teachers College Press.

Duke, D. (2004). The turnaround principal: High-stakes leadership, *Principal, 84*(1), 12-23.

Duke, D., & Stiggins, R. (1985). Evaluating the performance of principals: A descriptive study. *Educational Administration Quarterly 21*(4) 71-98.

Dwyer, D. (1986). Understanding the principal's contribution to instruction. *Peabody Journal of Education, 63*(1), 3-18.

Dwyer, D., Lee, G., Rowan, B., & Bossert, S. (1983). *Five principals in action; Perspectives on instructional management*. San Francisco, CA: Far West Laboratory for Educational Research. ERIC Document number 231085.

Eagly, A.H., Karau, S., & Johnson, B.T. (1992). Gender and leadership style among school principals: A meta-analysis. *Educational Administration Quarterly,28*(1), 76-102.

Ebel, R. (1951). Estimation of the reliability of ratings. *Psychometrika, 16*, 407-424.

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

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

Edmonds, R. (1982). Programs of school improvement: An overview. *Educational Leadership, 40*(3), 4-11.

Edmondson, A., Roberto, M., & Watkins, M. (2003). A dynamic model of top management team effectiveness. *Leadership Quarterly, 14*, 297-325.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Hillsdale, NJ: Erlbaum.

Erickson, D.A. (1979). Research on educational administration: The state-of-the-art. *Educational Researcher, 8*(3), 9-14.

Erickson, D. (1967). The school administrator. *Review of Educational Research, 37*(4), 417-432.

Evans, R. (1996). *The human side of school change*. San Francisco: Jossey Bass.

Fiedler, F. (1967). *A theory of leadership effectiveness.* New York: McGraw-Hill.

Firestone, W. A. & Corbett, H. D. (1988). Planned educational change. In N. Boyan (Ed.) *Handbook of research on educational administration* (pp. 321-340). New York: Longman.

Foster, R. (2005). Leadership and secondary school improvement: Case studies of tensions and possibilities. *International Journal of Leadership in Education, 8*(1), 35-52.

Fraenkel, J., & Wallen, N. (1990). *How to design and evaluate research*. New York: McGraw Hill.

Fullan, M. (2001). *Leading in a culture of change.* San Francisco, CA: Jossey-Bass Inc.

Fullan, M. (2002). The change leader. *Educational Leadership, 59*(8),16-20.

Fullan, M. (2007). *The new meaning of educational change* (4th Edition). New York: Teachers College Press.

Fullan, M. (2006). *The development of transformational leaders for educational decentralization.* Toronto, Canada: Michael Fullan.

Fullan, M., & Hargreaves, A. (1992). *Teacher development and educational change*. London/Washington: Falmer

Gay, L. (1992). *Educational research.* New York: MacMillan.

Getzels, J., Lipham, J., & Campbell, R. (1968). *Educational administration as a social process.* New York: Harper & Row.

Gewertz, C. (2003, Jan. 8). N.Y.C. Chancellor Aims to Bolster Instructional Leadership, *Education Week, 22*(16), 7, 12.

Glass, G.V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education, 5*, 351-379.

Glenn, J. (1975). *Chief executive time: An empirical study of the time allocation of American college and university presiden*ts. Unpublished doctoral dissertation, Stanford university.

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

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

Goldring, E., Cravens, X. C., Murphy, J., Porter, A., & Elliott, S. (2012). *The convergent and divergent validity of the Vanderbilt Assessment of Leadership in Education™ (VAL-ED): Instructional leadership and emotional intelligence.* Paper presented at the Annual Conference of Association for Education Finance and Policy, Boston, MA.

Goldring, E., Cravens, X, Murphy, J., & Elliot, S. (2009). The evaluation of principals: What and how do states and urban districts assess leadership? *Elementary School Journal. 110*(1), 19-39.

Goldring, E., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration, 46*(3), 332-352.

Goldring E., & Pasternak, R. (1994). Principals' coordinating strategies and school effectiveness. *School Effectiveness and School Improvement, 5*(3), 239-253.

Goldring, E., Porter, A., Murphy, J., Elliot, S. N., & Cravens, X. (2009). Assessing learning-centered leadership: Connections to research, professional standards, and current practices. *Leadership and Policy in Schools, 8*, 1-36.

Goldring, E. (2010). *The next generation of principal professional development: Feedback and coaching as a tool for leadership development*. Paper presented at the Keynote paper presented at the conference: The challenges and prospects of school improvement in the new era. Taipei, Taiwan

Goldring, E., & Sullivan, A. (1996). Beyond the boundaries: Principals, parents & communities shaping the school environment. In K. Leithwood, J. Chapman, D. Carson, P. Hallinger, & A. Hart (Eds.), *International Handbook of Educational Leadership and Administration* (Vol. 1) (pp. 195-222). Dordrecht, The Netherlands: Kluwer Academic Press.

Glover, J., Rainwater, K., Friedman, H., & Jones, G. (2002). Four principles for being adaptive (Part Two). *Organizational Development Journal, 20/4, 18-38.*

Greenleaf, R. (1977). *Servant leadership*. Mahwah, N.J.: Paulist press.

Griffin, M. A. (1997). Interaction between individuals and situations: Using HLM procedures to estimate reciprocal relationships. *Journal of Management*. *23*(6), 759-773.

Griffiths, D.E. (1979). Intellectual turmoil in educational administration. *Educational Administration Quarterly*, *15*(3), 43-65.

Grobman, H., & Hynes, V. (1956). What makes a good principal? *NASSP Bulletin 40*(223), 5-16.

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

Gronn, P. (2009). Hybrid leadership. In K. Leithwood, B., Mascall, & T. Strauss (Eds.), *Distributed leadership according to the evidence*. London: Routledge.

Gronn, P. (2003). *The new work of educational leaders: Changing leadership practice in an era of school reform*. London: Sage.

Gross, N., & Herriot, R. (1965). *Staff leadership in schools*. New York: Wiley.

Grubb, W. N., and J. Flessa. (2009). "A job too big for one": Multiple principals and other non-traditional approaches to school leadership. In K. Leithwood, B., Mascall, and T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 137-164). London: Routledge.

Hair. J. F., Black. B., Babin. B., Anderson, R. E. & Tatham, R. L., (1998). *Multivariate data analysis.* New Jersey: Prentice Hall.

Hakstian, A. R., & Whalen, T. E. (1976). A *k*-sample significance test for independent *alpha* coefficients. *Psychometrika, 41*(2), 219-231.

Hall, G. & Hord, S. (2002). *Implementing change: Patterns, principles, and potholes*. Boston: Allyn & Bacon.

Hall, V., & Southworth, G. (1997). Headship. *School Leadership & Management, 17*(2), 151-170.

Haller, E. (1979). Questionnaires and the dissertation in educational administration. *Educational Administration Quarterly, 15*(1), 47-66.

Hallinger, P. (2011a). A review of three decades of doctoral studies using the *Principal Instructional Management Rating Scale*: A lens on methodological progress in educational leadership. *Educational Administration Quarterly, 47*(2) 271-306.

Hallinger, P. (1983). *Assessing the instructional management behavior of principals*. Unpublished doctoral dissertation, Stanford University, Stanford, CA. ERIC Document No. 8320806.

Hallinger, P. (1995). Culture and leadership: Developing an international perspective in educational administration. *UCEA Review, 36*(1), 3-7.

Hallinger, P. (2011b). Developing a knowledge base for educational leadership and management in East Asia. *School Leadership and Management, 31*(4), 305-320.

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools, 4*(3), 221-240.

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

Hallinger, P. (2011c). Leadership for learning: Lessons from 40 years of empirical research. *Journal of Educational Administration, 49*(2), 125-142.

Hallinger, P. (2010). Making education reform happen: Is there an "Asian" way? *School Leadership and Management, 30*(5), 401-408.

Hallinger, P. (1982, 1990). *Principal instructional management rating scale.* Sarasota, FL: Leading Development Associates.

Hallinger, P. (In press). Reviewing reviews of research in educational leadership: An empirical analysis. *Educational Administration Quarterly.*

Hallinger, P., & Bridges, E. (2007). *Problem-based management education: Developing "managers for action."* Dordrecht, Netherlands: Springer.

Hallinger, P., & Bryant, D.A. (2013a). Mapping the terrain of research on educational leadership and management in East Asia. *Journal of Educational Administration, 51*(5), Online First May, 5, 2013. www.emeraldinsight.com/journals.htm?articleid=17087698.

Hallinger, P., & Bryant, D.A. (2013b). Review of research publications on educational leadership and management in Asia: A comparative analysis of three regions. *Oxford Review of Education*, 39(3), 307-328.

Hallinger, P., & Heck, R. H. (2010). Collaborative leadership and school improvement: Understanding the impact on school capacity and student learning. *School Leadership and Management, 30*(2), 95-110.

Hallinger, P., & Heck, R. H. (2011a). Conceptual and methodological issues in studying school leadership effects as a reciprocal process. *School Effectiveness and School Improvement, 22*(2), 149-173.

Hallinger, P., & Heck, R. H. (2011b). Exploring the journey of school improvement: Classifying and analyzing patterns of change in school improvement processes and learning outcomes. *School Effectiveness and School Improvement 22*(1), 1-27.

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

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

Hallinger, P., & Heck, R. H. (1996b). The principal's role in school effectiveness: An assessment of methodological progress, 1980-1995. In K. Leithwood et al. (Ed.), *International handbook of research in educational leadership and administration* (pp. 723-784). New York: Kluwer Press.

Hallinger, P., & Heck, R. H. (2002). What do you call people with visions? The role of vision, mission and goals in school leadership and improvement. In K. Leithwood, P. Hallinger and Colleagues (Eds.), *The Handbook of Educational Leadership and Administration (Second Edition)*. Dordrecht: Kluwer.

Hallinger, P., & Lee, M.S. (2013). Exploring principal capacity to lead reform of teaching and learning quality in Thailand. *International Journal of Educational Development, 33,* 305-315.

Hallinger, P., & Lee, M.S. (In press). Mapping instructional leadership in Thailand: Has education reform impacted principal practice? *Educational Management, Administration and Leadership.*

Hallinger, P., & Leithwood, K. (1994). Exploring the impact of principal leadership. *School Effectiveness and School Improvement, 5*(3), 206-218.

Hallinger, P., & Leithwood, K. (1998). Unseen forces: The impact of social culture on leadership. *Peabody Journal of Education, 73*(2), 126-151.

Hallinger, P., & Murphy, J. (1985). Assessing the instructional leadership behavior of principals. *Elementary School Journal, 86*(2), 217-248.

Hallinger, P., &Murphy, J. (1986). The social context of effective schools. *American Journal of Education, 94*(3), 328-355.

Hallinger, P., Murphy, J., Weil, M., Mesa, R., P., & Mitman, A. (1983). Effective schools: Identifying the specific practices and behaviors of the principal, *NASSP Bulletin, 67*(463), 83-91.

Hallinger, P., Taraseina, P., & Miller, J. (1994). Assessing the instructional leadership of secondary school principals in Thailand. *School Effectiveness and School Improvement, 5*(4), 321-348.

Hallinger, P., Walker, A., & Bajunid, I. (2005). Educational leadership in East Asia: Implications for education in a global society. *UCEA Review, 1*, 1-4.

Hallinger, P., & Wang, W.C., & Chen, W.C. (2013). Assessing the measurement properties of the Principal Instructional Management Rating Scale. A meta-analysis of reliability studies. *Educational Administration Quarterly.*

Hargreaves, A. & Fullan, M. (1998). *What's worth fighting for out there.* New York: Teachers College Press.

Harris, A. (2008). Distributed leadership: According to the evidence. *Journal of Educational Administration, 46*(2), 172-188.

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

Harris, A. (2006). Leading change in schools in difficulty. *Journal of Educational Change, 7*(1/2), 9-18.

Harris, A. (2003). Teacher leadership and school improvement. In A. Harris, C. Day, D. Hopkins, M. Hadfield, A. Hargreaves, & C. Chapman (Eds.), *Effective leadership for school improvement* (pp. 72-83). London: Routledge/Falmer.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Hawley W., & Rosenholtz, S. (1984). Good schools: What research says about improving school achievement. *Peabody Journal of Education, 61*, 117-124.

Hayes, D., Christie, P., Mills, M., & Lingard, B. (2004). Productive leaders and productive learners: Schools as learning organisations. *Journal of Educational Administration, 42*(4/5), 520–538.

Heck, R. H. (2006). Assessing school achievement progress: Comparing alternative approaches. *Educational Administration Quarterly, 42*(5), 667-699.

Heck, R. H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Educational Administration Quarterly, 36*(4), 513-552.

Heck, R. (1996) Leadership and culture: Conceptual and methodological issues in comparing models across cultural settings, *Journal of Educational Administration, 30*(3), 35–48.

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), 1-4.

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

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

Heck, R. H. & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement**.** *American Educational Research Journal, 46*, 626-658

Heck, R.H., & Hallinger, P. (1999). Conceptual models, methodology, and methods for studying school leadership. In J. Murphy & K. Seashore-Louis (Eds.), *The 2nd handbook of research in educational administration* (pp. 141-162). San Francisco: McCutchan.

Heck, R. H., & Hallinger, P. (2011). Collaborative leadership effects on school improvement: Integrating unidirectional- and reciprocal-effects models. *The Elementary School Journal, 111*(2).

Heck, R., & Hallinger, P. (1999). Next generation methods for the study of leadership and school improvement. In J. Murphy & K. S. Louis (Eds.), *Handbook of research on educational administration*, 2nd edition (pp. 141-162). San Francisco: Jossey-Bass.

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

Heck, R.H., Larson, T., & Marcoulides, G. (1990). Instructional leadership and student achievement: Validation of a causal model. *Educational Administration Quarterly*, *26*(2), 94-125.

Hersey, P. and Blanchard, K. H. (1977). *Management of organizational behavior: Utilizing human resources* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

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

Hoffer, T.B., Welch, V., Williams, K., Hess, M., Webber, K., & Lisek, B. (2005). *Doctorate recipients from United States universities: Summary report 2004.* Chicago: National Opinion Research Center.

Hofstede, G. (1997). *Culture and organizations: Software of the mind*, New York: McGraw Hill.

Hofstede, G. (2001). *Culture's consequences: International differences in work-related values.* Thousand Oaks, CA: Sage.

Holmes, H. & Tangtongtavy, S. (1996). *Working with Thais: A guide to managing in Thailand.* Bangkok: White Lotus.

Hooijberg, R., & Schnieder, M. (2001). Behavioral complexity and social intelligence: How executive leaders use stakeholders to form a systems perspective. In S. Zaccaro and R. Klimoski (Eds.), *The nature of organizational leadership: Understanding the performance imperatives confronting today's leaders* (pp. 104-131). San Francisco: Jossey-Bass.

Horng, E.L., Klasik, D., & Loeb, S. (2010). Principal time-use and school effectiveness. National Center for the Analysis of Longitudinal Data in Education research. Retrieved June 1st 2010 from www.stanford.edu/.../Principal%20Time-Use%20Research%20Paper%20(revised).pdfHunter Foundation. (2005). New drive to improve school leadership in Scotland. http://www.thehunterfoundation.co.uk/projects/?project_id=5.

House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies.* Thousand Oaks, CA: Sage.

Howell, J. M., & Avolio, B. J. (1993). Transformational leadership, transactional leadership, locus of control, and support for innovation: Key predictors of consolidated-business-unit performance. *Journal of applied psychology, 78*(6), 891-902.

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

Jacobson, S. (2011). Leadership effects on student achievement and sustained school success. *International Journal of Educational Management, 25*(1), 33-44.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: ASCD.

Kane, M. T., Gilmore, G. M., & Crooks, T. J. (1976). Student evaluations of teaching: The generalizability of class means. *Journal of Educational Measurement, 13*(3), 171-183.

Keller, R. T. (2006). Transformational leadership, initiating structure, and substitutes for leadership: A longitudinal study of research project team performance. *Journal of Applied Psychology, 91*(1), 202-210.

Kerlinger, F. (1966). *Foundations of behavioral research*. New York: Holt Rinehart & Winston.

Kerr, S, & Jermier, J. (1978). Substitutes for leadership: Their meaning and measurement. *Organizational behavior and human performance. 22*(3), 375-403.

Kimberly, J., & Miles, R. (1980). *The organizational life cycle*. San Francisco: Jossey-Bass.

Kirby, P., Paradise, L., & King, M. (1992). Extraordinary leaders in education: Understanding transformational leadership. *Journal of Educational Research, 85*(5), 303-311.

Kliene-Kracht, P. (1993). Indirect instructional leadership: An administrator's choice. *Educational Administration Quarterly, 18*(4), 1–29.

Kmetz, J., & Willower, D. (1982). Elementary school principals' work behavior. *Educational Administration Quarterly, 18*(4), 62-78.

Knapp, M., Copland, M., Honig, M., Plecki, M., & Portin, B. (2009). *Learning-focused leadership and leadership support: Meaning and practice in urban systems.* Center for the Study of Teaching and Policy, University of Washington, Seattle, WA.

Kotter, J. (1996). *Leading change*. Boston, HBS Press.

Kotter, J., & Heskett, J. (1992). *Corporate culture and performance*. New York: The Free Press.

Kouzes, J. M., & Posner, B. Z. (1995). *The leadership challenge.* San Francisco: Jossey-Bass.

Kouzes, J. M., & Posner, B. Z. (2002). *The leadership challenge* (3rd ed.). San Francisco: Jossey-Bass.

Kouzes, J., & Posner, B. (2007). *The leadership challenge*. San Francisco: Jossey Bass.

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

Kyriakides, L., & Creemers, B. P. M. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement, 19*(2), 183–205.

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

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

Lang, G., & Heiss, G.D. (1998). *A Practical guide to research methods*. New York: University Press of America.

Latham, G. & Wexley, K. (1981). *Increasing productivity through performance appraisal*. Menlo Park, CA: Addison Wesley.

Lee, V., & Bryk, A. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62*(3), 172-192.

Leithwood, K. (1994). Leadership for school restructuring. *Educational Administration Quarterly.* 30(4), 498-518.

Leithwood, K. (2001). School leadership in the context of accountability policies. *International Journal of Leadership in Education, 4*(3), 217-235.

Leithwood, K., Anderson, S., Mascall, B., & Strauss, T. (In press). School leaders' influences on student learning: The four paths. In T. Bush, L., Bell and D. Middlewood (Eds.), *The principles of educational leadership and management*. London: Sage publishers.

Leithwood, K., Begley, P. & Cousins, B. (1990). The nature, causes and consequences of principals' practices: An agenda for future research. *Journal of Educational Administration, 28*(4), 5-31.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). *Seven strong claims about successful school leadership*. National College of School Leadership, Nottingham, UK.

Leithwood, K., Hallinger, P., Chapman, J., Corson, D. & Hart, A. (Eds.; 2002). *The second international handbook of educational leadership and administration*. Dordrecht, Netherlands: Kluwer Press.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management, 28*(1), 27-42.

Leithwood, K., & Jantzi, D. (2005). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research*, 79(1), 464-490.

Leithwood, K. & Jantzi, D. (2000). The effects of transformational leadership on organizational conditions and student engagement with the school. *Journal of Educational Administration, 38*(2), 112-129.

Leithwood, K., & Jantzi, D. (2005). A review of transformational school leadership research 1996–2005. *Leadership and Policy in Schools, 4*(3), 177-199.

Leithwood, K. & Jantzi, D. (1999). The relative effects of principal and teachers sources of leadership on student engagement with school. *Educational Administration Quarterly, 35,* 679-706.

Leithwood, K., Louis, K.S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. Center for Applied Research and Educational Improvement, University of Minnesota.

Leithwood, K., & Louis, K.S. (2000). *Understanding schools as intelligent systems*. New York: JAI Press.

Leithwood, K., Mascall, B., & Strauss, T. (2009). New perspectives on an old idea. In K. Leithwood, B., Mascall, & T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 1-14). London: Routledge.

Leithwood, K., Mascall, B., & Strauss, T. (2009). What we have learned where we go from here (pp. 269-282). In K. Leithwood, B. Mascall, B., & T. Strauss (Eds.), *Distributed leadership according to the evidence*. New York: Routledge.

Leithwood, K., & Montgomery, D. (1986). *Improving principal effectiveness: The principal profile*. Toronto: OISE Press.

Leithwood, K., Patten, S., & Jantzi, D. (2010). *Testing a conception of how school leadership influences learning*. Unpublished paper. Toronto: U. of Toronto/OISE.

Leithwood, K. A., & Stager, M. (1989). Expertise in principals' problem solving. *Educational Administration Quarterly 25*(2), 126-161.

Leithwood, K. & Steinbach, R. (1991). Indicators of transformational leadership in everyday problem-solving of school administrators. *Journal of Personnel Evaluation in Education, 4*(3), 221-244.

Leithwood, K., & Sun, J.P. (2012). The nature and effects of transformational school Leadership: A meta-analytic review of unpublished research. *Educational Administration Quarterly* published online DOI: 10.1177/0013161X11436268.

Leithwood, K., Tomlinson, D., & Genge, M. (1996). Transformational school leadership. In K. Leithwood, J. Chapman, D. Corson, P. Hallinger, & A. Weaver-Hart (Eds.), *International handbook of educational leadership and administration* (pp. 785-840). New York: Kluwer.

Levitt, B., & March, J. (1988). Organizational learning. In W. Scott and J. Blake (Eds.), *Annual Review of Sociology, 14,* 319–340.

Linacre, J. M. (2005). *WINSTEPS Rasch measurement* [Computer software]. Chicago: Winsteps.com.

Lipham, J. (1961), *Effective principal, effective school*, National Association of Secondary School Principals, Reston, VA.

Lipham, J. (1964). Organizational character of education: Administrative behavior. *Review of Educational Research, 34*(4), 435-454.

Lissitz, R. W. & Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher, 36,* 437-448.

Locke, E. (2003). Leadership: Starting at the top. In C. J. Pearce & C. Conger (Eds.), *Shared leadership: Reframing the hows and whys of leadership* (pp. 271-284). Thousand Oaks, CA: Sage.

Lord, R. (2001). The nature of organizational leadership: Conclusions and implications. In S. Zaccaro & R. Klimoski (Eds.), *The nature of organizational leadership* (pp. 413-436). San Francisco: Jossey-Bass.

Louis, K.S., Dretzkea, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement, 21*(3), 315-336.

Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). Investigating the links to improved student learning: Final report of research findings. University of Minnesota, USA. Downloaded from http://www.cehd.umn.edu/carei/Leadership/Learning-from-Leadership_Final-Research-Report_July-2010.pdf on August 14, 2011.

Louis, K. S., & Marks, H. (1998). Does professional community affect the classroom? *American Journal of Education, 106*(4), 532-575.

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

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

MacBeath, J., & Cheng, Y.C. (2008). *Leadership for learning: International perspectives*. Rotterdam: Sense Publishers.

March, J. (1978). The American public school administrator: A short analysis. *School Review,* 86(2), 217-250.

Marcoulides, G., & Heck, R. (1993). Organizational culture and performance: Proposing and testing a model. *Organization Science, 4*(2), 209-225.

Marks, H., & Printy, S. (2003). Principal leadership and school performance: An integration of transformation and instructional leadership. *Educational Administration Quarterly, 39*(3), 370-397.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133-163.

Marshall, K. (1996). How I confronted HSPS (Hyperactive Superficial Principal Syndrome) and began to deal with the heart of the matter. *Phi Delta Kappan, 76*(5), 336-345.

Marshall, K. (2004). How I recovered from HSPS (Hyperactive Superficial Principal Syndrome: A progress report. *Phi Delta Kappan, 84*(9), 701-709.

Martin, W., & Willower, D.J. (1981).The managerial behavior of high school principals. *Educational Administration Quarterly*, *17*(1), 69-90.

Martinko, M.J. & Gardner, W.L. (1990). Structured observation of managerial work: A replication and synthesis. *Journal of Management Studies, 27*, 329-357.

Mathieu, J., Ahearne, M., & Taylor, S. (2007). A longitudinal cross-level model of leader and salesperson influences on sales force technology use and performance. *Journal of Applied Psychology, 92*(2), 528-537.

Maurer, R. (1996). *Beyond the wall of resistance.* Austin, TX: Bard Books.

May, H., Huff, J., & Goldring, E. (2012). A longitudinal study of principals' activities and student performance. *School Effectiveness and School Improvement*. DOI: 10.1080/09243453.2012.678866

May, H., & Supovitz, J. (2011). The scope of principal efforts to improve instruction. *Educational Administration Quarterly, 47*(2) 332–352.

McLaughlin, M. (1990). The Rand change agent study revisited. *Educational Researcher, 5,* 11-16.

Meindl, J. (1995). The romance of leadership as follower-centric theory: A social constructionist approach. *The Leadership Quarterly, 6,* 329-341.

Merton, R. (1957). The role set: Problems in sociological theory. *British Journal of Sociology,*

Meyer, J., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83,* 340-363;

Miller, G. (1960, April). *What is the role of the principal in promoting good relationships with staff?* Proceedings of the annual meeting of the National Association of Secondary School Principals, 19-22.

Mintzberg, H. (1973). *The nature of managerial work.* New York: Harper & Row.

Mohr, L. (1983). The implications of effectiveness theory for managerial practice in the public sector. In K. Cameron & D. Whetton (Eds.), *Organizational effectiveness: A comparison of multiple models* (pp. 225-39). New York: Academic Press.

Monge, P. (1990). Theoretical and analytical issues in studying organizational processes. *Organization Science, 1*(4), 406-430.

Mortimore, P. (1993). School effectiveness and the management of effective learning and teaching. *School Effectiveness and School Improvement, 4*, 290-310.

Mulford, B., & Silins, H. (2003). Leadership for organisational learning and improved student outcomes - What do we know? *Cambridge Journal of Education, 33*(2), 175-195**.**

Mulford, B., & Silins, H. (2009). Revised models and conceptualization of successful school principalship in Tasmania. In B. Mulford and B. Edmunds (Eds.) *Successful school principalship in Tasmania.* Launceston, Tasmania: Faculty of Education, University of Tasmania.

Murphy, J. (2005). *Connecting teacher leadership and school improvement*. Newbury Park, CA: Corwin Press.

Murphy, J. (1988). Methodological, measurement and conceptual problems in the study of instructional leadership. *Educational Evaluation and Policy Analysis, 10*(2), 117-139.

Murphy**,** J**.** (2002). Reculturing the profession of educational leadership: New blueprints. *Educational Administration Quarterly, 38*(2), 176-192.

Murphy, J., Elliott, S. N., Goldring, E., & Porter A. (2007). Leadership for learning: A research-based model and taxonomy of behaviors. *School Leadership & Management, 27*(2), 179-201.

Murphy, J. & Hallinger, P. (1989). Equity as access to learning: Curricular and instructional differences. *Journal of Curriculum Studies, 21*(2), 129-149.

Murphy, J., Hallinger, P., Lotto, L., & Miller, S. (1987). Barriers to implementing the instructional leadership role. *The Canadian Administrator, 27*(3), 1-9.

Murphy, J., Hallinger, P., & Mitman, A. (1983). Problems with research on educational leadership: Issues to be addressed. *Educational Evaluation and Policy Analysis, 5*(3), 297-306.

Murphy, J. & Meyers, C. (2008). *Turning around failing schools: Leadership lessons from the organizational sciences*. Thousand Oaks, CA: Corwin.

Murphy, J., & Shipman, N. (2003). Developing standards for school leadership development: A process and rationale. In P. Hallinger (Ed.), *Reshaping the landscape of school leadership development: A global perspective*. Lisse, Netherlands: Swets & Zeitlinger.

Murphy, J., & Vriesenga, M. (2005). Developing professionally anchored dissertations: Lessons from innovative programs. *School Leadership Review, 1*(1), 33-57.

Murphy, J., & Vriesenga, M. (2006). Research on school leadership preparation in the United States: An analysis. *School Leadership and Management, 26*(2), 183-195.

Murphy, J., Weil, M., Hallinger, P., & Mitman, A. (1982). Academic press: Translating high expectations into school policies and practices. *Educational Leadership, 40*(3), 22-27.

Murphy, K. R., & Deshon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53*(4), 873–900.

Neumerski, C. (2012). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly*. Online first.

Nettles, S., & Herrington, C. (2007). Revisiting the importance of the direct effects of school leadership on student achievement: The implications for school improvement policy. *Peabody Journal of Education, 82*(4), 724-736.

Nunnally, J. (1978). *Psychometric methods*. New York McGraw Hill.

Nicolaidou, M., & Ainscow, M. (2005). Understanding failing schools: Perspectives from the inside. *School Effectiveness and School Improvement, 16*(3), 229-248.

O'Donnell, R., & White, G. (2005). Within the accountability era: Principals' instructional leadership behaviors and student achievement. *NASSP Bulletin, 89*, 645, 56-71.

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

Ogawa, R. & Hart, A. (1985). The effect of principals on the instructional performance of

schools. *Journal of Educational Administration, 22(*1), 59-72

Opdenakker, M., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal, 33*(2), 179-206.

O'Toole, J. (1995). *Leading change.* San Francisco: Jossey Bass.

Ouston, J. (1999). School effectiveness and school improvement: Critique of a movement. In T. Bush, L. Bell, R. Bolam, R. Glatter, & P. Ribbins (Eds.), *Educational management: Redefining theory, policy and practice* (pp. 166-177). London: Paul Chapman.

Peterson, K.D. (1977-78). The principal's tasks. *Administrator's Notebook, 26*(8), 1-4.

Pitner, N. (1988). The study of administrator effects and effectiveness. In N. Boyan (Ed.), *Handbook of research in educational administration* (pp. 106-132). New York: Longman.

Ployhart, R., Holtz, B., & Bliese, P. (2002). Longitudinal data analysis applications of random coefficient modeling to leadership research. *Leadership Quarterly, 13*(4), 455-486.

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

Podsakoff, P., MacKenzie, S., & Fetter, R. (1993). Substitutes for leadership and the management of professionals. *Leadership Quarterly, 4*, 1-44.

Porter, A., Murphy, J., Goldring, E., & Elliott, S. N. (2008). Vanderbilt Assessment of Leadership in Education. Retrieved from http://www.valed.com

Pounder, D. G., Ogawa, R. T., & Adams, E. A. (1995). Leadership as an organization-wide phenomena: Its impact on school performance. *Educational Administration Quarterly, 31*(4), 564-588.

Printy, S.M. (n.d.). *How principals influence instructional practice: Leadership levers.* Unpublished paper. East Lansing, MI: Michigan State University.

Printy, S. M., Marks, H. M. & Bowers, A. (2009). *Integrated leadership: How principals and teachers share instructional influence*. Unpublished manuscript.

Purkey, S., & Smith, M. (1983). Effective schools: A review. *The Elementary School Journal, 83*(4), 426-452.

Purkey, S., & Smith, M. (1985). School reform: The district policy implications of the effective schools literature. *The Elementary School Journal, 85*(3), 352-389.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Institute of Educational Research. (Expanded edition, 1980. Chicago: The University of Chicago Press.)

Reynolds, D., Teddlie, C., Hopkins, D., & Stringfield, S. (2000). Linking school effectiveness and school improvement. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 206-231). London: Falmer Press.

Riehl, C. (2000). The principal's role in creating inclusive schools for diverse students: A review of normative, empirical, and critical literature on the practice of educational administration. *Review of Educational Research, 70*(1), 55-81.

Robinson, V. (2006). Putting education back into educational leadership. *Leading & Managing, 12*(1), 62-75.

Robinson, V. (2007). *School leadership and student outcomes: Identifying what works and why.* Melbourne: Australian Council for Educational Leaders Monograph No. 41.

Robinson, V., Lloyd, C., & Rowe, K. (2008). The Impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635-674.

Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient *alpha*. *Psychological Methods, 11*(3), 306-322.

Rogers, E. (2003). *Diffusion of innovations*. New York: Simon and Schuster.

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, New York.

Rowan, B., Bossert, S., & Dwyer, D. (1983). Research on effective schools: A cautionary note. *Educational Researcher, 12*, 24-31.

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

Rutter, M. (1983). School effects on pupil progress. Research findings and policy implications. In L. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 3-41). White Plains, NY: Longman.

Saphier, J., & King, M. (1985). Good seeds grow in strong cultures. *Educational Leadership, 42*(6), 67-74.

Sammons, P., Nuttall, D., Cuttance, P., & Thomas, S. (1995). Continuity of school effects: A longitudinal analysis of primary and seconday school effects on GCSE performance. *School Effectiveness and School Improvement, 6*, 285-307.

Saphier, J., & King, M. (1985). Good seeds grow in strong cultures. *Educational Leadership, 42*(6), 67-74.

Sarason, S. (1982). *The culture of the school and the problem of change* (Revised Edition). Boston: Allyn & Bacon.

Schlueter, E., Davidov, E. & Schmidt, P (2007). Applying autoregressive cross-lagged and latent growth curve models to a three-wave panel study. In K. Montfort, J. Oud, and A. Satorra (Eds.) *Longitudinal models in the behavioral and related sciences* (pp. 315-337). New York: Routledge.

Schmitt, N. (1996). Uses and abuses of coefficient *alpha*. *Psychological Assessment, 8*, 350-353.

Schoen, L., & Fusarelli, L. (2008). Innovation, NCLB, and the fear factor: The challenge of leading 21st-century schools in an era of accountability. *Educational Policy, 22,* 181-203.

Selig, J. P. & Preacher, K. P. (2009). Mediation models for longitudinal data in developmental research. Research in *Human Development, 6*(2-3), 144-164.

Sheppard, B. (1996). *Exploring the transformational nature of instructional leadership, The Alberta Journal of Educational Research, 42*(4), 325–344.

Silva, J., White, G., & Yoshida, R. (2011). The direct effects of principal–student discussions on eighth grade students' gains in reading achievement: An experimental study. *Educational Administration Quarterly*, published online April 5, 2011.

Silins, H. (1994). The relationship between transformational and transactional leadership and school improvement outcomes. *School Effectiveness and School Improvement, 5*(3) 272-298.

Silins, H., & Mulford, B. (in press). Re-conceptualising school principalship that improves student outcomes. *Journal of Educational Leadership, Policy and Practice.*

Sivasubramaniam, N., Murry, L., Avolio, B., & Jung, D. (2002). A longitudinal model of the effects of team leadership and group potency on group performance. *Group & Organization Management, 27* (1), 66-96.

Sleegers, P., Geijsel, F., & Van den Berg, R. (2002). Conditions fostering educational change. In K. Leithwood (Ed.), *Second international handbook of educational leadership and administration* (pp. 75-102). Dordrecht: Netherlands: Kluwer Academic.

Smith, P.C., & Kendall, L.M. (1983). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology, 47*, 149-155.

Southworth, G. (2002). Instructional leadership in schools: Reflections and empirical evidence. *School Leadership and Management, 22*(1), 73-92.

Southworth, G. (2003). *Primary school leadership in context: Leading small, medium and large sized schools*. London: Taylor and Francis.

Snyder, J. & Ebmeier, H. (1992). Empirical linkages among principal behaviors and intermediate outcomes: Implications for principal evaluation. *Peabody Journal of Education, 68*(1), 75-107.

Spillane, J. (2006). *Distributed leadership*. San Francisco: Jossey-Bass.

Spillane, J. P. (2005). Primary school leadership practice: how the subject matters. *School Leadership and Management, 25*(4), 383-397.

Spillane, J.P. Camburn, E.M., Pareja, A. (2007). Taking a distributed perspective to the school principal's workday. *Leadership and Policy in Schools, 6*, 103-125.

Steers, R. (1975). Problems in the measurement of organizational effectiveness. *Administrative Science Quarterly, 20*(4), 546-58.

Stricherz, M. (2001, Sept. 12). D.C. principal's training designed to boost instructional leadership. *Education Week, 21*(2), 13.

Stricherz, M. (2001b, Sept. 19). Leadership grant aimed at schools in South. *Education Week, 21(*3), 21.

Sweetland, S., & Hoy, W. (2000). School characteristics and educational outcomes: Towards an organisational model of student achievement in middle schools. *Educational Administration Quarterly, 36*(5), 703–729.

Tate, B. (2008). A longitudinal study of the relationships among self-monitoring, authentic leadership, and perceptions of leadership. *Journal of Leadership and Organizational Studies, 15*(1), 16-29.

Teddlie, C., Stringfield, S., & Reynolds, D. (2000). Context issues within school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *International handbook of school effectiveness research* (pp. 160-183). London: Routledge.

Teece, D. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization, 3*, 39-63.

Thomas, A. (1988). Does leadership make a difference to organizational performance? *Administrative Science Quarterly, 33*(22), 401-17.

Thompson, J. (1968). *Organizations in action.* New York: McGraw Hill.

Thompson, J. & McEwen, W. (1958). Organizational goals and environment: Goal-setting as an interaction process. *American Sociological Review, 23*, 23-31.

Timperly, H. (2009). Distributing leadership to improve outcomes for students. In K. Leithwood, B. Mascall, & T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 197-222). London: Routledge.

Turner, C. (2003). A critical review of research on subject leaders in secondary schools. *School Leadership & Management, 23*(2), 209-227.

Tyack, D., & Hansot, E. (1982). Managers of virtue: public school leadership in America: 1920-1980. New York: Basic Books.

Uhls, H. (1962). What's important Mr. Principal? *NASSP Bulletin, 46*(273), 108-111.

van de Grift, W. (1990). Educational leadership and academic achievement in elementary education. *School Effectiveness and School Improvement, 1*(3), 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* (pp. 227-242). Cardiff/Groningen: School of Education/RION.

Villanova, R., Gauthier, W., Proctor, P., & Shoemaker, J. (1981). *The Connecticut school effectiveness questionnaire*. Hartford, CT: Bureau of School Improvement, Connecticut State Department of Education.

Vogelaar, L. W., & Kuipers, H. (1997). Reciprocal longitudinal relations between leader and follower effectiveness. *Military Psychology, 9*, 199-208.

Wahlstrom, K. & Louis, K. S. (2008). How teachers experience principal leadership: The roles of professional community, trust, efficacy, and shared responsibility. *Educational Administration Quarterly, 44*(4), 458-495.

Wallace, M., & Poulson, L.(2003). *Learning to read critically in educational leadership and management*. London: Sage.

Walker, A.D., Hu, R.K., & Qian, H.Y. (2012). Principal leadership in China: An initial review, *School Effectiveness and School Improvement.* DOI: 10.1080/09243453.2012.678863

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1-19.

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

Willms, J. D. (1992). *Monitoring school performance.* London: Falmer.

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

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

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mahwah, NJ: Lawrence Erlbaum Associates.

Witziers, B., Bosker, R., & Kruger, M. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 34*(3), 398-425.

Wright, B. D., Linacre, J. M., Gustafen, J. E., & Martin-Lof, P. (1994) Reasonable meansquare fit values. *Rasch Measurement Transactions, 8*, 370.

Wright, B., & Masters, G. (1982). *Rating scale analysis.* Chicago, IL: MESA.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research, 20*, 557-585.

Wu, M., Adams, R. J., & Wilson, M. (1997). *ConQuest* [Computer Program]. Australian Council of Educational Research.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research, 74*(3), 255-316.

Yukl, G. (2002). *Leadership in organizations (5th ed.).* Upper Saddle River, NJ: Prentice Hall.

---

[i] See Appendix A for a complete list of PIMRS studies.

[ii] This chapter draws extensively on materials published in Hallinger, 2005 2011

The authors are grateful for helpful comments received from Ed Bridges, Ellen Goldring, Ken Leithwood, George Marcoulides, Bill Mulford, Viviane Robinson, and two anonymous reviewers on earlier versions of the paper.

[iii] As is common in the school effectiveness literature, we use the term school effects to indicate statistically significant associations between variables. These associations do not need to be causal in nature.

[iv] The conversion table can be obtained from the author.

[v] The author is in the process of revalidating the scale, a process that will yield standards of desirable performance within each of the functions.

[vi] Whereas this review focused solely on PIMRS studies of instructional leadership, Bridges' (1982) review examined studies across a broad range of topics in educational administration. Thus, the interpretation of findings from his citation analysis was clouded somewhat by the selection criterion. For example, one would not necessarily expect studies in school finance to be cited in studies of principal work activities or school climate. In contrast, the citation

analysis in this review had a built-in ‘control’ since all of the studies were conducted in the same sub-domain of the field.