Principal Evaluation: A Description of Current Practices

by

Leighangela Brady

A dissertation submitted to the faculty of

San Diego State University

In partial fulfillment of the requirements for the degree

Doctor of Educational Leadership

February 29, 2012
The Undersigned Faculty Committee Approves the
Dissertation of
Dr. Leighangela Brady
Principal Evaluation: A Description of Current Practices

Margaret R. Basom, Chair
College of Education, Educational Leadership

Nina Potter
College of Education, Educational Leadership

Sharmila Kraft
Encinitas Union School District

2/29/12
Approval Date
Copyright © 2012

by

Leighangela Brady
DEDICATION

To my son. Continue to dream big and Live Friendly.
ABSTRACT

Strong principal leadership is critical in establishing and maintaining effective elementary schools. However, evaluation methods to measure elementary principal effectiveness still lack in development and consistency. The complexity of the principal role makes it difficult to align evaluation processes to desired research based effective principal behaviors, and, unfortunately, principal evaluation is still a much underdeveloped topic in educational research. States and districts need to critically address principal evaluation practices and create systems that accurately measure principal performance and support principals in their professional growth.

This qualitative study sought to better describe the status of current elementary school principal evaluation procedures and identify how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness. Interviews of 10 elementary school principals representing different school districts within Southern California served as the primary source for data collection. Additionally, principal evaluation instruments, district documents, websites, demographic data, and field notes supported this inquiry. A constant comparative method helped the researcher to organize and analyze data.

The study describes the status of current elementary school principal evaluation procedures in 10 districts. Specifically, the study reports on processes and tools used in principal evaluation systems, the degree to which current principal evaluation practices align across districts, how practices align to research on effective leadership practices, and the degree to which elementary school principals perceived evaluation procedures influenced or supported improvement of their leadership effectiveness.
Findings suggest that policies and practices across districts varied, principals did not feel that current evaluation practices influenced their leadership, and participants had little input into the development of the evaluation tool used by their districts. Data also point to the need for evaluation processes that focus on building trust between principals and supervisors, a desire for greater visibility of evaluators, and a need for ongoing and specific feedback regarding their performance. Participants’ other suggestions for improvement included building relationships and trust, providing more formative feedback, engaging in regular conversations about leadership, visiting schools more often, being consistent in completion of evaluations, and aligning professional development to expectations.

Study results support the need to revise and align current evaluation practices with established research on effective leadership behaviors. Processes and tools need to be consistent, and impact principal behaviors. Principals need to feel valued and have confidence in receiving beneficial feedback about their performance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER 1—OVERVIEW OF THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>Background of the Study</td>
<td>2</td>
</tr>
<tr>
<td>Problem Statement</td>
<td>4</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>4</td>
</tr>
<tr>
<td>Overview of the Methodology</td>
<td>5</td>
</tr>
<tr>
<td>Limitations and Delimitations of the Study</td>
<td>6</td>
</tr>
<tr>
<td>Definitions of Key Terms</td>
<td>7</td>
</tr>
<tr>
<td>Outline of the Research Document</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER 2—LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>Historical Perspective of the School Principal’s Role</td>
<td>12</td>
</tr>
<tr>
<td>Key Leadership Behaviors of Today’s Principals</td>
<td>18</td>
</tr>
<tr>
<td>Leadership Theories</td>
<td>18</td>
</tr>
<tr>
<td>Leadership Responsibilities</td>
<td>21</td>
</tr>
<tr>
<td>Evaluation Overview</td>
<td>27</td>
</tr>
<tr>
<td>Current Principal Evaluation Practices</td>
<td>31</td>
</tr>
<tr>
<td>Home Recipes</td>
<td>32</td>
</tr>
<tr>
<td>Guidelines and Standards</td>
<td>34</td>
</tr>
<tr>
<td>Chapter/Section</td>
<td>Page</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Instruments</td>
<td>85</td>
</tr>
<tr>
<td>Description of tools.</td>
<td>85</td>
</tr>
<tr>
<td>Reliability of instruments.</td>
<td>93</td>
</tr>
<tr>
<td>Perceived Effects on Leadership.</td>
<td>95</td>
</tr>
<tr>
<td>Alignment.</td>
<td>96</td>
</tr>
<tr>
<td>Influence.</td>
<td>102</td>
</tr>
<tr>
<td>Benefits.</td>
<td>105</td>
</tr>
<tr>
<td>Success.</td>
<td>108</td>
</tr>
<tr>
<td>Suggestions for Improvement.</td>
<td>109</td>
</tr>
<tr>
<td>Trust.</td>
<td>110</td>
</tr>
<tr>
<td>Visibility.</td>
<td>113</td>
</tr>
<tr>
<td>Feedback.</td>
<td>115</td>
</tr>
<tr>
<td>Summary.</td>
<td>116</td>
</tr>
<tr>
<td>CHAPTER 5—SUMMARY AND DISCUSSION.</td>
<td>118</td>
</tr>
<tr>
<td>Overview of the Problem.</td>
<td>118</td>
</tr>
<tr>
<td>Methodology.</td>
<td>119</td>
</tr>
<tr>
<td>Summary and Discussion of Findings.</td>
<td>120</td>
</tr>
<tr>
<td>Status of Principal Evaluation Practices.</td>
<td>121</td>
</tr>
<tr>
<td>Perceived Effects on Leadership.</td>
<td>125</td>
</tr>
<tr>
<td>Suggestions for Improvement.</td>
<td>128</td>
</tr>
<tr>
<td>Limitations.</td>
<td>129</td>
</tr>
<tr>
<td>Implications.</td>
<td>130</td>
</tr>
<tr>
<td>Recommendations for Future Research.</td>
<td>133</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.</td>
<td>Participant Screening and Interview Protocol.</td>
<td>72</td>
</tr>
<tr>
<td>Table 2.</td>
<td>Board Policies on Personnel Evaluation.</td>
<td>81</td>
</tr>
<tr>
<td>Table 3.</td>
<td>Principal Evaluation Tool Components.</td>
<td>86</td>
</tr>
<tr>
<td>Table 4.</td>
<td>Data Codes.</td>
<td>87</td>
</tr>
<tr>
<td>Table 5.</td>
<td>Comparison of 21 Leadership Responsibilities and District Evaluation Tools.</td>
<td>99</td>
</tr>
<tr>
<td>Table 6.</td>
<td>Comparison of 21 Leadership Responsibilities and Principal Responses.</td>
<td>101</td>
</tr>
<tr>
<td>Table 7.</td>
<td>Comparison of 21 Leadership Responsibilities, District Evaluation Tools, and Principal Responses.</td>
<td>103</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

PAGE

Figure 1. State API trends by principal. ................................. 104
Figure 2. Principal perceptions of current evaluation practices. .......... 110
ACKNOWLEDGMENTS

When I began this journey, I had no idea the impact it would have on my life. Yet, with the support of some very dedicated people, I was able to weather the storms, while achieving perhaps one of my greatest accomplishments to date.

First, I would like to thank my dissertation committee. Peg Basom was a dedicated chair with a heart of gold. Having studied with her before at San Diego State, she was my draw to come back to the university. Her patience, mentoring, and friendship enriched my experience beyond words. I chose Nina Potter as a committee member when I discovered that she was my match in her passion for the topic of evaluation. Her smile and good humor always reminds me to embrace each experience with a positive attitude and lighthearted spirit. And Sharmila Kraft is a colleague I was destined to meet and share a career. She pushes me to achieve my personal best, and for that I will always be eternally grateful.

Of course, I would be remiss not to acknowledge others who encouraged me from the sidelines. Tim Baird, my superintendent, inspired me to study this topic with his unique approach to principal evaluation and reminded me to look at my research with passion, purpose, and power. Lil Volkening, my office manager, has finally earned her little “e” to be placed after her name for supporting yet a third principal through the doctoral process. I cannot thank her enough for listening to me so many times with genuine interest and curiosity.

And most of all, I owe gratitude to my family. I love my husband, Chris, down to my very core. Through this process, he never wavered in his support and held our lives together through the toughest of times. He is an amazing blessing to me. And with this
wonderful man, I share a beautiful son, Noah. Noah recognizes what it means to be a devoted student and cheered me on along the way. May I one day return the favor. Finally, my mother Sharyn taught me the importance and value of education and how it is the one thing in life that can never be taken away. Though I have achieved in learning, I hope one day to be as wise as she. To my creator—welcome back into my life.
CHAPTER 1—OVERVIEW OF THE STUDY

Demands for continuous improvement have inundated American schools for decades. As leaders, principals are at the center of these efforts, and hold the greatest accountability for student success. Therefore, strong principal leadership is an essential component in developing and maintaining effective schools (Edmonds, 1980; Elmore, 2000; Ginsberg & Thompson, 1992; Lezotte, 1992; Mace-Matluck, 1987). Yet to achieve this, principals need to be aware of leadership behaviors that will maximize their effectiveness.

A substantial body of research has identified behaviors of successful principals and has determined which leadership practices have the greatest impact on school success (Cotton, 2003; Marzano, Waters, & McNulty, 2005; Murphy, Elliott, Goldring, & Porter, 2010; Stiggins & Duke, 2008). Standards for school principals, such as those developed by the Interstate School Leaders Licensure Consortium (ISLLC, 2008), indicated that a school administrator is responsible for ensuring student success by facilitating a shared vision, promoting a school culture, managing day-to-day operations, building relationships, acting in a fair and ethical manner, and influencing a larger cultural context. A meta-analysis by Marzano et al. (2005) further identified distinct principal responsibilities yielding the greatest effect on student academic achievement. However, researchers have paid little attention to studying effective ways to measure or evaluate these responsibilities.

Careful evaluation of principals can lead to effective principals and effective schools (Thomas, Holdaway, & Ward, 2000). Unfortunately, studies indicate that principal evaluation systems and processes are a neglected area in educational research.
(Albanese, 2003; Davis, Kearney, Sanders, Thomas, & Leon, 2011; Ginsberg & Berry, 1990; Murphy, Hallinger, & Peterson, 1985; Studebaker, 2000). In addition, there is lack of implementation, usefulness, and reported effects of current evaluation processes (Gaziel, 2008; Goldring, Cravens, et al., 2009; Marcoux, Brown, Irby, & Lara-Alecio, 2003; Oyinlade, 2006). Furthermore, practices are often inconsistent between states, districts, and schools (Albanese, 2003; Studebaker, 2000). Therefore, prudent attention should be paid to identifying and examining effective principal evaluation systems.

**Background of the Study**

Almost all professions participate in some type of formal evaluation practices. Evaluation processes can provide employers with valuable information about employee performance, and provide employees with feedback for improvement (Longnecker & Nykodym, 1996). Due to the dynamic nature of professions and job requirements, it is important to regularly evaluate employees to maintain high quality performance (McGaghie, 1991). Yet, despite the long history of performance evaluations in the workplace, they continue to remain complicated and controversial (Longnecker & Nykodym, 1996; Roberts, 2003).

These findings are amplified in the research surrounding principal evaluation. An in-depth review of the literature concluded that, over time, principal evaluation practices have remained varied and inconsistent with sparse research to indicate the effectiveness of individual systems (Glasman, 1992; Glasman & Heck, 1992; Glasman & Martens, 1993; Heck & Glasman, 1993). Increasing requirements of the principal role make it even more difficult to standardize assessment practices (Glasman, 1992; Heck & Glasman, 1993; Heck & Marcoulides, 1992).
Many scholars have identified a need for improved principal evaluation methods and emphasized the importance of aligning these methods and research-based criteria for effective principal leadership (Albanese, 2003; Ginsberg & Berry, 1990; Goldring, Cravens, et al., 2009; R. Johnson, 2005; Murphy et al., 1985; Studebaker, 2000). Yet, only a few studies have been conducted to further explore effective principal evaluation practices (Murphy et al., 1985; Stufflebeam & Nevo, 1993; Toler, 2006). Moreover, numerous methods found in the literature outlined systematic procedures but did not offer research to support the effectiveness of the practices. Leadership standards attempt to outline principal responsibilities and provide guidelines for evaluation, but tools remain underdeveloped or inconsistently applied (Babo, 2009; Brooks & Voss, 2008; Derrington & Sharratt, 2008).

Overall, researchers agree on the necessary responsibilities of effective principals, though many principal evaluation systems are summative in nature and do not provide principals with the opportunity to improve leadership practices as promoted in the literature (Albanese, 2003; Conca, 2008). Involving stakeholders in the process of creating evaluation tools could provide increased benefits such as better employee attitude, increased understanding of process, and perceived reduction of bias/favoritism; however, this is not a common practice in current principal evaluation methods (Roberts, 2003). Roberts (2003) found “when employees possess a meaningful role in the appraisal process, employee acceptance and satisfaction with the appraisal process is strongly enhanced” (p. 95). Moreover, two decades ago, Ginsberg and Berry (1990) reported that a lack of attention to principal evaluation methods resulted in the poor quality of principal evaluation systems. This same finding exists in today’s literature.
The role of an elementary school principal is far-reaching and complex (Catano & Stronge, 2006; Marzano et al., 2005). Elementary school principals are feeling the pressures of a position without bounds (Archer, 2002). Many elementary school principals have limited administrative support: secretarial, special education, discipline, social services, before and after school care, and other like responsibilities. They are being spread thin and are “expected to be the top administrators, budget directors, and human-resource managers in their buildings” (Archer, 2002, p. 2). Yet, in spite of the additional responsibilities and the continually evolving role of elementary school principals, practices for evaluating these leaders have remained stagnant. This study will focus on the present status and the perceived impact of evaluation practices as they apply to elementary school principals.

Problem Statement

Elementary principals are held accountable for guaranteeing student learning (Catano & Stronge, 2006; Maxwell, 2008). Yet, evaluation methods to measure this accountability still lack in development and consistency (Brooks & Voss, 2008; Conca, 2008). In addition, the intricacy of the principal role makes it difficult to align evaluation processes to desired principal behaviors (Glasman, 1992; Heck & Glasman, 1993; Heck & Marcoulides, 1992). This study investigated current principal evaluation systems and the ways elementary school principals perceived evaluation procedures support improvement of their leadership effectiveness.

Purpose of the Study

This study explored current elementary principal evaluation practices by examining methods of principal evaluation in 10 different school districts. The study
aimed to better describe the status of current elementary school principal evaluation procedures, and identify how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness. To help determine this, the following research questions were investigated:

1. What do processes and/or practices used in principal evaluation systems entail?
2. How consistent are principal evaluation practices across districts?
3. Do current principal evaluation practices align to research on effective leadership practices?
4. How do principals perceive their evaluation systems influence their leadership behaviors?

This topic was clearly an area in need of studying as substantiated by the lack of empirical studies in the literature. Results of this study add both theoretical and empirical research to inform superintendents, districts, and other administration responsible for supervising and evaluating principals of practices that principals perceive most support leadership effectiveness.

**Overview of the Methodology**

The study began with an in-depth look at the literature. Areas of research and scholarship included an historical look at the principal role, key leadership behaviors of today’s principals, an overview of evaluation, current principal evaluation practices, and factors influencing principal evaluation. An exhaustive literature review demonstrated a lack of empirical research describing effective principal evaluation practices and their influence on leadership behaviors.
This study used a qualitative approach to better understand the current nature of principal evaluation methods. “The purpose of qualitative research is to describe and understand social phenomena in terms of the meaning people bring to them” (Boeije, 2010, p. 11). Due to a paucity of theoretical research around principal evaluation methods, this was an area of scholarship that warranted further exploration.

Using a qualitative descriptive research design, this study produced descriptive data that generated findings potentially useful to principal leadership. Ten elementary school principals participated in this study, each representing a different school district within the region of Southern California. Data were collected primarily through interviews. Principal evaluation instruments, documents, websites, demographic data, and field notes served as secondary sources of data. Data were organized, analyzed, and categorized into themes. Ethics and trustworthiness of the data were carefully considered. Data analysis was used to draw conclusions about principal evaluation practices compared across districts.

**Limitations and Delimitations of the Study**

Glatthorn and Joyner (2005) denoted delimitations as boundaries of a study and ways in which a study may lack generalizability. In this study, both the small sample size and localized geographic region could be considered delimitations. Other limiting factors that potentially influenced study data are principals who may have misrepresented information and possible internal bias of the researcher.

Since the study considered perceptions of elementary school principals regarding current evaluation systems of their districts, the perspectives of the interviewees could be
considered an additional delimitation. Perspectives may have been exaggerated or
downplayed which could potentially influence study findings.

Finding 10 qualified willing participants for the study proved to be a challenge. Out of the 41 school districts in San Diego County, 15 districts were either too small or remote to gain access to principals who met the criteria of the study. Six additional districts were high school only districts and therefore could not be used for the study. Furthermore, finding principals that were able and willing to set aside time from their busy schedules to participate in this study was difficult. Subject identification was made through personal and professional connections. Emails and/or phone calls were used to identify an appropriate and willing participant pool. All attempts were made not to pressure any of the principals contacted to participate in this study.

Finally, to limit internal bias, the literature review provided a framework in the development of the interview questions. Principals not participating in the study were asked to provide feedback regarding question clarity after interview questions were developed. This feedback was used to revise the interview questions only, and no input from these principals was added to the data corpus of the study.

During the study, the researcher maintained confidentiality by using a password-protected computer and a locked filing cabinet. All identifying documents were permanently deleted at the conclusion of the study.

**Definitions of Key Terms**

The following definitions are offered to provide an understanding and the context of key terms used in this study:
Principal: The term principal refers to the multifaceted role of the building administrator in today’s schools (Catano & Stronge, 2006; Marzano et al., 2005). This study will focus primarily on principals at the elementary level.

Evaluate: “To determine the significance, worth, or condition of usually by careful appraisal and study” (“Evaluate,” 2010).


Rater: A supervisor completing an employee evaluation or appraisal (Arvey & Murphy, 1998; Balkin, 1994; Curtis, Harvey, & Ravden, 2005; Feldman, 1981).

Meta-analysis: A variety of techniques used to quantitatively synthesize a vast amount of research (Marzano et al., 2005).

Standards: Standards in this study will refer to the knowledge and skills required for successful school leadership. Standards provide a foundation to effectively evaluate the performance of principals (Derrington & Sharratt, 2008).


Summative evaluation: The goal of summative evaluation is to precisely evaluate performance. Summative evaluations are often used to facilitate decisions over compensation or tenure (Peterson, 1991).

Outline of the Research Document

This dissertation consists of five chapters. Followed by a brief introduction, Chapter 1 outlines a general background of the study, the problem statement, the purpose
of the study, an overview of the proposed methodology, delimitations of the study, and definitions of key terms.

Chapter 2 expounds on the empirical and theoretical literature around principal evaluation and leadership responsibilities that impact student academic achievement. The literature review is organized into five sections: a history of the principal role, key leadership behaviors of today’s principals, an overview of evaluation, current principal evaluation practices, and factors influencing principal evaluation. Embedded throughout Chapter 2, an analysis of the research summarizes gaps relating to the need for this study.

The overall methodology and research design are described in Chapter 3. Following the outline presented in Glatthorn and Joyner (2005), this chapter clearly describes the research methodology, context, participants, instruments and materials used, procedures followed, and proposed data analysis procedures.

Data gathered and results of the study are reported in Chapter 4. Results of the findings in terms of the posed research questions is summarized and discussed in Chapter 5. Conclusions and implications are presented as well as recommendations for future research.
CHAPTER 2—LITERATURE REVIEW

The role of the principal has continuously evolved since the birth of the principalship over a century ago. Throughout history, principals have served as principal teachers, instructional managers, instructional leaders, and transformational leaders (Glasman & Heck, 1992). Over the past 35 years, researchers have searched to identify and isolate leadership roles and practices that have the biggest impact on school effectiveness. In a meta-analysis of over 69 studies examining effective school leadership in the United States, Marzano et al. (2005) synthesized the six most popular theories on school leadership and derived 21 leadership responsibilities for principals.

Marzano et al. (2005) made statistically based generalizations emerging from studies of 2,802 schools, approximately 1.4 million students, and 14,000 teachers. The generalizations led to specific responsibilities that help principals better define their role in schools. The meta-analysis quantitatively synthesized an immense amount of research on specific leadership behaviors for school administrators, and added empirical evidence to existing theories. Marzano et al. discovered that principals had a greater impact on student achievement when engaging in these explicit actions.

Knowing what principals need to do, however, is only half of the challenge. To be effective in their role, principals must receive ongoing feedback regarding their performance (Thomas et al., 2000). In any professional role, feedback and monitoring are used as a basis for improvement of professional performance. This most often occurs through assessment and evaluation measures. Evaluation methods for school leaders are only valuable to the degree they assess principal effectiveness, successfully address leadership weaknesses, and lead to improved principal performance (Ginsberg & Berry,
Yet, current principal evaluation methods are often summative in nature, outdated, or fail to reflect the complexities of the job (Albanese, 2003; Ginsberg & Berry, 1990; Murphy et al., 1985; Studebaker, 2000).

To further explore the issue of assessment and evaluation of principals, this literature review explores five primary areas of the research and scholarship including a history of the principal role, key leadership behaviors of today’s principals, an overview of evaluation, current principal evaluation practices, and factors influencing principal evaluation. Beginning with an historical view of the principalship, the review considered the principal’s role in school effectiveness and the implications of these responsibilities for principal evaluation. Brookover and Lezotte (1979) and Edmonds (1980) led the way for research around the effects schools have on student achievement. Specifically, they helped define the role of the principal in school effectiveness. Elmore (2000) described imperatives for public school leadership and the restructuring of this leadership. Years later, studies indicated a continued need for improvement of principal evaluation methods (Albanese, 2003; Ginsberg & Berry, 1990; Murphy et al., 1985; Studebaker, 2000). In addition, this review of scholarship and research analyzed the current research on principal evaluation practices and explored the degree to which these methods addressed, or failed to address, important leadership knowledge, skills, and dispositions (Goldring, Cravens, et al., 2009; Green, 2004; Marcoux et al., 2003; Oyinlade, 2006).

Effective performance-based evaluation methods have potential to provide principals opportunities to reflect upon and adjust their own practices (Brooks & Voss, 2008; Derrington & Sharratt, 2008; Maxwell, 2008; Oyinlade, 2006). Due to ongoing discrepancies between defined principal evaluation measures and current practices, this
study sought to determine if current evaluation practices influenced principal leadership behaviors. Studying various methods may unveil areas of strength, while uncovering potential ways to maximize the effects of principal evaluation.

**Historical Perspective of the School Principal’s Role**

The system of public education in the United States continually evolves in response to societal needs, economic demands, and governmental mandates, to name a few. However, despite over a century of school reform, the overall structure and function of schooling have changed very little (Cuban, 1988; Kafka, 2009; Rousmaniere, 2007; Tyack & Cuban, 1995). Classes are still organized by age-level groups, grades and credits are assigned as evidence of learning, and principals are charged with supervising teachers to monitor educational pedagogy (Tyack & Cuban, 1995).

Tyack and Cuban (1995) outlined the organization and reform of American schools dating back to the mid 1800s. As American schools expanded in population during the 19th century, graded classrooms resulted. Schools became far larger than the former single-room schoolhouses. Administrative tasks arose, including staff supervision, student discipline, community relations, and crisis management (Rousmaniere, 2007). Because district administrators tended to govern remotely, they assigned master teachers to oversee daily administration of local schools (Kafka, 2009; Rousmaniere, 2007). Thus, the principal position was born.

The first principals, called master or principal teachers, had both teaching and administrative duties. As years passed, principal responsibilities increased and teaching time decreased.
By the 1920s, the modern school principalship had been established and looked markedly similar to the position today: Principals had bureaucratic, managerial, instructional, and community responsibilities. They were expected to lead and instruct teachers, to monitor students, to communicate with the district, and to work with parents and members of the wider community. Moreover, they were seen as pivotal figures in any school reform effort. (Kafka, 2009, p. 323)

Eventually, principals were relieved of teaching duties altogether, but the professional status of principals remained unclear. Operating somewhere in middle management between the district office and the classroom, principals were considered neither administration nor teachers (Rousmaniere, 2007).

Although there is sparse research outlining the historical role of the principal (Kafka, 2009; Rousmaniere, 2007), the principal role continued to evolve along with the many educational reforms of the 20th century. By the late 1960s, the public expressed a growing dissatisfaction with American schools. Some of the discontent can be attributed to The Coleman Report, a notable study on student performance (Coleman et al., 1966). Study researchers believed that a correlation existed between the input and output of schools. This assumption implied that if resources were similar in schools, student performance would also be similar (Coleman et al., 1966).

Based on nationally conducted surveys, The Coleman Report (Coleman et al., 1966) concluded that most students in the United States, spanning across races and demographics, had similar access to resources in schools. However, a gap existed in student performance across these groups, despite the same input. Therefore, the results were interpreted to suggest that school performance must be directly tied to factors
outside the school rather than within the school’s control (Coleman et al., 1966). This report was a major factor that led to the creation of compensatory education programs, such as Title I and the Elementary Secondary Education Act, to adequately address the needs of disadvantaged students (Lezotte, n.d.).

Of course, the claim that schools had little effect on the achievement of students sparked a reaction from scholars that led to many studies. These studies acted as a catalyst to the Effective Schools Movement. Researchers of this movement not only set out to refute the findings of The Coleman Report, but to find examples of good schools and study their characteristics. Although new studies agreed with the general finding that there are easily measurable differences among schools, the implications of those differences were quite different (Cuban, 1998; Purkey & Smith 1983).

Researchers explored the impact of schools on student performance and found that schools needed to better address student performance inequities and that effective schools shared common characteristics (Brookover & Lezotte, 1979; Edmonds, 1980; Mace-Matluck, 1987). The term effective schools is pervasive throughout the literature; however, a single definition does not exist for what makes a school effective. Many researchers have identified common characteristics of effective schools including strong instructional leadership of the principal, high expectations of the staff for student learning, clear goals and a focused school mission, a safe and orderly environment, staff training opportunities, and an established system for monitoring student progress (Brookover & Lezotte, 1979; Edmonds, 1980; Mace-Matluck, 1987).

Specifically, the leadership role of the school principal emerged as significant in this movement. Researchers concluded that strong instructional leadership of the
principal was a main characteristic of effective schools (Brookover & Lezotte, 1979; Edmonds, 1980; Mace-Matluck, 1987). Of the many case studies cited in Mace-Matluck (1987), Weber (1971) was noteworthy about the effects of strong leadership. In this study, 95 schools were identified and four schools chosen as successful urban schools nominated for intensive study. All four chosen schools demonstrated strong leadership as success factors in student achievement. In three of the four schools, the principal represented the strong leadership referenced (Weber, 1971).

Brookover and Lezotte (1979) added to this research finding that teachers’ and principals’ attitudes toward student capabilities, whether high or low, appeared to be strongly related to student performance. In an intensive study of 10 schools, Brookover and Lezotte compared schools with inclining and declining achievement scores. Data were collected through interviews and questionnaires. Of the 10 major findings, the most significant was that teachers and administrators of the improving schools had higher opinions of their students’ abilities than those in lower achieving schools (Brookover & Lezotte, 1979).

In 1983, The National Commission on Excellence in Education (NCEE) released a letter to the American people entitled *A Nation At Risk*. This publication begins with the following statements:

All, regardless of race or class or economic status, are entitled to a fair chance and to the tools for developing their individual powers of mind and spirit to the utmost. This promise means that all children by virtue of their own efforts, competently guided, can hope to attain the mature and informed judgement needed to secure gainful employment, and to manage their own lives, thereby
serving not only their own interests but also the progress of society itself. (NCEE, 1983, p. 1)

It went on to blame the decline in educational performance on inadequacies of the educational process and advocated for school reform. It called for a commitment to excellence for individuals, schools, and society. It also stated that the primary goal of education should be to have high, rigorous academic standards for all.

Although *A Nation at Risk* (NCEE, 1983) did not address principals specifically, it expressed “that the essential raw materials needed to reform our educational system are waiting to be mobilized through effective leadership” (p. 7). Scholars, policy makers, and the public at large have agreed that principals are key figures in schools, charged with ensuring improvement (Ginsberg & Thompson, 1992; Glasman & Heck, 1992; Murphy et al., 1985; Rallis & Goldring, 1993). Therefore, principals became increasingly accountable for educational inadequacies.

Under increased scrutiny from the American public, the principal’s role shifted, obliging principals to be effective instructional leaders. Manasse (1985) reviewed research on effective schools and effective principals. She prepared a paper for the National Institute of Education identifying focus areas for local districts and agencies to support effective principal behaviors. Among suggestions, Manasse recommended training principals for the complexities of their work, developing pedagogical skills to ensure effective instructional leaders, and providing ongoing feedback through performance appraisal systems.

Almost two decades later, the need for school improvement remained. In 2001, the U.S. Department of Education demanded school accountability by proposing and
implementing the No Child Left Behind Act (NCLB). This legislation significantly raised expectations for states and schools to ensure educational reform. No Child Left Behind addressed student inadequacies across varying demographics, and instituted a plan to guarantee the success of every child. States began monitoring schools through the implementation of state standards and increasing progress on state assessments.

Principals were, and continue to be, at the center of NCLB expectations. With NCLB, the role of the principal became even more complex and demanding. Crow (2006) suggested that the scrutiny of the public and the increased expectations for student learning, “has added to the complexity of the principal’s job, requiring principals to be entrepreneurial, to be more focused on student outcomes and instructional processes, and to be more connected with their communities” (p. 316).

In a publication entitled Beyond NCLB: Fulfilling the Promise to Our Nation’s Children, Thompson and Barnes (2007) reinforced the idea that “research has consistently shown that high-performing schools have principals who are effective leaders” (p. 32). In their summary of recommendations, they dedicated an entire section to ensuring teacher and principal effectiveness. Recommendations regarding principals included establishing a definition for a highly effective principal, improving professional development for school leaders, and accurately identifying the needs of principals. Although Thompson and Barnes suggested that the need for high performing principals is based on research, their recommendations are not empirically supported. The very nature of their recommendations further illustrate that we still do not fully understand the complexities and needs of the principal role nor the implications for evaluation.
Key Leadership Behaviors of Today’s Principals

There is a body of research that has attempted to better understand the complexities and needs of the role of the school principal. The quality of principal leadership constitutes a primary indicator of school effectiveness (Elmore, 2000; Ginsberg & Thompson, 1992; Lezotte, 1992). Since school effectiveness is most often measured according to student academic performance, leadership effectiveness is likewise determined in relation to the school’s core functions, teaching and learning (Cotton, 2003; J. Johnson, 2008; Lezotte, 1992; Orr, Berg, Shore, & Meier, 2008; Valentine, 1987). In order to be effective, leaders must be both knowledgeable of, and involved with, current curriculum and assessment practices. They must also know how to review student data and understand how to link assessment results to instructional improvement (Leithwood, Jantzi, & Jantzi, 2008; Marzano et al., 2005; Stiggins & Duke, 2008).

With the reform of schools and an emphasis on accountability for student performance, principals have had to reform as well. Many studies have been conducted over the past 35 years to try and determine what exactly the role of the school principal is. Researchers have searched to identify which leadership practices have the biggest impact on school effectiveness. One of the most noted contributions is the meta-analysis by Marzano et al. (2005).

Leadership Theories

The six prominent leadership theories derived from the meta-analysis (Marzano et al., 2005; Stiggins & Duke, 2008) consisted of transformational leadership, transactional leadership, total quality management, servant leadership, situational leadership, and instructional leadership. Transformational and transactional leadership
are terms used in both business and education. These types of leaders are skilled at influencing others to act toward the achievement of common goals. Although similar, Marzano et al. (2005) emphasized a slight difference between the two. While transactional leadership is the act of replacing one thing with another, transformational leadership is more about change itself.

Marzano et al. (2005) summarized transactional leadership as three basic forms: management-by-exception-passive, management-by-exception-active, and constructive transactional leadership. Management-by-exception-passive leaders believe that it is their job to maintain the status quo and only influence leadership behaviors if problems arise. In management-by-exception-active, leaders believe that it is their job to aggressively manage set standards, behavior, and any issues that arise. It is not their job, however, to take risks or initiative. Constructive transactional leaders set goals and desired outcomes, monitor and provide feedback, and recognize accomplishments (Marzano et al., 2005).

Marzano et al. described transformational leadership as encompassing factors of individual consideration, intellectual stimulation, inspirational motivation, and idealized influence. Transformational leaders use these four factors to enable others to act. Specifically, principals exhibit individual consideration when they provide individual attention to the teachers and other staff members. They employ intellectual stimulation by guiding teachers to find new approaches to existing problems. Inspirational motivation is observed when transformational leaders stimulate others to meet high expectations. Lastly, school principals exhibit idealized influence by modeling the desired behavior of teachers.
A third leadership theory is total quality management. Marzano et al. (2005) gave credit to Edward Deming (1986) as founding and providing the framework for total quality management (TQM). First introduced in business models, TQM also had strong implications for school leadership. The 14 principles of TQM can be synthesized into five factors that most affect principals: change agency, teamwork, continuous improvement, trust building, and eradication of short-term goals. Change agency refers to a principal’s ability to identify and affect change in a school, teamwork to establish and support effective teams, continuous improvement to enlist employees in incremental growth toward active goals, trust building to motivate and support employees, and eradication of short-term goals to design a long-term perspective to goal design and implementation (Marzano et al., 2005).

Marzano et al. (2005) also recapitulated servant leadership. Servant leaders are motivated by their desires to help others. The servant leader nurtures and supports, does not supervise or manage. They are found in the center of the organization, as opposed to the top of the hierarchy, and often viewed as connected to all facets of the organization. Marzano et al. described servant leaders as being able to understand personal needs, heal wounds caused by conflict, act as stewards of resources, develop the skills of employees, and listen effectively.

Marzano et al. (2005) recognized the work of Paul Hersey and Kenneth Blanchard (Hersey, Blanchard, & Johnson, 1996) when outlining the fifth leadership theory of situational leadership. Because there is not one best way to influence others, leaders use situational leadership to adjust to the needs and abilities of employees (Hersey, 1984). Telling, participating, selling, and delegating are four styles associated with situational
leadership. When an individual or group needs direction and is low in ability and willingness, the situational leader will need to use a “telling” or directed approach (Hersey et al., 1996, p. 201). Leaders engage in the “selling” or persuasive approach when employees are still unable to perform a task, but they are trying (Hersey et al., 1996, p. 202). If employees are able, but lack confidence to perform a task, the leader will use a “participating” or guiding approach (Hersey et al., 1996, p. 203). However, if employees are willing, able, and confident situational leaders can use a “delegating” approach (Hersey et al., 1996, p. 205).

Finally, Marzano et al. (2005) cited instructional leadership as “perhaps the most popular theme in educational leadership over the last two decades” (p. 18). Several definitions exist for instructional leadership encompassing roles such as facilitator, coach, and communicator; and functions such as defining the vision and mission, organizing curriculum and instruction, and promoting a positive culture. All definitions of instructional leadership view the principal as a resource provider and a supporter of instruction.

**Leadership Responsibilities**

Marzano et al. (2005) also identified 21 specific leadership responsibilities that have the greatest impact on student achievement. These responsibilities emerged out of quantitative meta-analyses that produced significant correlations between leadership behavior and the average academic achievement of students. The following outlines the 21 responsibilities by Marzano et al. that help principals better define their role in schools. The analyses produced an average correlation of .25, with correlations ranging from .18-.33. These correlations represented the relationship between each leadership
behavior and student achievement. Simply put, principals had approximately a 6-11% greater impact on student achievement when engaging in the following actions (correlations represented in parentheses):

1. Affirmation (.19): The principal systematically addresses both positive and negative performance of the students, staff, and school. Affirmation is the accountability of the principal for delivering the results and progress of school goals.

2. Change Agent (.25): The principal challenges the status quo. As a change agent, the principal is empowered to take risks and lead initiatives despite uncertain outcomes. Change agents are characterized by always considering better alternatives for everyday practices.

3. Contingent Rewards (.24): The principal recognizes and rewards staff accomplishments. Recognitions call attention to what is valued. The principal must be proactive in finding opportunities to provide contingent rewards of various staff performance.

4. Communication (.23): The principal establishes strong, effective communication between students and staff. The principal is accessible to teachers and staff and maintains open lines of communication.

5. Culture (.25): The principal builds a positive culture to influence teachers and students. A positive culture indirectly affects student achievement, and the principal is at the center of these efforts. A principal creates a sense of community by promoting cohesion, cooperation, and well-being; and developing a shared vision and purpose.
6. Discipline (.27): The principal protects teachers’ instructional time from internal and external distractions. Discipline refers to the act of buffering, or creating structures around teachers and classrooms so that their focus remains solely on teaching and learning.

7. Flexibility (.28): The principal adapts to various styles and situations. A flexible principal encourages opposing opinions and is comfortable with initiating decisions and change.

8. Focus (.24): The principal establishes clear goals for curriculum, instruction, and assessment; and promotes high expectations to meet them. An effective principal focuses efforts on maintaining the advancement of the school goals.

9. Ideals and Beliefs (.22): The principal shapes the conditions of a school through strong ideals and beliefs. The principal effects change by establishing, sharing, and modeling these beliefs.

10. Input (.25): The principal involves staff in important decision-making. A strong principal effectively listens and provides opportunity for input from all staff.

11. Intellectual Stimulation (.24): The principal keeps the teachers informed of the most current and relevant research on effective schooling. Systematic discussions about this research keep a principal’s staff intellectually stimulated.

12. Involvement in Curriculum, Instruction, and Assessment (.20): The principal becomes personally involved with designing curricular activities, addressing instructional issues, and designing assessment opportunities. This direct
involvement is at the core of instructional leadership. Principals accomplish this through regularly meeting with teachers, visiting classrooms, and providing input.

13. Knowledge of Curriculum, Instruction, and Assessment (.25): The principal is not only involved, but is also knowledgeable of curriculum, instruction, and assessment. The principal acquires and possesses extensive knowledge about effective instructional, curricular, and assessment issues by staying abreast of current best practices.

14. Monitoring and Evaluating (.27): The principal monitors and evaluates school practices in terms of the impact on student achievement. The principal provides feedback on improving these practices.

15. Optimizer (.20): The principal inspires through a positive attitude. As an optimizer, the principal is the driving force behind major initiatives and innovation.

16. Order (.25): The principal creates a smooth running school by establishing clear boundaries and routines. Order is defined by structures, rules and procedures for staff and students.

17. Outreach (.27): The principal communicates to all stakeholders such as parents, the district, and the community. The principal develops partnerships and is an advocate for the school.

18. Relationships (.18): The principal demonstrates an awareness of the personal lives of the staff. The school leader makes an effort to maintain personal
relationships with teachers and acknowledge significant events in the lives of staff members.

19. Resources (.25): The principal proactively ensures that teachers have the necessary materials, equipment, and training that they need.

20. Situational Awareness (.33): The principal is aware of all aspects of the school. With effective situational awareness, a principal can accurately predict what could go wrong or what issues may arise.

21. Visibility (.20): The principal interacts regularly with teachers, students, and parents. To establish effective visibility, the principal makes frequent visits to classrooms, attends school functions, and initiates regular contact with students. (Marzano et al., 2005, pp. 41-61)

Marzano et al.’s (2005) 21 leadership responsibilities are similar to the 25 categories of principal behaviors identified by Kathleen Cotton (2003). Cotton based her findings on a narrative review of the literature surrounding principal practices that positively affected student and teacher attitudes and behavior. Cotton and Marzano et al. overlap in all areas. Some of the responsibilities identified by Marzano et al., such as optimizer; focus; and monitoring/evaluating, align with several of the practices described by Cotton. The main difference is that, unlike Marzano et al., Cotton did not quantitatively measure the effects of principal leadership, but rather looked for trends and patterns in the literature using a narrative approach. Though Marzano et al. and Cotton took different approaches to quantifying the research, both analyses agree, “leadership is vital to the effectiveness of a school” (Marzano et al., 2005, p. 4).
Nevertheless, despite the large volumes of theoretical and empirical research on the effects of principal leadership, some scholars continue to question the validity and reliability of the tools used to collect the data cited in the conclusions of the previously mentioned studies. In particular, self-report surveys are one type of tool commonly used to collect data on principal leadership practices (Hallinger & Heck, 1996). These surveys are most often administered in a way that asks principals to report their practices from memory. Unfortunately, respondents may not accurately recall events and, in effect, may inadvertently misrepresent the emphasis of their leadership practices.

A closer look by Camburn, Huff, Goldring, and May (2010) found self-report surveys to be unreliable. Camburn et al. (2010) designed a study to compare survey measures of principal leadership practice with comparable daily log measures. Camburn et al.’s study found that evidence from daily logs gave more adequate data of the time principals spent on various leadership activities compared to asking them to report from memory on a year-end survey. Moreover, “researchers using the same log data used here found that principals’ emphasis on different leadership tasks can vary quite a bit from day to day” (Camburn et al., 2010, p. 320).

Like Hallinger and Heck (1996) and Camburn et al. (2010), many scholars believe that more valid studies are needed to better understand how principals truly allocate their time across the various leadership practices. However, most researchers would agree that there are at least strong indirect effects of principal leadership on school achievement (Camburn et al., 2010; Cotton, 2003; Hallinger & Heck, 1996; Marzano et al., 2005; Mattson Almanzán, Sanders, & Kearney, 2011). Therefore, it is vital that principal leadership be effectively evaluated.
Evaluation Overview

According to William McGaghie (1991), the western industrial revolution redefined the term profession. He described this term as “hundreds of occupational groups—most with university ties, elaborate credentialing procedures, and some licensing restrictions” (p. 3). Many professions require a certain type of credential before a person is allowed to enter the field. Professions are continually evolving, as are the employees of these professions. Consequently, regularly evaluating the competence of professionals is important to ensure quality, effectiveness, and safety.

Some form of performance evaluation exists in all professions. The term appraisal is often used interchangeably with the term evaluation. The Merriam-Webster Online Dictionary defines evaluate as “to determine the significance, worth, or condition of usually by careful appraisal and study” (“Evaluate,” 2010) and appraisal as “a valuation of property by the estimate of an authorized person” (“Appraisal,” 2010). More specifically, evaluation can be described as a measure of an employee’s value to a company based on their work behavior and performance (Arvey & Murphy, 1998). Jack Feldman (1981) defined performance appraisal as “the outcome of a dual process system of evaluation and decision making” (p. 127), arguing that there is a cognitive process beyond simple evaluation. For the purposes of this literature review, the terms evaluation and performance appraisal will be used synonymously.

Evaluation was designed to measure results, identify strengths and barriers, and make recommendations for improvement (Guerra-Lopez & Leigh, 2009). To achieve this, performance evaluations often include goal setting, coaching, tracking progress toward desired goals, judgments, feedback, and taking corrective action when needed.
A supervisor is usually charged with completing these evaluations; however, there is a body of research that also promotes self-assessment as a form of evaluation. Campbell and Lee (1988) found that, when used correctly, self-assessment augments performance evaluation. However, they also determined that self-assessments should not be used as the only form of accountability or measure of one’s own effectiveness.

Evaluation and performance appraisal research dates back over 50 years (Arvey & Murphy, 1998; Curtis et al., 2005), though practices of evaluating the competence of professional personnel have existed for centuries (McGaghie, 1991). Between 1950 and 1980, research centered mainly on improving evaluation tools (Arvey & Murphy, 1998). In 1978, Landy, Barnes, and Murphy stated that many personnel decisions were made based on performance evaluations. Moreover, evaluation success depended on an employee’s confidence in the evaluation system. McGaghie (1991) stated five reasons for employee dissatisfaction with evaluation procedures including the narrow focus of evaluations, a bias toward assessment of acquired knowledge, little attention devoted to the assessment of practical skills, sparse attention paid to professional or personal qualities, and measurement and validity issues.

Effectively designed formal performance appraisal can provide employers with many benefits, yet, despite the long history of performance evaluations in the workplace, they remain complicated and controversial (Longnecker & Nykodym, 1996; Roberts, 2003). Several issues exist in performance appraisals. Bias, such as rater accountability and personal bias of the supervisor, is an inherent flaw in evaluation with a single evaluator (Arvey & Murphy, 1998; Balkin, 1994; Curtis et al., 2005; Feldman, 1981).
Multiple raters could mediate such biases; however, they may cause a different dilemma of managing disagreements between raters (Balkin, 1994). Other flaws include inflated appraisals, lack of rater training, lack of rater time to complete appraisals, and absence of performance documentation (Roberts, 1998).

Around 1980, research shifted, from a focus on evaluation tools, to a focus on the raters themselves and the judgments they made on employee performance (Arvey & Murphy, 1998). Arvey and Murphy (1998) wrote that studies attempted to develop “a better understanding of the way raters form impressions and judgments of their subordinates’ performance” (p. 143). In a quantitative study, Landy et al. (1978) sought to identify elements that might account for an individual’s perception of fairness and accuracy of performance evaluations. Landy et al. mailed 950 questionnaires, consisting of 12 items, to professionals in a large manufacturing company. Seven hundred eleven people from areas throughout the United States answered the questionnaire within the 2-week time frame, constituting a 74% response rate. Using a stepwise multiple regression analysis, Landy et al. found that evaluations were perceived as fair and accurate when supervisors were familiar with performance, evaluated often, set goals, and provided feedback. In this study, the evaluator was more important to the evaluation than the tool being used.

In 1996, Longnecker and Nykodym researched the effectiveness of formal appraisal systems in the public sector. The case study aimed to encourage organizations to assess their own appraisal systems. Longnecker and Nykodym surveyed 357 professional members of a large organization in the Midwest United States of which 254 people responded: 77.1% of managers, 69.9% of subordinates, and 71% overall. The
organization in this study had a performance appraisal approach typical in many public sector organizations including a professional design, rater training, and annual reviews of employees. The study examined the perceptions of management and subordinates and sought to find ways to improve the appraisal process.

Longnecker and Nykodym’s (1996) study used Chi-square analysis to determine the statistical significance between perceptions of managers and subordinates of the formal appraisal process. Results indicated that both managers and subordinates believed that the formal appraisal process provided subordinates input about their performance and feedback for improvement. However, performance appraisals were not viewed as effective for providing motivation or improving manager/subordinate relationships. Longnecker & Nykodym suggested several recommendations for the improvement of appraisal systems:

1. Make sure managers and subordinates understand the appraisal system.
2. Assess the effectiveness of your current appraisal system.
3. Appraisal skills training for your managers is a must.
4. Increase your manager’s willingness to conduct effective appraisals.
5. Start with effective performance planning.
6. Make informal appraisals an ongoing activity.
7. Provide resources necessary to link pay to performance.
9. Use employee anniversary dates to stagger appraisals. (pp. 160-161)

Since performance appraisals are an important function for employee growth and development, Longnecker and Nykodym recommended that organizations continually
look for ways to keep performance appraisals at the forefront of management responsibilities.

**Current Principal Evaluation Practices**

Current principal evaluation practices are varied in approach. They are often not consistent between districts or states (Albanese, 2003; Lashway, 2003; Studebaker, 2000). Ginsberg and Berry (1990) discovered that, of the wide array of principal assessment practices being used, little systematic research exists in support of one over another. Explicitly, Ginsberg and Berry outlined five sources that illustrate the lack of empirical data to support principal evaluation practices:

1. Home recipes—insitutions, methods, and opinions presented in the literature with minimal empirical support for any of the approaches offered.

2. Literature reviews—a number of reviews which all derived the theme that little analysis or research on principal evaluation exists.

3. Guidelines and textbooks—published guides for practitioners, though no research-based evidence was found to substantiate any particular approach.

4. Surveys of practices—much of the research involved self-report surveys on practices in various school districts and states. The review concluded that many of these surveys were flawed methodologically, though the state of current practices was documented.

5. Research and evaluation studies—investigations examining specific instruments, control of principals, practices in effective districts, and various aspects of principal evaluation were found to suggest ways to improve principal evaluation, and implied the need for further research. (p. 67)
Although, this summary dates back 20 years, the same sparseness exists in current literature. Likewise, evaluation structures remain virtually unchanged despite the added complexities to the principal role. The following sections organize current principal evaluation practices into three similar categories such as those delineated by Ginsberg and Berry.

**Home Recipes**

Samuel DePaul (2006), a superintendent of Stanley County Schools in Albemarle, North Carolina, described a principal evaluation structure currently used by many districts. DePaul’s structure contained three basic components: goal setting, conferences, and site visits. First, the superintendent works directly with a principal to develop a school improvement plan and professional growth needs. Then, the superintendent guides the principal in setting specific measurable goals. The principal works toward these goals throughout the year and attends conferences with the superintendent in August, January, and June to check in on progress. Intermittently, the superintendent visits the site to walk and talk with the principal. Finally, the superintendent completes a summative write-up at the end of the year.

DePaul (2006) outlined a systematic approach to principal evaluation; however, he failed to cite any research to suggest that this method of evaluation adequately addresses the present day complexities of the principal role. Without any empirical research, any successes from this method are not generalizable to other situations. Therefore, without further study, this method of principal evaluation cannot be relied upon to guarantee principal effectiveness.
Similarly, additional evaluation practices described in the literature lack empirical support linking them with improved principal performance. For example, Mathews (2005) suggested that evaluation conferences for principals could be more beneficial if they were self-led. In a prescriptive article, Mathews contended that rather than a focus on compliance, self-led conferences would inspire and motivate principals to perform better. “Our principals can respond and do so robustly with suggestions of how we can improve student learning” (p. 40), wrote Mathews. Yet, no data were reported to support this assertion and therefore his suggestions cannot be generalized.

Bippus (2007), a superintendent from Indiana, suggested the use of a disinterested third party to perform principal evaluations so that the superintendent could act as a coach rather than an evaluator. He wrote, “Far too many employees take suggestions for improvement as criticism of them personally rather than using a performance evaluation to examine their own professional behaviors and actions” (p. 43). Bippus went on to say that a professional consultant could be hired to collect information regarding an employee’s performance.

In Bippus’s (2007) model, consultants collected data through interviews with various stakeholder groups and then met jointly with the principal and superintendent to discuss the findings. The principal then had an opportunity to provide data to support or refute the findings. Bippus concluded that when using this model, the superintendent acted as a coach rather than an evaluator. This method of principal evaluation is a self-report of practice, but Bippus does not present evidence to support his methodology or describe any success accrued from using the model.
Conversely, other educators such as Derrington and Sanders (2011) have been following and studying principal evaluation issues for several years. Based on their extensive research and experience they conceptualized a model to better address principal evaluation needs. In a peer reviewed article, they proposed a principal evaluation system focused on four components. First, supervisors should build a relationship and trust with the principals they evaluate. Next, evaluation practices should be aligned to research-based leadership standards. Third, competencies should be described and measured in relation to the data that will be collected. Finally, principal performance should be based on rubric descriptors (Derrington & Sanders, 2011). Though the authors offered sound advice based on a review of research, the article lacked any empirical evidence demonstrating any effectiveness of their proposed model.

**Guidelines and Standards**

Concomitantly, performance standards are also being implemented to enhance principal evaluation practices. Standards attempt to provide a foundation to effectively evaluate the performance of principals (Derrington & Sharratt, 2008). Probably the most widely used standards for principal evaluation are the Interstate School Leaders Licensure Consortium (ISLLC) standards. First published in 1996, the ISLLC standards have been adopted by the National Policy Board for Educational Administration. These standards provide a framework for states to use when developing their own standards. The ISLLC (2008) Standards Research Panel compiled a master template of 83 empirical studies consisting of qualitative, quantitative, mixed methods, meta-analyses, field trials, case studies, narrative reviews, and descriptive research studies. Each study included a
According to the National Policy Board for Educational Administration (ISLLC, 2008), principals are effective when they employ the following six standards:

Standard 1: An education leader promotes the success of every student by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by all stakeholders.

Standard 2: An education leader promotes the success of every student by advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff professional growth.

Standard 3: A school administrator is an educational leader who promotes the success of all students by ensuring management of the organization, operations, and resources for a safe, efficient, and effective learning environment.

Standard 4: An education leader promotes the success of every student by collaborating with faculty and community members, responding to diverse community interests and needs, and mobilizing community resources.

Standard 5: A school administrator is an educational leader who promotes the success of all students by acting with integrity, fairness, and in an ethical manner.

Standard 6: A school administrator is an educational leader who promotes the success of all students by understanding, responding to, and influencing the larger political, social, economic, legal, and cultural context. (ISLLC, 2008, pp. 1-3)

Each standard is followed by specific functions that could be observed by an administrator who is accomplished in the standard. States such as Missouri and Delaware
use these standards for principal evaluation (Brooks & Voss, 2008; Maxwell, 2008); however, the instruments remain weak and more empirical data are needed to assure instrument validity and generalizability for practical application (Babo, 2009; Brooks & Voss, 2008). Principal Evaluation tools should measure the performance of an administrator. They should be aligned to descriptors/indicators that demonstrate an effective principal, yet many are not (Brooks & Voss, 2008). In addition, few instruments are published, lack rigor, have low reliability, are outdated, and have not been field tested or vetted (Condon & Clifford, 2009).

In 2001, a large Southern California elementary district of 45 schools used the ISLLC standards to develop performance standards for principals. These standards were then incorporated into a peer evaluation process where principals participated in evaluating each other. The purpose was to promote professional and personal growth and development through honest communication and immediate feedback. Peer groups used several approaches to gather data about each other’s performance including classroom observations, analysis of student work, formal interviews with staff and parents, and regular meetings. Feedback centered on the following performance standards:

1. The principal is accountable for staff performance.
2. The principal is accountable for building leadership capacity.
3. The principal is accountable for customer satisfaction.
4. The principal is accountable for acting with integrity and fairness.
5. The principal is accountable for managing the school to be a safe, efficient, and effective learning environment. (Gil, 2001b, p. 38)
Although the district found that this process increased trust and made principal evaluation more meaningful, the district still questioned if the process was measuring what they wanted, whether performance tool was effective, and if it was an adequate way to measure growth (Gil, 2001b).

In 2009, Gerard Babo conducted a quantitative research project to determine if superintendents apply a preferred hierarchy of the ISLLC standards when evaluating principals. Babo constructed a 66-item, forced response survey to gather quantitative data. Forty email addresses, randomly selected from each of the 50 states, comprised the 2,000-person study sample. Out of the potential sample, study data accounted for only a 14%-15% return rate. Ninety-five percent of the respondents were current school superintendents or chief school administrators, 70% of participants male, and 30% female. Babo performed a series of Chi-square analyses, confirmed by a Friedman Test for related samples, to determine a preferred hierarchy of the ISLLC standards. Of the six standards, superintendents indicated instruction, vision, ethics, management, community, and larger context as the priority order of each standard. Yet, despite this apparent preference, Babo contended that, “a clearer picture on how to best measure and evaluate the actual operationalization of these standards needs to be the focus of discussion by all those with a vested interest in developing and producing quality principals” (p. 17).

Also using the ISLLC standards as a model, California has developed the California Professional Standards for Educational Leaders (CPSEL; WestEd, 2004), and groups such as the Association of California School Administrators (ACSA) and WestEd have joined efforts to help districts reconsider principal evaluation aligned to these
Kearney (2005) explained the following benefit of the CSPEL standards:

The CSPELs articulate a consensus about the broad range of knowledge and skills required for successful school leadership. The Descriptions of Practice provide the next level of detail within the standards. They further identify some of the research-based behaviors that most frequently yield positive student results.

(p. 20)

Kearney argued that the CSPELs help districts develop principal evaluation systems that are fair, rigorous, and coherent. Though Kearney referenced the work of district teams, explicitly their revised policies and specific plans, no specific qualitative or quantitative studies were offered as evidence of these practices.

However, also in 2005, a school leadership study commissioned by the Wallace Foundation, scholars began to question if the ISSLC and CPSEL standards were enough (Davis, Darling-Hammond, LaPointe, & Meyerson, 2005). Authors of this report argued that the ISSLC and CPSEL standards left out or underemphasized important components of effective leadership (Davis et al., 2005). Specifically, they referenced the principal’s role in curriculum design and implementation, support of effective instruction and assessment, and situational leadership. The authors referenced ongoing empirical studies and work such as Marzano et al. (2005) and recommended that the standards continue to be revised to reflect this new critical knowledge (Davis et al., 2005).

One attempt to remedy the weak leadership standards was made by Michael Bossi, ACSA’s director of leadership coaching, in 2008. He wrote a blog introducing the Six Dimensions of Leadership. He referred to the ISSLC and CPSEL standards as
incomplete and unchanging, citing that they do not include enough of a focus on collaborative and strategic leadership (Bossi, 2008). Bossi (2008) also made reference to Marzano et al. (2005) and Cotton’s (2003) already well-researched lists of effective leadership behaviors; however, he described them as too cumbersome.

Therefore, Bossi’s (2008) team at ACSA came up with these six encompassing leadership indicators: visionary/cultural leadership, operational leadership, instructional leadership, learning leadership, collaborative leadership, and strategic leadership. Under each indicator, Bossi offered definitions the dimensions and encouraged leaders to use them as an integrated resource for leadership. He concluded by talking about the future elaboration of the Six Dimensions of Leadership; however, no references to the development of the Six Dimensions of Leadership could be found in the literature.

The Essential Behavioral Leadership Qualities (EBLQ) is another standards-based approach to principal assessment. This method attempts to measure a leader’s effectiveness against systematically derived standards. It is important to note that these standards include the opinion of teachers and what they perceive to be effective behaviors of a leader. Oyinlade (2006) conducted a mixed methods case study employing the EBLQ approach. The case study led to the creation of a leadership assessment tool of 18 measurable qualities for effective principals. Teachers then rated their principals against the assessment tool. Oyinlade examined the collection of data in the study to determine areas of perceived strength and weakness.

The study by Oyinlade (2006) is limited, however, in that it only reflected data for 25 principals. In addition, the study only included principals of schools for students who are blind or visually impaired. Oyinlade argued that this approach could be used in other
settings and with different demographics of leaders and suggested that the instrument be modified based on the opinions of the given subordinates. With so many variables to control for, more studies using the EBLQ approach would be needed to determine the assessment’s generalized usefulness.

**Research and Evaluation Studies**

Despite overall limited research on principal evaluation methods, portfolio-based evaluation has gained increasing empirical support (Green, 2004; Marcoux et al., 2003; Russo, 2004). Qualitative data indicate improvements in performance of principals using this approach. Increased principal self-reflection and improved leadership are reported results of portfolio-based evaluation methods (Green, 2004; Marcoux et al., 2003). In addition, it provides principals with a more formative evaluation (Russo, 2004).

Utilizing a portfolio evaluation method, principals demonstrate accountability by gathering evidence of performance. Green (2004) performed a case study of principals in Pomona, CA who used portfolio-based evaluation methods. Green quoted Patrick Leier, Pomona’s superintendent, as saying, “Our principals are becoming more focused on outcomes rather than compliance. The process of preparing portfolios helps to close the gap between self-knowledge and action by the self” (p. 32).

Another case study conducted by Marcoux et al. (2003) examined the impact of portfolio-based evaluation methods on leadership effectiveness, student achievement, professional development of teachers, and the reflective practice of the principal. For this phenomenological study, the researchers selected a purposeful sample of one rural New York school district that was using portfolio-based evaluation methods. The sample consisted of one superintendent, two assistant superintendents, five principals, and
10 teachers. Marcoux et al. collected data from various sources including interviews, focus groups, test score data, and survey responses.

Marcoux et al. (2003) identified 17 themes that emerged from the data corpus demonstrating the impact of the principal evaluation portfolio process. Themes were aligned to the research questions they answered:

1. How has the principal portfolio evaluation process impacted leadership effectiveness?
2. How has the principal portfolio evaluation process impacted student achievement?
3. How has the principal portfolio evaluation process impacted teacher professional development?
4. How has the principal portfolio evaluation process impacted reflective practice?

The following summarizes the themes that materialized for each of the corresponding questions:

1. Communication and open dialogue, common language and a common vision throughout the district, on-going self-assessment, and increased principal visibility.
2. Principal support of professional reading and book studies, a focused effort on state standards and student achievement, and principal being able to identify individual teacher strengths.
3. Focused and well-planned professional development, increased teacher input and empowerment, and a commitment to shared learning.
4. Increased collaboration and communication, increased trust, redirected opportunities, and increased ability to problem-solve. (Marcoux et al., 2003)

Results of the portfolio study by Marcoux et al. (2003) revealed improved leadership effectiveness and increased self-reflection of principals. Students and teachers also benefited from the portfolio assessment process. Findings revealed a relationship between participation in this type of evaluation and principals’ increased focus on student achievement and teacher professional development (Marcoux et al., 2003). In addition, portfolios provided principals with both formative and summative feedback through an ongoing self-reflective process. Although Marcoux et al.’s study discovered some initial benefits of portfolio use in principal evaluation, more studies with greater number of participants and varied demographics would need to be conducted to ensure generalizability of these findings.

A current phenomenological study by Parylo, Bengtson, and Zepeda (2010) confirmed the lack of published research on principal evaluation. In a qualitative study, Parylo et al. (2010) aimed to determine the meaning of the principal evaluation process based on the lived experiences of the participants. Parylo et al. conducted semi-structured interviews with 16 principals from four school systems in Georgia, and collected documents from research sites where available. Interviews were audiotaped and transcribed verbatim. Participants represented eight males and eight females, ranging from less than 1 year to 9 years of experience as a principal. The school breakdown consisted of one primary, six elementary, five middle, and four high schools. Parylo et al. used a two-step data analysis process to examine the transcripts and apply theories about supervision and learning.
Results of the Parylo et al. (2010) study indicated that principals viewed their evaluations as an evaluation of the school. They used words such as important, multi-faceted, fair, complex, and transparent when describing evaluation processes. Seven themes materialized from the data including principal awareness of evaluation processes, principal evaluation as a process rather than an event, transparency of the process, open dialogue, trust and respect with supervisors, feedback, and support to become a better leader. Principals reported a shift in focus of evaluations from managerial to student achievement; however, descriptions of existing versus ideal procedures demonstrated a need for improvement in principal evaluation processes (Parylo et al., 2010).

Another principal evaluation method found in the literature advocates a 360º feedback approach to principal evaluation. In the field of business, a 360º feedback approach, or multi-source multi-rater (MSMR) assessment system, is often used for short-term coaching.

Multi-source multi-rater assessment systems, otherwise known as 360 degree feedback, typically involve a target manager providing a self-rating on various competencies, which is then compared with the ratings of relevant others such as supervisors, peers, subordinates and occasionally clients. (Jones & Fletcher, 2002, p. 145)

According to an article written by the Association to Advance Collegiate Schools of Business (AACSB) International (2006), little empirical evidence has been collected on the effectiveness of these tools. The article cites a quantitative study by Robert Hooijberg, a professor of organizational behavior at the Institute for Management Development (IMD) in Lausanne, Switzerland. Hooijberg surveyed 900 students who
had experienced 360º feedback to determine the effects of short-term coaching interactions using the 360º feedback approach. Approximately 240 students responded and reported that the 360º feedback approach was helpful in receiving short-term coaching regarding personal development areas if the coaches inspired them and gave advice. In the future, Hooijberg plans to investigate whether this method will be effective in improving employee performance.

In education, a 360º approach may provide principals with better feedback for personal and leadership development (Dyer, 2001; Green, 2004; Maxwell, 2010; Moore, 2009, Porter et al., 2008). This approach applies an evaluation instrument that gathers feedback about a principal from various stakeholders. The instrument allows multiple raters to provide principals with honest information about their performance (Moore, 2009; Porter et al., 2008). Dyer (2001) cautioned, however, that the purpose of a 360º approach should be to provide developmental feedback and not be used as a stand-alone tool of evaluation.

Various 360º feedback tools have been developed including one at Vanderbilt University between 2006-2010. The Vanderbilt Assessment of Leadership in Education (VAL-ED), is a tool aimed at assessing principal performance based on core components of effective schools: high standards for student learning, rigorous curriculum (content), quality instruction (pedagogy), culture of learning and professional behavior, connections to external communities, and performance accountability. Principals are rated on the key behaviors of planning, implementing, supporting, advocating, communicating, and monitoring that they can engage in to meet these six components. These behaviors,
derived from the research of Murphy, Elliott, Goldring, and Porter (2007), also intersect with the ISLLC standards.

In 2008, 218 schools of all levels from across the nation participated in a national field trial of the 360º VAL-ED tool. National norms were created from data gathered. These percentile norms along with outlined performance levels of distinguished, proficient, basic, and below basic became the tool’s benchmark measurements of principal performance using the VAL-ED. An article by Goldring, Porter, Murphy, Elliott, and Cravens (2009) introduced a conceptual framework and provided further rationale for the VAL-ED tool. Each key process of the instrument is based on research of how the leadership behavior adds value to student achievement. In addition, the authors explain that this new leadership assessment focuses solely on leadership behaviors that are linked to student learning. Yet, although over 4 years have been spent on developing, testing, and revising the tool, little data have been gathered from real use of the tool (Porter et al., 2010b).

A brief published in 2009 by Learning Point Associates sought to identify current, publicly available, rigorous, and reliable principal performance assessment instruments. Authors Condon and Clifford (2009) established three criteria that instruments had to meet to be considered for the brief: instruments must be intended for use as a performance assessment, psychometrically tested for reliability and validity with a minimum reliability rating of 0.75, and publicly available for purchase. Condon and Clifford identified 20 potential instruments through Google Scholar. Of these 20 instruments, 8 met the preset criteria for rigor.
Condon and Clifford (2009) reviewed and summarized each of the eight identified instruments by approach, time required, content and construct validity, and reliability. Of the eight reviewed instruments, Condon and Clifford classified four with poor reliability (below 0.80), three with moderate reliability (0.80-0.90), and one with high reliability (above 0.90). Only two of the eight instruments, the Leadership Practices Inventory (LPI) and the Vanderbilt Assessment of Leadership in Education (VAL-ED), were developed within the past decade, while the others dated back at least 10-20 years. The LPI’s 30-item survey measure of general leadership received a poor test-retest reliability score of 0.79, while the VAL-ED’s 360º assessment tool received a reliability score of 0.98.

Overall, the review indicated that, although the literature suggests an interest and need for principal accountability, few instruments have been developed or reviewed for continued validity to meet the needs of current principals. And, of the best tool evaluated, the VAL-ED, little data exist to demonstrate the tool’s effectiveness in actual school settings.

**Other Research**

Practices and procedures around principal evaluation are beginning to gain the attention of several large educational groups such as New Leaders and West Ed. New Leaders is a national nonprofit focused on developing strong school leaders and leadership policies. In 2010, New Leaders published a paper focused on evaluating principals. The paper aimed at finding ways to help districts better balance principal accountability and professional growth. The report offers four main ways to better improve principal evaluation methods:

1. Base principal evaluations on student outcomes.
2. Ensure that the central office staff are also held accountable for principal effectiveness.

3. Create performance expectations and accountability, while allowing for professional growth and improvement.

4. Ensure that the evaluation system can be modified and improved over time (New Leaders, 2010).

WestEd, another nonprofit, public research and development agency, also continues to examine practices and procedures around principal evaluation. A literature review of 68 peer reviewed and nonpeer reviewed publications on principal evaluation methods conducted by WestEd reported findings consistent with this literature review (Davis et al., 2011). Davis et al. (2011) organized their literature review into two sections highlighting primary and secondary sources. Primary sources consisted of peer reviewed and professional publications focused on principal evaluation. Practitioner journals and best practice recommendations were considered secondary sources.

Twenty-eight studies were analyzed in the first section and reported under the following themes: implementation (15 studies), instrumentation (7 studies), portfolio-based evaluation (3 studies), and component analysis (3 studies). Of the implementation studies, the literature review found that most district-developed principal evaluation tools lack validity and reliability, alignment between district evaluation standards and professional standards is inconsistent, a variance exists in methods and tools used to evaluate principals, and superintendents valued tools more than principals when the tools were not explicit or developed collaboratively. The instrumentation findings discovered very few valid and reliable instruments except for the Val-Ed instrument previously
described in this review. Portfolio-based evaluation proved to be rarely used and time consuming. Although good for principal reflection and for gathering evidence of performance, no proof was found that portfolio-based evaluation methods actually improved principal performance or student achievement. Finally, the three component analysis studies reiterated a need for collaborative involvement of stakeholders in the development in quality evaluation tools. Moreover, analysis of principal behaviors strengthens the quality of principal evaluation procedures (Davis et al., 2011).

Davis et al. (2011) considered 40 more secondary sources in the next section of their literature review and reported findings under four perspectives: status of principal evaluation systems, critical descriptions and commentaries, recommendations for more effective principal evaluations, and descriptions of best practices in principal evaluation. Nine secondary sources described the status of principal evaluation systems and focused on evaluation policies, processes, and features. Davis et al. found, that aside from practices being varied, evidence-based evaluation systems are gaining in popularity. In addition, using summative and formative purposes is increasing and districts are considering context more when designing tools for principal evaluation. Sixteen publications offered descriptions and commentaries of principal evaluation systems. A summary of these sources agreed that principal evaluation tools tend to be locally developed, not well aligned with professional standards or the literature on leadership effectiveness, and offer little feedback to principals. Moreover, data are sparse regarding the effects on principal leadership. Another 33 sources offered suggestions for improvement of principal evaluation methods. Recommendations from these publications suggested that methods should focus on principal growth, align to school and
personal goals, involve multiple stakeholders, consider multiple methods, be flexible and collaborative, measure growth, include valid and reliable tools, and should be based on a conceptual framework that includes research on effective leadership. Lastly, 11 publications proposed exemplary evaluation practices used by various school districts; however, much of this literature was published in the 1980s (Davis et al., 2011).

The literature review conducted by Davis et al. (2011) emphasized the lack of research in the area of principal evaluation practices. Yet, increasing accountability of principals augments the importance of creating accurate, effective, and fair evaluations of principal performance. Expectations and standards for leadership should be clear and provide useful feedback for professional growth and development (Davis et al., 2011). Thus, it is critical to study current principal evaluation practices and consider how to improve practices to better align with current leadership expectations for improving student achievement.

Another group from WestEd, Leon, Davis, Kearney, Sanders, and Thomas (2011), published findings from a principal evaluation study conducted to explore how principal effectiveness is assessed. Following a comprehensive research review, Leon et al. (2011) conducted focus groups with superintendents, administrators, principals, and professors, as well as held conversations with individual practitioners, experts, and researchers in education and personnel evaluation, to derive “best practices” in principal evaluation. Leon et al. found that the most effective evaluation systems emphasized the following:

2. Research and address state and national professional standards.
3. Supportive district policy.
5. Principal involvement in setting and prioritizing individual professional goals and objectives.
6. Professional development and growth of the principal.
7. Ongoing opportunities to review evaluation evidence and receive feedback.
8. The use of multiple forms of data/evidence.
9. Adaptable to adjust to various contexts, needs, and circumstances.
10. Ongoing training for evaluators and principals.
11. Validity and reliability of processes, data, decisions, and outcomes.
12. A regular review of the system (Leon et al., 2011).

These features were also compared with the personnel evaluation standards described in the Joint Committee on Standards for Educational Evaluation (JCSEE, 2009) framework. To further understand the current status of principal evaluation systems, WestEd also completed an in-depth study at how six different states were implementing principal evaluation systems (Mattson Almanzán et al., 2011). Six states were identified in this analysis: Delaware, Iowa, New Mexico, North Carolina, Ohio, and South Carolina. Each state’s system was described in terms of policy context, system structure, and implementation. A closer look at the six states showed that all six states had a system and/or tool in place, were developed based on state and national professional standards, and focused on professional growth and improving principal performance. Though systems were clearly outlined, flexibility and practice of implementation existed across states and districts (Mattson Almanzán et al., 2011).
Although California was not a state included in the analysis, WestEd specifically discussed the implications for principal evaluation policies in California (Morse, 2011). Already, the California Education Code (2011) requires the evaluation school administrators, but it does not specify the methods required to do so. In 2001, California developed the California Professional Standards for Education Leaders (CPSEL) from the Interstate School Leaders Licensure Consortium (ISLLC) Standards for School Leaders. These standards provided guidelines for successful leadership behaviors specific to California schools; however, in many districts they are not specifically tied to principal evaluations. Due to increasing debates over evaluation policies in California, the need for evaluation reform is at hand (Morse, 2011). This is just one more reason why we should look more closely at the principal evaluation practices in California.

Factors Influencing Principal Evaluation

Principals are held accountable for their work primarily through principal evaluations. Yet, the complexity of the principal role and the context in which they serve makes it challenging to standardize assessment practices (Glasman, 1992; Heck & Glasman, 1993; Heck & Marcoulides, 1992). Moreover, groups such as school boards, teachers, students, and parents further influence expectations for principals. These stakeholder demands vary and are perhaps at odds with one another (Catano & Stronge, 2006), making it even more difficult to determine what to emphasize in principal evaluations. Researchers agree that there is lack of consistency in evaluation processes, usefulness, and effects on principal actions (Gaziel, 2008; Goldring, Cravens, et al., 2009; Murphy et al., 1985; Stufflebeam & Nevo, 1993; Toler, 2006).
Some scholars have attempted to clarify principal evaluation as a means to improve administrator practice (Glasman, 1992; Glasman & Heck, 1992; Glasman & Martens, 1993; Heck & Glasman, 1993). In a qualitative study, Murphy et al. (1985) explored principal evaluation procedures of effective schools and districts, seeking to understand how these districts assess principal work. They examined the supervision and evaluation of principals in 12 effective school districts in California. From over 1,000 possible school districts, the researchers selected four unified, three high school, and five elementary districts on the basis of consistently high student test scores. Murphy et al. conducted interviews with each superintendent of the selected districts and found that all but two of these superintendents were directly responsible for evaluating principals.

According to the data collected by Murphy et al. (1985), superintendents were highly involved in the principal evaluation process, met with principals more often, and made more frequent visits to individual school sites. Ten of the superintendents reported that reviewing curriculum and instruction with principals was either a primary or highly important activity when visiting schools. Furthermore, progress on yearly objectives and test scores were key components of evaluation in all districts. Yet, despite the implications of this research, limited studies have been conducted since to further explore effective principal evaluation practices (Murphy et al., 1985; Stufflebeam & Nevo, 1993; Toler, 2006).

A recent study by Toler (2006) emphasized a discrepancy between defined principal evaluation measures and current practices. An extensive review of principal evaluation practices indicated that researchers, educators, and policy makers across the Commonwealth of Virginia agreed that principal evaluation processes must be in place in
order to improve leadership effectiveness, on behalf of schools and their communities. However, through questionnaires returned from 91 school divisions, analyzed in combination with documents returned by 61 of these respondents, researchers determined that approximately a third of the responding school divisions did not even have formal procedures in place.

Scholars concur about the need for improvement in principal evaluation methods (Amsterdam, Johnson, Monrad, & Tonnsen, 2005; Bippus, 2007; Cuban, 1988; Kearney, 2005; Moore, 2009; Peterson, 1991; Rooney, 2009). Ginsberg and Berry (1990) described a lack of attention to principal evaluation, resulting in the poor quality of current methods. An increasing body of research not only demonstrates the need for overall high-quality principal evaluation methods, but also highlights the importance of establishing congruence between these methods and research-based criteria for effective leadership associated with school performance (Albanese, 2003; Ginsberg & Berry, 1990; Goldring, Porter, et al., 2009; R. Johnson, 2005; Murphy et al., 1985; Studebaker, 2000).

Amsterdam et al. (2005) presented a case study that, “describes the process of developing and validating a principal evaluation system in which stakeholders contributed to the quality of the criteria and instrumentation” (p. 222). Through a collaborative process, stakeholders such as the South Carolina Board of Education, superintendents, and principals, assisted with the planning and design of the principal evaluation. A pilot study was conducted to collect data using this new system. Feedback indicated that pilot members felt the evaluation system increased dialog between superintendents and principals. In addition, principals found the instrument useful for reflection regarding their role. Implications of this study suggested the importance of
involving stakeholders in the process of creating evaluation tools. Although Roberts (2003) described such a process in an article focused on public personnel management, this is still not a common practice in current principal evaluation methods.

Still, even when practices for principal evaluation exist, methods are often summative or punitive (Albanese, 2003). Albanese (2003) performed a qualitative study to explore assessment practices used to evaluate principals in the state of Rhode Island. Albanese used multiple methods for collecting data including interviews, document analysis of evaluation instruments, and other related information. Albanese interviewed 44 principals, with at least 1 principal representing each of Rhode Island’s 36 school districts.

Overwhelmingly, the principals in Albanese’s (2003) study felt that evaluations were subjective in nature, citing shortcomings that limited the ability of evaluations to influence and improve performance. Many expressed concerns that “evaluations did not take into account the day-to day reality of their positions, with certain aspects of the job being ignored” (Albanese, 2003, p. 114). Additionally, a majority of the principals felt that the superintendents lacked firsthand knowledge of occurrences within schools and of principal performance. Evidence reinforced the need to align evaluation methods with job expectations and responsibilities, developing outcome-based methods that have greater capacity to foster improved principal performance.

A quantitative study conducted in an urban setting by Goldring, Porter, et al. (2009) found that current principal evaluation documents failed to emphasize some of the most important factors related to improved learning, nor had the instrument been
validated as a measure of principal effectiveness. Specifically, Goldring, Porter, et al. wrote:

Our study provides a timely update on the state of principal assessment in urban districts. More importantly, it makes a case for the urgent need for researchers and practitioners to sharpen the conception of school leadership with a learning-centered focus and to operationalize such a conception through an assessment process characterized by desired psychometric properties. (p. 36)

These findings support earlier claims advanced by Edmonds (1980) 30 years ago regarding the need to address student performance inequities, and connects these assertions with the necessary functions fulfilled through principal evaluation. Once more, incongruence existed between evaluation methods and established criteria for effective principal leadership.

Public demands for educational accountability further complicate the nature and purpose of principal evaluation practices (Glasman & Martens, 1993; Heck & Glasman, 1993; Smylie & Crowson, 1993; Snyder & Ebmeier, 1992). In most cases, scholars agree on necessary responsibilities of effective principals. However, recent studies scrutinize the degree to which evaluation methods coincide with leadership practices advocated in the literature (Conca, 2008), as well as the degree to which principal accountability is directly tied to student achievement (Maxwell, 2008). Publicly, principals are being held accountable for educational outcomes as a measure of their performance (Catano & Stronge, 2006; Maxwell, 2008), yet the methods to ensure this accountability are still being superficially or inconsistently implemented (Brooks & Voss, 2008; Conca, 2008).
Therefore, it is important to analyze current principal evaluation practices and identify areas for improvement.

**Summary**

Framed by five sections, a history of the principal role, key leadership behaviors of today’s principals, an overview of evaluation, current principal evaluation practices, and factors influencing principal evaluation, this literature review substantiated a lack of research studies linking effective principal evaluation methods with leadership behaviors. Likewise, researchers have found little consistency in evaluation procedures (Gaziel, 2008; Goldring, Cravens, et al., 2009; Murphy et al., 1985; Stufflebeam & Nevo, 1993; Toler, 2006). With increased accountability on principals for ensuring student achievement, it is imperative that we discover the essential components and processes that best evaluate principal’s leadership capabilities.

Like Ginsberg and Berry (1990), this review of the literature found a variety of tools and approaches being used for principal evaluation with meager empirical research to support claims. DePaul (2006), Mathews (2005), and Bippus (2007) offered prescriptive solutions without evidence of a link to improved principal performance. Several others advocated for the development and use of standards in principal evaluation, however they did not present adequate data to demonstrate the application or effectiveness of processes (Babo, 2009; Brooks & Voss, 2008; Derrington & Sharratt, 2008; Gil, 2001a; Kearney, 2005; Maxwell, 2008; Oyinlade, 2006). Although some evaluation studies such as the VAL-ED and the portfolio process were supported with compelling empirical research, limited quantity of studies indicates that this is still an area in need of much more attention.
Evaluation is an important tool for measuring the competence and performance of an employee. Evaluation practices exist in all professions, however many factors influence the effectiveness of evaluation methods. In educational research, there is a proven link to effective leadership behaviors and increased student achievement; therefore, evaluation methods could provide the necessary feedback for principals to effectively improve their performance.

Current gaps between effective leadership behaviors and current principal evaluation methods underscore the need for further studies in the area of principal assessment. In addition, the lack of empirical evidence warrants the need for deeper understanding of principal evaluation methods as they relate to leadership behaviors. Chapter 3 will discuss the research methods to be used in this study. It will also provide an overview of the research design, as well as the proposed data collection and analyses methodology.
CHAPTER 3—RESEARCH DESIGN AND METHODOLOGY

The literature review unveiled a lack of research studies around effective principal evaluation methods, specifically as they relate to leadership behaviors. Thus, this study served a dual purpose of examining the explicit characteristics of principal evaluation methods that inform principal practices, as well as determining ways in which elementary principals perceive evaluation procedures support improvement of their leadership effectiveness. Through a qualitative descriptive analysis, this study explored current elementary principal evaluation methods as described by principals in 19 different Southern California school districts.

Merriam (1998) defined qualitative research as “an umbrella concept covering several forms of inquiry that help us understand and explain the meaning of social phenomena with as little disruption of the natural setting as possible” (p. 5). Guided by the philosophy of phenomenology, Merriam described qualitative researchers as “interested in understanding the meaning people have constructed, that is how they make sense of their world and the experiences they have in the world” (p. 6). This study embodied what Merriam characterized as interpretive research, seeking understanding and meaning through an inductive mode of inquiry.

According to Merriam (1998), the most common form of qualitative research in education is basic or generic qualitative study. Basic or generic qualitative study is characterized by three main attributes:

1. Includes description, interpretation, and understanding.

2. Identifies recurrent patterns in the form of themes or categories.

Data for this type of study are usually collected through interviews, observation, or document analysis. This study used both interview and document review methods.

One of the main reasons to conduct a qualitative study is for the purpose of exploring a topic that not much has been written about (Creswell, 2009). The dearth of studies in the literature made principal evaluation a significant topic in need of further exploration. Conducting a basic or generic qualitative study of current elementary principal evaluation practices across multiple districts enabled this researcher to better describe the current state of principal evaluation practices being used in Southern California which will perhaps enlighten school districts, superintendents and other administration responsible for supervising and evaluating elementary principals of practices that principals perceive most support their leadership effectiveness.

Research Questions

Two overarching research questions guided this work:

1. What is the status of current elementary school principal evaluation procedures?

2. How do elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness?

To help determine this, the following research questions were investigated:

1. What do processes and/or practices used in principal evaluation systems entail?

2. How consistent are principal evaluation practices across districts?

3. Do current principal evaluation practices align to research on effective leadership practices?
4. How do principals perceive their evaluation systems influence their leadership behaviors?

Elementary school principals participated in interviews to describe their current evaluation systems and provided details regarding how they perceive evaluation procedures support improvement of their leadership effectiveness. Interview questions concentrated specifically on what salient actions, processes, and themes exist in the principal evaluation process; how evaluation processes measure principal effectiveness; what feedback, if any, evaluation methods provide principals; and how principal evaluation practices improve leadership effectiveness. Data from the interviews helped to address a final research question focused on how principal evaluation practices compare across districts. Document reviews of forms added additional data to the aforementioned questions.

**Methodology**

This study is best described as a qualitative descriptive study. “The purpose of descriptive research is to describe a phenomenon” (Glatthorn & Joyner, 2005, p. 43). Marshall and Rossman (2006) argued that qualitative research should be used for research that stresses context, setting, and participants’ frames of references. This study considered the frame of reference of various elementary school principals across various settings within the context of principal evaluation methods, and therefore lent itself to a qualitative research approach. Qualitative methods of data collection can include document reviews, interviewing, and observations (Creswell, 2009; Marshall & Rossman, 2006; Merriam, 1998); however, this study focused mainly on interviewing as the primary
method of data collection with document analysis providing additional information relevant to the research questions.

**Context**

To gain a broad perspective of the current state of elementary principal evaluation methods, 10 principals across 10 different school districts were interviewed. There are 41 school districts in San Diego County; however, not all districts met the criteria for this study. Districts varied in size and student population; the smallest district with 2 schools and the largest district with over 100 elementary schools. Most districts averaged between 8-20 schools, and only elementary schools were selected for the study. This excluded six San Diego school districts that only serve secondary students.

The study included a mix of public urban and suburban districts from both K-6 and K-12 school districts. All of the selected schools were from the Southern California region, specifically from San Diego County. Small districts where superintendents also serve as principals were not selected for this study, nor were very remote districts on the borders of the county. This excluded another 15 districts from the study.

**Sample**

A purposeful selection of 10 elementary school principals was chosen for this study. Interview studies tend to include an average of 5-15 interviews (Kvale & Brinkmann, 2009). However, scholars assert that the number of participants needed for a study depends on what information is needed to answer the research questions (Kvale & Brinkmann, 2009; Marshall & Rossman, 2006; Merriam, 1998). For this study, a group of 10 was chosen to provide sufficient data while keeping the study reasonable in size.
Having too many participants would have added to the complexity and limited the ability to deeply analyze the data (Kvale & Brinkmann, 2009; Marshall & Rossman, 2006).

The sample was intentionally chosen based on participants’ abilities to contribute insights into the issues of principal leadership and evaluation (Boeije, 2010; Creswell, 2009; Merriam, 1998). Merriam (1998) used the term *typical sample* to describe the average person that reflects the phenomenon being studied. Participants in the sample chosen for this study were all elementary school principals with adequate experience to describe principal evaluation methods used in their perspective districts. Each principal represented a different school district within the Southern California region of the United States. The variability in school districts provided an opportunity to examine trends, or lack of trends, in current principal evaluation procedures and practices.

Personal and professional connections provided an adequate pool of candidates from which to select willing participants for this study. Participants were selected from a defined population (Boeije, 2010) and met a general list of selection criteria (Merriam, 1998). Merriam (1998) suggested that the criteria chosen for the sample should reflect the purpose of the study. Therefore, this study identified current status as an elementary school principal, employment in a Southern California school, a minimum of 3 years experience as an elementary school principal, and a willingness to participate in the study as most important criteria for participants.

However, since the researcher did not personally know 10 different principals from 10 different districts, a convenience sample had to be utilized. A convenience sample begins with accessible connections and builds from those initial connections (Marshall & Rossman, 2006). The sample in this study was generated through
networking with other principals and colleagues known to the researcher. Referrals to other principals were then generated through these initial connections. Emails and phone calls provided the means to identify, network, and connect with an appropriate and willing participant pool.

Data Collection Methods

Interviews served as the primary source for data collection. Additionally, a document review of principal evaluation instruments, district documents, websites, demographic data, and field notes supported this study.

Interviews

Data were collected primarily through in-depth interviews. In describing the in-depth interviewing method, Marshall and Rossman (2006) recommended framing the interview with a few general topics, while letting the phenomenon of interest unfold. Therefore, face-to-face, semi-structured interviews attempted to elicit views and opinions from participants. “Interviews provide an opportunity for researchers to learn about social life through the perspective, experience, and language of those living it” (Boeije, 2010, p. 62). In addition, interviews allow participants to convey information, feelings, insights, and interpretations that are otherwise not observable (Kvale & Brinkmann, 2009; Merriam, 1998).

To gain a true understanding of the nature of current elementary principal evaluation practices, interview questions were not entirely prestructured, nor left entirely open (Boeije, 2010; Marshall & Rossman, 2006). The literature review provided the theoretical framework in the development of the interview questions. Types of interview questions such as introductory questions, probing questions, and direct questions as
outlined by Kvale and Brinkmann (2009) also served as a guide during question design. Introductory questions allowed for spontaneous answers and rich descriptions, probing questions helped elicit deeper understanding of answers, and direct questions provided specific context to topics being discussed (Kvale & Brinkmann, 2009).

Principals not participating in the study were asked to provide feedback regarding question clarity after interview questions were developed. This feedback was used to revise the interview questions only, and no input from these principals was added to the data corpus of the study. In addition, feedback from the reviewing principals was used to weed out both leading and yes-or-no type questions to ensure effective questions (Merriam, 1998).

Interviews took place at mutually convenient, pre-agreed upon times in private areas to ensure confidentiality. Interviews lasted between 15 and 50 minutes, were audiotaped, and then transcribed verbatim. Names were stated on audiotapes only for data reference purposes during analysis. Participants were ensured confidentiality and guaranteed anonymity in final study results. No names, schools, or districts were specifically cited in the findings. Sensitivity to the needs of the subjects was maintained at all times. All interview data were stored in a password-protected computer at all times during the study.

During the interviews, the researcher took notes to keep track of answers to ensure comparable data were collected from each participant and to record thoughts leading to subsequent questions. Ensuing interview questions such as follow-up questions, specifying questions, and interpreting questions were infused as necessary. Follow-up questions were asked to seek further elaboration to answers, specifying questions to gain
more precise descriptions, and interpreting questions to ensure clarification of answers (Kvale & Brinkmann, 2009). Participants were informed that follow-up interviews might be necessary if information was gathered during other interviews that was not part of the original question set, but was determined not necessary at the end of the data collection process.

**Document Review**

Where available, principal evaluation instruments and documents were collected and provided a secondary source of data. Marshall and Rossman (2006) stated that reviewing documents could provide knowledge of the history and context surrounding a specific setting. All documents were collected either during or after the interviews. The researcher later analyzed documents for alignment to identified themes emerging from interviews. All paper documents collected were kept in a locked filing cabinet in the researcher’s home. Any electronic data provided was kept in the researchers password-protected computer.

In a chapter entitled “Mining Data From Documents,” Merriam (1998) depicted physical material collected as artifacts grounded within the setting of the study. These instruments can serve as a way to check on information gathered through the interview process. In this study, documents were gathered from participants for the specific purpose of learning more about principal evaluation methods across districts.

**Ethics**

The ethical nature of the subjects and the potential political outcomes resulting from evaluation processes was considered. Participation was voluntary, and participants were free to withdraw at any time without consequence. Since evaluation is often tied to
a person’s employment, it can be a sensitive topic (Boeije, 2010). No personal information of the participants was disclosed to anyone, as privacy was a very important consideration. Furthermore, participants were assured that any links between codes, names, or districts would be permanently deleted upon completion of the dissertation. Each interviewee granted written permission in accordance with informed consent regulations for human subjects. Consent forms explicitly detailed the purpose, description, confidentiality, benefits, and direct effects of participation in the study. Forms were kept with the other paper documents in the locked filing cabinet in the researcher’s home. If interested, participants had the opportunity to request and receive an outcome brief denoting the results at the conclusion of the study.

All aspects of the study including the consent form, interview questions, and recruitment materials were reviewed and approved by the San Diego State University Institutional Review Board. To further protect participants, and attempt to avoid potentially vulnerable disclosures, questions were designed to focus more specifically on the process of principal evaluation than evaluations of participants themselves. Due to the honest and intimate answers of the interviewees, it is a reasonable assumption that the participants felt confident that no identifying information would be disclosed.

**Data Analysis**

Marshall and Rossman (2006) described a seven-phase process to data analysis including organizing the data, immersion into the data, generating categories and themes, coding the data, offering interpretations through analytic memos, searching for alternative understandings, and writing the report to present the study. This process was used for analyzing data in this study. Data from interviews were transcribed, organized, and coded
and sorted into categories and themes. “Coding data is the formal representation of analytic thinking” (Marshall & Rossman, 2006, p. 160).

Categories and subcategories were created using a constant comparative process between interviews and documents. These categories and themes were determined based on information that recurred regularly in the data (Boeije, 2010; Creswell, 2009; Merriam, 1998). Boeije (2010) described qualitative analysis as the segmenting and reassembling of data. Through this segmenting and reassembling process, transcribed interview data were broken down and then synthesized into relevant and meaningful parts in order to generate understanding of current principal evaluation processes. The researcher used NVivo, a qualitative data analysis software, to assist in this analysis process.

As a way of triangulating data, documents, evaluation instruments, websites, demographic data, and field notes were also analyzed against the research questions. A content analysis approach was implemented when reviewing all documents and artifacts reviewed. Content analysis is “a method for describing and interpreting the artifacts of a society or social group” (Marshall & Rossman, 2006, p. 108). According to Marshall and Rossman (2006), the ability to perform content analysis without disturbing the setting can add strength to a study; however, interpretations made by the researcher can be perceived as potential weaknesses. With this as a consideration, all documents and artifacts collected were clarified with the interviews of the participants providing the documentation whenever needed.

This triangulation of the data was only one way the researcher addressed internal validity. Member checking was also implemented to safeguard the accuracy of the data
analysis (Creswell, 2009). Transcriptions were sent to all participants to review for correctness. Verbatim transcriptions were also shared with a professional colleague to check interpretations of themes.

**Trustworthiness of Data**

The quality of data collected can be greatly increased when the researcher is able to develop trusting relationships with the study participants and the credibility of the study can be reasonably assured (Marshall & Rossman, 2006). By using personal and professional connections to select the sample for this study, there was a greater potential for developing trust between the researcher and the participants.

Credibility, transferability, dependability, and confirmability also contribute to the trustworthiness of the data. Marshall and Rossman (2006) described each of these the following way:

1. Credibility: The ability to demonstrate that the subject of the inquiry was appropriately identified and described.

2. Transferability: The ability of the researcher to argue that the findings of the study will be useful to others in similar situations, with similar research questions.

3. Dependability: The ability to account for changing conditions within the context of the study.

4. Confirmability: The ability to demonstrate that another could confirm interpretations made by the researcher. (Marshall & Rossman, 2006, pp. 201-203)
Working as an elementary school principal in San Diego qualified the researcher to identify appropriate subjects for this study. Given the lack of studies in the field, it can be assumed that the findings of this study would prove beneficial to those generalizing to similar studies.

Rich descriptions addressed dependability and confirmability aspects of data collected. In addition, the researcher used a professional service to transcribe interviews and requested certification of verification that the transcriptions were true and reliable (Appendix A). The researcher also compared transcriptions to field notes when analyzing and coding data.

Summary

This chapter outlined the research design and methodology for this study. This qualitative study used a purposeful selection of 10 elementary school principals across 10 different school districts in Southern California to describe current principal evaluation practices and principal leadership behaviors. Data were collected primarily through interviews with principals. Evaluation instruments and documents were used as a secondary source of data. Websites, demographic data, and field notes were added as additional data to both clarify and triangulate findings.

Participation was completely voluntary, and ethics were carefully considered. The seven-phase process to data analysis described by Marshall and Rossman (2006) guided the data analysis approach of this study. The trustworthiness of the data was addressed through the lens of credibility, transferability, dependability, and confirmability. Data gathered and results of the study are delineated in Chapter 4.
CHAPTER 4—FINDINGS

This study examined characteristics of principal evaluation methods that inform principal practices and sought to determine ways in which elementary principals perceive evaluation procedures to support improvement of their leadership effectiveness. The preceding chapters outlined the overall extent of the research, a comprehensive literature review, and an outline of the methodology used. This chapter presents the findings and data analysis from the research project.

This chapter is organized into sections aligned to the two overarching research questions presented in Chapter 1. It first reports findings on the status of current elementary school principal evaluation procedures. Specifically, it describes processes and implementation of practices and/or tools used in principal evaluation systems and reports the degree to which current principal evaluation practices align across districts. Secondly, it investigates how those evaluation practices align to research on effective leadership practices and additionally how elementary school principals perceive evaluation procedures influence or support improvement of their leadership effectiveness. Finally, the chapter presents unexpected, yet important, data of suggestions for improvement made by participating principals.

The bulk of data comes directly from principal interviews; however, data from evaluation instruments, documents, websites, demographic data, and field notes were also used in the data analysis. Data in each section are organized under distinct themes that emerged during the analysis including processes, tools, alignment, impact, trust, visibility, and feedback.
Status of Principal Evaluation Practices

This first section delineates the status of current principal evaluation practices as outlined by 10 elementary school principals across 10 different school districts in Southern California. Principals were purposefully selected from a defined population of current elementary school principals employed in Southern California schools. All participants had a minimum of 3 years experience as an elementary school principal, and demonstrated a willingness to participate in the study (Merriam, 1998). A convenience sample was generated through networking with other principals and colleagues known to the researcher. The researcher began the study by asking friends and close acquaintances meeting the criteria to consider participating in the study. Referrals to other principals were then generated through these initial connections. Participants first made personal connections with potential candidates. The researcher then followed up via email with candidates that expressed a willingness to be included in the study (Appendix B).

Face-to-face interviews were completed over a 7-week period between July and August 2011. Participants selected meeting locations. Interview locations included the office of the researcher, offices of participants, the home of a participant, and coffee houses. All participants signed an informed consent agreement (Appendix C) that the researcher then stored in a locked filing cabinet at the researcher’s home.

Interviewees represented mostly suburban districts and ranged in experience from 3-17 years. Interviews lasted between 15-49 minutes depending on the amount of detail provided by the interviewee. The researcher used a semi-structured approach combined with probing questions to elicit comparable data from each participant (Appendix D).
The researcher took field notes during the interviews to keep track of responses, as well as to make notation of other nonverbal communication for later reference.

Interviews were digitally audio recorded, and verbatim transcriptions produced 100 pages of raw data. Nine out of 10 participants also provided a copy of the principal evaluation tool used in their district. All documents collected were compared to interview data and identified themes emerging from interviews. Table 1 summarizes participant screening and interview protocol.

Table 1

*Participant Screening and Interview Protocol*

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Date of interview</th>
<th>Location of interview</th>
<th>District type and size</th>
<th>Years employed as an elementary school principal</th>
<th>Tool/document submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7/6/11</td>
<td>Participant home</td>
<td>Suburban, K-8, &lt;10 schools</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>7/6/11</td>
<td>Participant home</td>
<td>Urban, K-12, &gt;100 schools</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>7/8/11</td>
<td>Participant office</td>
<td>Suburban, K-6, &lt;10 schools</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>7/14/11</td>
<td>Researcher office</td>
<td>Suburban, K-6, 25-50 schools</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>7/14/11</td>
<td>Coffee house</td>
<td>Suburban, K-6, 10-25 schools</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>7/14/11</td>
<td>Coffee house</td>
<td>Suburban, K-12, 10-25 schools</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>7/24/11</td>
<td>Coffee house</td>
<td>Suburban, K-6, &lt;10 schools</td>
<td>8</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>8/3/11</td>
<td>Coffee house</td>
<td>Suburban, K-12, 25-50 schools</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>8/10/11</td>
<td>Participant office</td>
<td>Suburban, K-5, 10-25 schools</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>8/24/11</td>
<td>Participant office</td>
<td>Suburban, K-6, &lt;10 schools</td>
<td>17</td>
<td>Yes</td>
</tr>
</tbody>
</table>
By becoming immersed in the collected data, the researcher was able to organize the data and generate categories and themes (Marshall & Rossman, 2006). Data were broken down, synthesized into categories and subcategories, and then coded by themes. The process, tools, principal input, and implementation of practice were all compared. Comparisons served to determine the status of processes and/or practices used in principal evaluation systems, consistency across districts, and whether principal evaluation practices align to research on effective leadership practices.

**Processes**

All principals interviewed described a principal evaluation process consisting of both a written and face-to-face component. Nine out of the 10 principals described a defined process occurring either annually or biannually. Notably, out of the 10 principals interviewed, no two processes were the same. Furthermore, evaluation experiences described by principals indicated a significant range in systems across districts. Variations included formal to informal, vague to clear, formative to summative, and consistent to nonexistent.

Several principals described a comprehensible process of goal setting, meetings, and performance descriptors. Two principals reported being evaluated by the superintendent; three by an assistant superintendent; two by either the superintendent or an assistant superintendent; and three by a team consisting of the superintendent, assistant superintendent, and other district level administrators. Seven of the principals described their evaluation as a formal process, while the other three expressed that the level of formality depended on the evaluator.
Unlike teacher evaluations that often vary depending on years of service, almost all principals interviewed believed that the process in their district was the same for both new and veteran principals. Only one principal stated the process was different and explained that the only difference is whether the formal process occurs or not. She emphasized her point by stating,

It may also be based on just whether the supervisors at the time that they’re evaluating have the time to even do what their intentions are to follow through on their goals or intentions. We all know that we get busy during the year, so I sometimes wonder if that affects what the process looked like.

A few districts allowed principals to self-select goals, while others dictated goals to the principals. “We’re back in August, we have a general administrative council meeting where we’re given the tool for the year. We’re actually given two of the three goals. We’re told what our goals are,” explained one of the principals. Three out of the 10 principals did not mention goals at all. Meeting frequency also varied by district. Five of the principals stated that the process was intended to include a minimum of three meetings consisting of a pre, mid, and post review. Others described a series of informal conversations with a summative write-up at the end. Two of the principals received only a final summative report. Another two reported the process was started yet never revisited nor completed.

Some principals seemed clear on the principal evaluation process used by their district expressing things like, “there are certain expectations that I know we all have. Achievement needs to stay where it is in a high-achieving district.” Yet, others felt the process lacked clarity. For example, one principal offered, “There’s not a lot of clarity.
A lot of it’s very subjective, I think. You know for the superintendent, it could be, well, she bases a lot on teacher complaints.” Other participants described an intended process, but explained that the process does not always occur. One principal was not sure and responded, “I think the purpose is to just to do it, to just go through the process because it’s something that’s expected and that’s it.”

In a few cases, participants discussed trying to manipulate the process. For instance, one principal described how principals in her district try to control the outcomes of the evaluation: “We kind of do a pinky swear. Don’t bring a binder. Okay. Are you bringing a binder? No, I’m not bringing a binder. So we all talk about it. And then whoever goes first, calls everybody else.” Another principal lamented putting personal reflections in writing that then ended up in her personnel file: “And so I learned from that. I’ll talk all I want, all I need to about what my areas of need are, but I shouldn’t put those in writing.” Still another principal found that her performance was tied closely to points on a rubric. To earn the points and receive a higher evaluation, she did the following:

After the first year when I didn’t get as many points as I could have in some areas, I created the hugest binder you can imagine to prove to my other boss that I was doing all this stuff because I felt she didn’t know what I did. And guess what? The binder was never looked at.

**Purpose.** When asked the purpose of their evaluation, principals had differing responses. Some felt that it was a growth model, others cited accountability, and some believed it was to give feedback and determine next steps. In fact, many principals seemed uncertain about the purpose of their evaluation. Almost every participant used a
phrase such as “I think” when offering their perceptions as demonstrated in the following statements:

Good question. What is the purpose? I think the purpose in our system currently is to keep us principals accountable for academic achievement.

Well, I mean, I think that the purpose of going to a tool was to be more succinct and to hold everybody to the same standard and be more objective maybe. I don’t know what the purpose overall is. I know that for me being evaluated just makes me hear that I’m appreciated.

It certainly doesn’t measure growth. I don’t feel like that’s the gist of the conversation. I think determine support needed is a strength of our tool.

Well, I do think it’s to provide feedback, because we do get feedback about our performance, and I don’t think, okay, how can I say this . . . I don’t think we get a lot of ongoing feedback unless it’s something bad.

I think it’s a growth model. I feel very supported as a principal in our district.

Probing into those responses, the researcher found that four principals felt the purpose of their evaluation was for accountability purposes, two felt their evaluation processes were intended to foster principal growth, and two believed the process was aimed at determining support needed. Participants were asked if they felt their evaluation was intended to measure leadership. Approximately half said no. The other half felt that the process somewhat measured goals, gave feedback about leadership, or could potentially measure leadership if the process was completed.

In addition to feeling uncertain about the purpose of their evaluation, most principals were not exactly sure what their role was in the process either. Some felt it
was to fill out the forms, others to listen to feedback, and still others to come to the table with ideas. Several of the principals saw themselves as active participants in the process. Three participants conveyed the need to be honest with themselves. Three others felt it was their responsibility to track data, collect evidence, and come prepared to meetings. Two principals receiving only a year-end review felt their role was to be a part of the discussion and “toot their own horn.”

**Feedback.** When describing their evaluations, half of the principals felt that the process was summative with formative components. Three described a summative only process, one formative, and one neither. “I think there’s formative along the way with our discussions, but I think the written evaluation at the end is summative,” described one principal. Other principals made similar comments, such as, “It’s summative in that we get a written summary of it, but I think it’s more formative as we have the conversations throughout the year about what we’re working on and where we’re going.” Even the principal who described the process as formative reflected, “We have a midyear meeting so that we discuss our progress towards the goals and bring data to support that, so it’s more formative and the process is new.”

Digging deeper into how principals reflected on receiving feedback from their evaluation experience, the researcher noted the following:

- Two out of 10 principals received only summative feedback.
- Four out of 10 principals received summative and limited formative feedback.

Two of the four were told that the postconference would occur in the subsequent fall after the previous year’s state test scores were received.
• Two out of 10 principals described the process as an ongoing, informal, “organic” process with a summative write-up at the end.

• Two out of 10 principals received neither formative nor summative feedback. Participants reported that they received feedback from their supervisors in other ways; however, most did not feel it was tied to their evaluation performance. Informal conversations, emails, compliments, and complaints were some of the ways principals stated they receive information about their performance. Six principals mentioned their evaluators also visited their schools and walked through classrooms with them; however, most participants did not give the impression that this was a regular occurrence. Three principals also cited data discussions as another way of sharing information with their supervisors, yet they were not held accountable to the data in the end. Exemplifying this point, one principal elaborated:

  We have data meetings after each benchmark with other peers, with our superintendent, assistant superintendent. We have coaches that come to our schools, mentors that are also looking at our coaching logs to see how often we’re in classrooms and giving feedback. So, there is some accountability in that way. It’s just that it’s not necessarily . . . it hasn’t been addressed in a formal evaluation meeting.

Only two principals felt they did not receive any feedback outside of their evaluation.

**Perceptions.** Flexibility and adaptability within the principal evaluation process received mixed reviews. Only two participants felt there was complete flexibility within the process. The first principal commented:
I have a lot of flexibility in it, especially with the new superintendent we have. It pretty much was open-ended. Open-ended in the sense that I could set my goals, but within the parameters of—here’s what the data says [sic]. So I don’t ever feel like goals are imposed on me.

The other principal remarked:

I would say there’s quite a bit of flexibility, just because I don’t know that they come with [goals]. I’ve never been in a situation where they’ve mandated the goals that I’m supposed to be working on. It seems like it’s a collaborative process.

Both participants believed their evaluations were adaptable based on the needs of the principals; however, one was not quite sure, due to inconsistencies in the evaluation process.

Three other participants experienced limited flexibility in their evaluations, while five reported there was really no flexibility at all. Where limited flexibility existed, principals reflected on having had some input into goals as a way in which they felt they could contribute. Of the five citing no flexibility, three explained that the goals were prescribed for them, while two were not given goals to work on at all. These five were not sure if their district’s process was adaptable to their needs, since the process was highly structured for them.

The researcher noted that most participants seemed surprised to consider what training their supervisors might have had to conduct principal evaluations. In fact, none of the 10 principals had any idea how their supervisors were trained to evaluate
principals. “My evaluator has never even been a principal,” responded one participant, then added, “so I think it’s fly by the seat of your pants.”

Three others assumed their supervisors most likely relied on their past experiences, and one suggested, “I don’t know other than maybe prior experiences of principal themselves and then knowing and applying what they’ve done as supervisors.” Two additional principals offered the thought that perhaps their training on how to evaluate teachers is what prepared their supervisors to evaluate them. “I think it’s the same training we receive as principals to evaluate teachers. Not a huge amount,” deduced one of the study participants. Another surmised that perhaps her supervisor had read the manual as part of her training, since the evaluation was purchased as a packaged process. Researcher field notes also reflected that when asked if their district had a board policy or an administrative regulation on principal evaluation, most of the principals gave a nervous laugh. Four of the principals expressed that they did not know. Five principals assumed that they must, but weren’t sure. “I would imagine there’s a board policy there, but have I ever really looked at it? No. That’s actually kind of funny,” was a typical type of response given by the principals interviewed. One principal confidently stated that her district did not.

As referenced in Chapter 2, the California Education Code (2011) requires the evaluation of school administrators, but it does not specify the methods required to do so. A website search of board policies from each of the principal’s districts unveiled that none of the districts studied had a board policy or administrative regulation specifically addressing principal evaluation. Three districts made general references to administrative evaluations, two districts delineated that the evaluation of all personnel was at the
discretion of the superintendent, two made no mention at all, and two districts’ board policies were unavailable. Of the 10 districts, it is noteworthy that 3 have distinct board policies for the evaluation of the school board, 6 have superintendent evaluation policies, and 5 have evaluation policies for teachers. Table 2 outlines evaluation board policies as described.

Table 2

Board Policies on Personnel Evaluation

<table>
<thead>
<tr>
<th>District</th>
<th>Board evaluation policy</th>
<th>Superintendent evaluation policy</th>
<th>Principal evaluation policy</th>
<th>Teacher evaluation policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>“Administration”</td>
<td>“Administration”</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>“Superintendent-All Personnel”</td>
<td>“Superintendent-All Personnel”</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>✓</td>
<td>“Administration”</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓</td>
<td>“Superintendent-All Personnel”</td>
<td>“Superintendent-All Personnel”</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>“Administration”</td>
<td>“Administration”</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Board policies not available through public web access.

**Implementation.** The majority of principals went on to describe inconsistencies in the implementation of the process. Eight of the principals interviewed described the process in their district as changing from year to year, or changing depending on the evaluator completing the tool. “One never knows because in the 4 years I’ve been in this district, it’s been different all 4 years,” stated one principal. Another principal with
16 years of experience at her school commented on only having been evaluated twice in her tenure, and each was with a different superintendent. A third principal reflected, “The process, I think it changes depending on who the superintendent is. There’s not a real defined timeline for evaluation.” And yet another explained that she had been a principal for 6 years and commented, “I’m pretty sure that in the 6 years, I’ve had a different . . . superintendent each year.”

Participants were asked to expound on the degree of implementation of their district’s principal evaluation processes. Only two felt that the process was implemented fully as they described. Other principals explained that the process was not implemented with fidelity, consistent between evaluators, nor in some cases even completed. Two principals even guessed that the process was most likely different for them than it was for other principals in the district:

I don’t really know what it looks like with the other principals. Like I know how my evaluations always feel and what the process is with me, but I honestly can’t say if that looks different with principals that have a different success rate, because I know that I’m one of the more successful principals in the district and that even the school board trusts me to take some risks that they may not trust some of the other principals to take.

Though they had differing thoughts, the principals offered several reasons for this inconsistency of implementation. Time was a prevalent response. Either there wasn’t enough time available, or the supervisor ran out of time to complete the evaluation process. A principal in a medium-sized district sympathized with her evaluator:
She had to evaluate 31 principals and meet with them twice, plus do all the other things that were going on throughout the district, I just, I don’t think it was a realistic, I don’t think it’s a realistic expectation for her to do it all in the way she wanted to, nor do I think it fit into . . . it wasn’t top priority.

Comments encompassing other reasons for inconsistency in implementation of the principal evaluation process such as clarity, priorities and personalities of evaluators, and accessibility were also offered:

There was a lot of discrepancy. So that showed me that there’s just not clarity and there’s not equity either.

One of the three . . . leaders takes it very seriously and follows it step by step, and another one, possibly two, doesn’t and implements it, but not with fidelity and I do feel that although it was more work the year that I was asked to do it with fidelity, I did feel like I had more feedback, I had more growth; I had more understanding of where I needed to head, where I needed to put my energy the next year.

It’s pretty much relegated to the two associate supts, but it’s an interesting question because they are like polar opposites in their styles, so I would not be surprised if it looks different for the other people that are being evaluated by the other supt.

I know that he has those conversations with everybody, but my gut is he’s at my site more only because it’s close because he can walk down from his office instead of having to get in the car and go somewhere else.
The principals added that there was not much accountability to their evaluations either. Aside from self-accountability which three principals mentioned, most principals interviewed felt there was limited or no accountability. A few principals talked about high expectations but did not explain what would happen if those expectations were not met. Others discussed assistance plans and wondered:

I think if there’s glaring things that are happening, then they try to hold you accountable. Like if you have maybe tons of parent complaints, then maybe they’ll hold you accountable for that, or if your test scores are just always really, really low, then they’d hold you accountable. But as long as, I kind of feel that as long as things seem to be going smoothly, then okay good, we don’t have to worry about that school.

Yet, even with the thoughts that there might be higher accountability for underachieving principals, none of the principals referenced anyone losing their position due to poor performance. In fact, a principal reinforced this by offering the following thought:

I think that as much in teaching as in administration that we need to be looking at, there are some people who get hired into the positions that are good at what they do, there are some that are not and that this process should, in some way, shape, or form, lead a district into knowing whether or not a person is satisfactory or not and should continue in the position. As told the day I started my administrative credential that you do not have tenure when you’re an administrator, you can be released at any time for any reason. I know very few principals who are released.
Instruments

The researcher also discovered variations among instruments. In this study, four principals described their district’s tool as predominantly narrative, two as rating scales or rubrics, and four a combination of both. Four district tools had goal setting components; however, seven principals also discussed goal setting as part of the overall process. Eight of the nine district evaluation tools collected were self-developed tools, or what the principals perceived to be self-developed. “I think this was self-developed by the district. I also know that a couple of years ago, they were working on a new principal evaluation system. I’m not really sure what happened with that,” and “I feel like it was self-developed and my perception and impression of that when it was given to us is that it had been the task of the new assistant superintendent of HR to develop it and to write it or create it,” were typical statements about the tools. Only one principal did not submit a tool; however, she indicated that she believed it was also self-developed.

None of the principals reported being a part of the development of the tool, but one principal reported having input into the process. Several principals had the impression that their evaluation tool encompassed components of evaluation tools from other districts. For example, a principal offered a thought that her supervisor “may have taken bits and pieces from other districts.” Another principal speculated, “I think my assistant superintendent kind of took ideas from the county and revised it.” Table 3 outlines the components included in each district’s evaluation tool.

**Description of tools.** To reinforce interview descriptions of current principal evaluation tools, the researcher requested instruments from each of the participants. These documents served a specific purpose of learning more about principal evaluation
### Table 3

**Principal Evaluation Tool Components**

<table>
<thead>
<tr>
<th>Tool components</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation Philosophy</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directions-Tool Completion</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluator Comments</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Self-Reflection</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rubric/Rating Scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rubric/Rating Descriptors</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goals</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timelines</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Expected Behaviors/Standards</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Measurement of Progress</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Resources/Support Needed</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Academic Data</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conference Indicators</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accomplishments/Strengths</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Areas for Growth</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assistance Plan</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Methods across districts (Merriam, 1998). Table 4 provides codes used in the presentation of data in this section.

One principal did not submit the tool used for her evaluation, though she described it only as a narrative write-up from her supervisor with specific comments about her performance. “The superintendent we have now is not a big form person, and so the whole conversation was pretty much write it up however it works for you,” she explained.
While no two principals described the same process, it was interesting that two of the principals submitted virtually the same tool. Tools B and H were individualized for their respective districts, but contained almost all of the same components. Forms were broken into three leadership components: instructional, operational, and working with parents and community members. Overall, the only major difference between the two tools was the discreet descriptors under each of these three categories.

More specifically, these forms from principals 3 and 9 encompassed rating scales followed by narrative summaries. Directions provided to the evaluator specified to rate the evaluatee on each element within the leadership component and then on each leadership component overall using rating scale descriptors of meets standard, requires improvement, and unsatisfactory. Additionally, the evaluation included sample indicators of effective performance for each component. The final page of each of the tools offered a summary of the evaluation and a place for evaluator and evaluatee comments. Tools B and H also delineated a timeline for completion for administrator receiving a release of service notice and for administrators demonstrating successful performance.

Of the two principals submitting this evaluation form, only principal 3 described it as the current tool. She explained that the tool was used differently depending on who the evaluator was completing it:
I think it depends on the person who’s evaluating, because I have one area superintendent who did include some data in here. She included charts on our attendance and our literacy and math benchmarks, but she’s the only one who did that. The other ones basically checked off the boxes meet standards and that was that.

The other principal, principal 9, submitted a tool dated 2009-2010, but gave the impression that the evaluation process in her district had changed:

So last year, this was brand new; our assistant superintendent, who was my direct supervisor, met with each principal individually in the fall, September, October and sat down and talked to us about what we thought three goals would be for our work for the year and they could be operational, instructional, a combination of it, just depending on schools’ test scores, our personal leadership styles. . . . It was really stuff I wanted to do anyway that year so I just figured let’s roll it up and make goals that make sense for what I’m going to do this year. And we actually typed them out. So, we typed them out and then we were just working on them all throughout the year. We never came back to them all year.

Likewise, principals 5 and 6 also submitted similar tools for review, yet neither of the two principals described the same process as the tool submitted. Tools D and E appeared to have similar components to the tools B and H, but were much more developed. Both tools D and E were designed as growth models and included supporting documents in excess of 20 pages.

Tool E specifically aligned to the CPSEL standards. Principal 6 noted that the principal evaluation system was based on the teacher system and that the principal forms
were only slightly different. The growth plan comprised a template for goal setting, an area for an action plan (steps, timeline, and resources), and an assessment section. A place to record notes from a mid-year goal reflection conference and an observation date were also part of the template. The handbook included separate forms for evaluatees’ midyear and end-of-year goal reflections, as well as instructions of how to complete all forms. An overall summative evaluation for the evaluatee rated each CPSEL indicator as meets standard or does not meet standard. When describing the tool, the principal noted:

And, again, because it came out of the teacher piece, that’s how they all kind of came up with the forms and identified it that it be this collaborative way with little, a section where you put how you’re going, how you yourself would measure your success towards reaching your goal and say it’s kind of self-directed. So, that’s how they initially came up with it and, essentially, the administrator tool is the same. It’s the same form.

Principal 5 also had a growth model tool and submitted document D with a handwritten note that stated, “Here’s what I found—not bad, but what I rec’d in my eval. was not anything from this packet.” The packet included the purpose of the administrative evaluation system, components and cycle of the evaluation, six proficiencies for instructional leaders (with a total of 65 indicators of success), a district/action plan review, and a professional growth plan. Templates highlighted areas for rating evaluatees and for recording narrative comments. The instrument utilized the same rating scale descriptors as the same as the previous tools (meets standard, requires improvement, and unsatisfactory), yet this document also included a rubric outlining the performance levels. Finally, an assistance plan form completed the packet.
Most significantly, tool D detailed the purpose, beliefs, and expectations of the principal evaluation system. As referenced in the literature, in the historical context of evaluation, evaluation was designed to measure results, identify strengths and barriers, and make recommendations for improvement (Guerra-Lopez & Leigh, 2009). This document outlined all of these components in the philosophy of the system. Nevertheless, principal 5 was not required to participate in the formal administrative evaluation system. She described her personal experience:

We’re given a date to meet, and we meet with our superintendent and our assistant superintendent of educational resources. The meeting takes place in person in the superintendent’s office and, to be honest, in 4 years of administrative experience, I really feel like last year was the first year I truly had any paperwork provided to me for my evaluation. . . . It was narrative, a short paragraph this last time. That’s the only time I can actually honestly tell you, because it’s the only time I’ve received a written document.

Other principals described evaluations that were more goal focused. Tools A and C included areas for three goals: a district or academic plan goal, a 21st century school or building goal, and a personal goal. Principals were expected to propose goals, develop action steps to accomplish the goals, and determine possible methods to measure their success. In theses goal focused tools, there was also a section for principals to discuss or stipulate support or resources needed from their supervisors. Tool C also included an opportunity to list timelines, accomplishments, and areas to grow. Evaluatee and evaluator reflections were attached separately to the tool at the end of the year.
Principal 2 noted that she felt the tool was effective when it was used; however, the use of the tool depended on the evaluator:

And then this year I did the same . . . process, but it was very different. A different person was the chair . . . and we discussed personal goals, but were not asked to put anything in writing. We met three or four times throughout the year and went and observed at multiple schools. Talked a lot throughout the year, but did not have an end-of-the-year culmination meeting or anything, and we were not asked to write any kind of summary of our year and were not given any reflection back on how we did for the year.

Principal 4 also submitted a goal-focused form, yet could hardly remember the tool enough to describe it:

We don’t really have a . . . well . . . do we have a tool? We have a form and that’s basically it. The form is something that was developed in-house and, basically, it just has what your goal is and, typically, we have at least three goals I would say. That’s probably the average. So you have a place where you write what your goal is, why you’re choosing this goal, and the steps you will use to achieve the goal, and that’s pretty much it.

Tool F offered a slight variation from the type of tool described above. This form specified predetermined goals. All three goals for principals focused on student achievement, instructional strategies, and teacher collaboration. For each of the goals, a template was provided for each principal to record a summary of data, description of process, objectives, evidence, an action plan, and support needed. A specific list of academic data outlined performance indicators to be collected. General expectations for
site administrators were listed on the front page, while instructional resources were
attached to the back. Principal 7 had the following to say about it:

It’s very narrative. We write the goals and then we respond to the goals in the
midyear as a narrative. Primarily, I will have to say that it’s the conversation. It’s
the meeting with the superintendent and the assistant supt. We pretty much
control the conversation. I’ve spoken to other principals and we go in there and
just kind of speak to the goals, speak to what’s happening at our sites. There’s a
place, one of the components is support needed, which is nice. So we can talk
about . . . we might need more release time or articles.

The final two instruments, G and J, were summative tools completed at the end
of the year. Both tools included a series of various leadership areas with elements or
descriptors listed under each area. Tool G listed educational leadership; school culture;
effective management; community relations; personal and professional leadership; and
professional development and staff effectiveness as the areas of feedback for principals.
Principals received a bulleted list of observable behaviors they individually demonstrated
during the year under each of these expectations.

Tool J broke the appraisal down in the areas of personal characteristics; leadership
characteristics; communication ability; management ability; and professional growth and
self-improvement. Several descriptors were listed under each area with over 30 total
descriptors on the tool. Principals received a rating of outstanding; good; acceptable;
needs improvement; unsatisfactory; or not rated for each descriptor. No quality indicators
were included to describe the areas on either tool.
Reliability of instruments. It is uncertain if any of the principal evaluation instruments collected during this study had undergone measures for reliability or validity. However, it is unlikely since 9 of the 10 instruments were believed by participants to be district-developed. Principal 8 indicated that she thought instrument J was purchased, but it had no accompanying documentation of its genesis. Furthermore, no principal participants were aware of any research tied to their district’s evaluation tools. None of the principals cited the 21 leadership responsibilities as identified by Marzano et al. (2005), nor any other research on effective leadership behaviors.

Despite not knowing if their evaluations encompassed any research-supported elements, many of the study participants articulated that they felt expectations of principals were clear. Of the seven that responded affirmatively, two cited that they knew the expectation based on what was included on the tool. A few principals assumed that the tool itself might be aligned to standards for educational leaders. A deeper look at the district instruments that included standards or expectations for site leaders revealed that four were very tightly aligned to the ISLLCS/CPSEL standards, one aligned to the Six Dimensions of Leadership, one promoted behaviors aligned to the theory of instructional leadership, and one did not seemed tied to any identified standards or research in educational leadership. Yet, only one principal was aware of any standards and specifically referred to the California Professional Standards for Educational Leaders.

Another participant explained that the superintendent distributed and communicated expectations for principals. “My assistant superintendent has literally shared, it’s literally called the expectations for principals. It’s a document that has and is grouped into four areas, which are district focus areas: high-quality teaching and learning,
data analysis, communication, and collaboration.” Yet, these expectations were not reflected in the evaluation tool. Still another participant expounded:

I would say that the paper version of expectations are there. They cover about 1/1000th of what I just said I do in my job description, because it’s very much instructionally driven and, because my boss has never been a principal, I don’t know that she can name that whole list.

Lastly, five of the principal evaluation instruments analyzed had some sort of rubric or rating scale included as part of the tool, yet only two of the five incorporated rubric or rating scale descriptors. Principals did not comment on whether these rubrics or rating scales were reviewed with them before their final evaluation. Eight principals also referenced data or evidence as part of the component of the tool, yet only two of the forms collected specified actual data to be collected. For instance, one participant reflected:

Our superintendent that we have right now is very data driven, but it’s not something that you’ll be . . . it’s not necessarily the evaluation piece. Like, he wants to know where are you as a baseline and how do you plan to improve on that, but it’s not specific you have to hit this certain target. And even if, so like if you’re looking at the achievement data and it’s up and down, he just wants to know that you know why, be able to explain what’s going on with the numbers and it’s not about the numbers themselves.

Changing instruments, lack of research, variations between evaluators, and consistency of implementation were all considered in the data analysis. Overall, there were few reliability measures. The next section will outline the alignment, influence, and
overall perceived effect on leadership of the aforementioned principal evaluation
processes and instruments.

**Perceived Effects on Leadership**

The literature review underscored the far reaching duties of the elementary school
principal, and scholars have conducted many studies linking effective leadership with
student achievement (Catano & Stronge, 2006; Marzano et al., 2005). During interviews,
the researcher asked principals to think about their professional duties in three ways: what
activities did they engage in on a regular basis that they felt had the greatest impact on
student achievement, how did they spend the majority of their time on a typical school
day, and what other tasks were they required to perform as principals.

Each principal reported two to four key responsibilities that they felt they engaged
in on a regular basis that had the greatest impact on student achievement. Many shared
the same responses resulting in only nine behaviors in all. The most common answers
were work with teachers and provide timely feedback; visit classrooms to observe
teaching and learning; and investing time to get to know students personally.
Approximately a third or more of the participants also mentioned building relationships,
attending or designing professional development, and tracking data. Other principals also
felt that it was important for them to ensure student safety; keep up on professional
reading; and identify support and resources for teachers.

Describing a typical day was a bit more difficult for the principals. “What is a
typical school day? There is no such thing,” exclaimed one principal. Another
responded, “The majority of my time, wow. You know what? It depends on the time of
the year.” In all, study participants came up with a total of 13 different behaviors that
they felt they engaged in daily. Of the 13, almost every principal again noted classroom visits in conjunction with supervision and safety. Meetings, paperwork, correspondence, and other operational procedures were among the other responsibilities mentioned. One principal best described the complexity of the job with this response: “I couldn’t even explain it because there’s so much to it. A typical day, I know there’s a million things I’m supposed to do. Pull out that job description; it will show you.”

Asking principals to list all other duties that they were responsible for was almost an impossible task. Principals offered long lists of tasks including planning, budgets, evaluations, coordinating with support staff, attending meetings, data analysis, professional development, facilities, scheduling, and student discipline. “Everything under the sun like that,” concluded one of the study participants. Over 26 different tasks were brought up during the 10 different interviews, yet the researcher noted that principals most likely could have given more with additional time and prodding.

**Alignment**

The researcher also asked principals how well they felt the evaluation tools and practices of their district aligned to the leadership practices that they reported engaging in. Three felt unsure or expressed that there was very little alignment. Three of the four principals whose evaluations were based on the CPSEL standards felt that there was some overlap, however did not encompass many of their daily duties. As described by two principals, “I mean certainly some of the components and elements definitely align to the things that I’ve decided to do here at the school, but I think that’s by accident,” and “I think that the system itself meaning that the tool could probably work, I just think that the fact that they give us these goals to meet—it doesn’t make it very meaningful.”
Other principals with more informal district processes found that the evaluation helped to give them focus, but again did not fully align to their everyday leadership. One principal stated:

Well, I think that by having to set goals specific to our sites, like setting a goal for morale. It really had me talk about, look at research, and measure ways to improve morale . . . still working on it. In fact, my secretary gave me a book this summer called *The No Complaint Rule* to read. But by having me really focus on that, by improving morale, it’s going to make the school a better place and the school being a better place for kids, and teachers, and parents is part of the goal of a good leader, I believe.

Another principal elaborated,

In some ways, it aligns because the first goal is so encompassing, student achievement in subgroups. You’ve got to be paying attention to special ed. You’ve got to be paying attention to your English language learners. You’ve got to be paying attention to low socioeconomic . . . . So in that regard, it’s aligned because I’m going to be answering to their results, how their student achievement is, but you know, it’s not the work you do every day.

Another discrepancy was noted between the processes described by the participants and the tools submitted for review. For instance, several principals mentioned goal setting, although goal setting was not a part of the submitted tools. Conversely, another participant stated that she talked about goals with her evaluator, but was not asked to write down anything. The tool, however, was based on documentation of three agreed upon goals. Still others described the process, or lack of process, that they
experienced, yet turned in a very complex tool that was not used in their particular evaluation.

In analyzing the evaluation instruments, the researcher noted a lack of alignment between principal responsibilities described by participants during interviews and what is actually measured by the evaluation instruments. Several tools were aligned directly to the CPSEL standards. Of the four districts with evaluations closely tied to these standards, all of the principals alluded to engaging in more responsibilities than were identified or measured on the tool.

Tables 5-7 compare Marzano et al.’s (2005) 21 leadership responsibilities with district evaluation tools and principal responses concerning what responsibilities they felt they engaged in regularly. One caveat to note is that the evaluation descriptors noted on tools and specific activities mentioned by participants were open to interpretation by the researcher and therefore provide somewhat of a subjective alignment. Nonetheless, this loose comparison provides insight as to the priorities of districts versus the participating principals priorities, both in comparison to the past research findings. In addition, it highlights variations across districts.

Table 5 demonstrates how often each of the 21 leadership responsibilities aligned to the district evaluation tools. Evaluation tools for Districts 1, 2, and 4 are based exclusively on goal setting. Therefore, these goal-oriented evaluations could encompass more of the 21 responsibilities depending on what goals are written by the principals. For example, if a principal wrote a personal goal to more often recognize and reward staff accomplishments, that particular evaluation could fall under contingent rewards. Also, the principal for District 1 did not submit an evaluation tool for analysis; therefore, the
**Table 5**

*Comparison of 21 Leadership Responsibilities and District Evaluation Tools*

<table>
<thead>
<tr>
<th>21 leadership responsibilities</th>
<th>District instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Affirmation</td>
<td>✓</td>
</tr>
<tr>
<td>Change Agent</td>
<td>✓</td>
</tr>
<tr>
<td>Contingent Rewards</td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>✓</td>
</tr>
<tr>
<td>Culture</td>
<td>✓</td>
</tr>
<tr>
<td>Discipline</td>
<td>✓</td>
</tr>
<tr>
<td>Flexibility</td>
<td>✓</td>
</tr>
<tr>
<td>Focus</td>
<td>✓</td>
</tr>
<tr>
<td>Ideals and Beliefs</td>
<td>✓</td>
</tr>
<tr>
<td>Input</td>
<td>✓</td>
</tr>
<tr>
<td>Intellectual Stimulation</td>
<td>✓</td>
</tr>
<tr>
<td>Involvement in Curriculum,</td>
<td>✓</td>
</tr>
<tr>
<td>Instruction, and Assessment</td>
<td></td>
</tr>
<tr>
<td>Knowledge of Curriculum,</td>
<td>✓</td>
</tr>
<tr>
<td>Instruction, and Assessment</td>
<td></td>
</tr>
<tr>
<td>Monitoring and Evaluating</td>
<td>✓</td>
</tr>
<tr>
<td>Optimizer</td>
<td>✓</td>
</tr>
<tr>
<td>Order</td>
<td>✓</td>
</tr>
<tr>
<td>Outreach</td>
<td>✓</td>
</tr>
<tr>
<td>Relationships</td>
<td>✓</td>
</tr>
<tr>
<td>Resources</td>
<td>✓</td>
</tr>
<tr>
<td>Situational Awareness</td>
<td>✓</td>
</tr>
<tr>
<td>Visibility</td>
<td>✓</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
</tr>
</tbody>
</table>
comparison was based solely on the description given by the principal during the interview. It is possible that the tool itself encompassed more than was reported.

Overall, most district tools seemed to incorporate some aspect of focus. Evaluation documents reflected that principals were responsible for establishing clear goals for curriculum, instruction, and assessment; and promoting high expectations to meet those goals. Principals were also expected to focus efforts on advancing the goals of the school. Other heavily emphasized responsibilities for principals were developing culture; providing intellectual stimulation; involvement in curriculum, instruction, and assessment; and maintaining order. Districts providing standards or expectations for principals had a higher alignment to the 21 leadership responsibilities than those only focused on goal setting.

All of the 21 leadership responsibilities were found in more than one tool studied except for discipline, flexibility, relationships, situational awareness, and visibility. Only one tool submitted included these five behaviors; however, the principal interviewed reported that she never received the tool as part of her evaluation. None of the district instruments studied made any notation of contingent rewards.

Table 6 demonstrates how principal responses to the responsibilities they felt they engaged in regularly aligned to each of the 21 leadership responsibilities. The researcher determined an overall percentage rate of how often each of the 21 leadership responsibilities were referenced. Rough estimates and loose interpretations provide a feel for the kinds of behaviors that principals most focus on during the course of their duties.

Table 6 clearly indicates that visibility was a focus for participating principals. Other top cited responsibilities were communication; monitoring and evaluating;
Table 6

*Comparison of 21 Leadership Responsibilities and Principal Responses*

<table>
<thead>
<tr>
<th>21 leadership responsibilities</th>
<th>Number of times cited by principals</th>
<th>Percent of total responsibilities cited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affirmation</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Change Agent</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Contingent Rewards</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Communication</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Culture</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Discipline</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Flexibility</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Focus</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ideals and Beliefs</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Input</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Intellectual Stimulation</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Involvement in Curriculum, Instruction, and Assessment</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Knowledge of Curriculum, Instruction, and Assessment</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Monitoring and Evaluating</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Optimizer</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Order</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Outreach</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Relationships</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Resources</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Situational Awareness</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Visibility</td>
<td>20</td>
<td>14</td>
</tr>
</tbody>
</table>
outreach; and resources. Comparing Tables 5 and 6, the researcher did not find any overlap in the five heaviest represented responsibilities on district evaluation tools compared to the greatest emphasis principals placed on their regular behaviors.

Table 7 compares the top five leadership responsibilities that Marzano et al. (2005) found to have the largest correlations to the top five emphasized responsibilities of district evaluation tools and to principal responses.

This comparison indicated that there was no overlap between the emphasis on district tools and Marzano et al.’s (2005) top 21 leadership behaviors with the largest correlations to student achievement. There was also no overlap between district tools and principal reported behaviors. Conversely, there was overlap in two areas between the most common responsibilities as reported by principals and the top leadership behaviors as identified in the research. These two areas were monitoring and evaluating; and outreach.

**Influence**

As previously discussed in Chapter 2, Marzano et al. (2005) identified 21 specific leadership responsibilities that have the greatest impact on student achievement. Using publicly accessible test score data from the state, the researcher ran a 3-year trend on each of the principals in the study. The intent of looking at the student achievement data was to determine if any principals experienced significant gains in student achievement and to compare those data to specific principal evaluation instruments. Because the state academic test scores are only one measure of academic achievement and are not statistically correlated to the evaluation tools, the results merely suggest relationships between evaluation tools and student achievement (Figure 1).
Table 7

Comparison of 21 Leadership Responsibilities, District Evaluation Tools, and Principal Responses

<table>
<thead>
<tr>
<th>21 leadership responsibilities</th>
<th>Responsibilities with largest correlations</th>
<th>Most common responsibilities found on evaluation tools</th>
<th>Most common responsibilities cited by principals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affirmation</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Change Agent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contingent Rewards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Culture</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Discipline</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flexibility</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideals and Beliefs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intellectual Stimulation</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement in Curriculum, Instruction, and Assessment</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of Curriculum, Instruction, and Assessment</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring and Evaluating</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Optimizer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outreach</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Relationships</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resources</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Situational Awareness</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visibility</td>
<td>√</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. State API trends by principal.
Of the 10 participants in the study, only 3 schools experienced an upward trend of test score data over a 3-year period. Three schools had fluctuating test data, while four actually showed an overall decline. Principals 5 and 9 had overall upward trends, yet expressed that they did not receive an evaluation on the instrument they submitted as part of this study.

Principal 7 also had an upward trend in student academic achievement. This participant revealed in her interview that principals received predetermined goals from the district for their evaluations. All three goals for principals were related to student achievement, instructional strategies, and teacher collaboration. In addition, the district instrument included data, principal expectations and instructional resources. Researcher notes, however, indicate that principal 7 viewed the evaluation process as merely a requirement rather than a benefit for leadership improvement. She illustrated this point with the following comment:

Well we don’t get an outcome of our evaluation. We have our meeting and then summer comes and we don’t even get, we don’t get a summary. We don’t get a synopsis. We don’t get an at-a-boy. It’s just we have a meeting and then we start the process over. So, I feel as if I make decisions based on what I think my kids and my teachers need, not what the superintendent said in June.

**Benefits.** Principals were asked how they felt their evaluations benefitted or supported them in their professional responsibilities. Three of the principals felt the evaluation was an opportunity to have conversations and receive feedback from their supervisor. Other principals felt the amount of benefit of the evaluation stemmed from principal motivation and individual emphasis placed on the evaluation. Still others
commented that trust in the evaluator directly affected how supported they felt in the process. For example, one participant commented:

Yeah, again, it depends on who the person is. Over the years, I’ve had a couple of . . . superintendents who I most definitely would not have sought feedback from them because I was not comfortable doing that. They had not worked on building a relationship with me so I didn’t know where I stood. But I’ve had a couple of others that did bother to build a relationship with me. I knew where I stood and I was comfortable going to them. I didn’t feel like I was admitting failure to them if I asked them a question. I felt safe.

The findings overall indicated that only 3 out of 10 of the principals could find any perceived benefit to their evaluations. Seven out of 10 of the principals did not feel that the evaluation tool itself lent any benefit for leadership improvement. Many of the comments made by principals, such as this participant, indicated that improvement came from within:

Generally, my evaluations have all been very positive where I’ve left feeling that they appreciate the work I’m doing and value me as a leader. There hasn’t been much of a focus on where I could improve, and I certainly know I can improve. So that hasn’t been a big piece of it, though maybe it’s because I think a lot of effective leaders are always reflecting themselves on where they can improve and generally have a sense of what they could be doing differently or areas of growth. So, maybe that’s kind of more inherent to what we know about ourselves.

When asked directly in what ways did participants feel their district’s evaluation system influenced their leadership practice, seven reported none at all. One principal felt
that her evaluation had minimal influence because it was successful; however, she clarified, “If it was bad, I probably would . . . I probably might make a few changes.” Two principals reported that they felt a greater awareness of their actions as a result of the feedback from their evaluations. In the words of the first principal, she reported proceeding with more caution:

I feel like you’re not given permission to make people unhappy or uncomfortable, and you’re not given that, also that opportunity to make mistakes because we all make mistakes. I feel like I let my teachers make mistakes. It’s okay; it’s not the end of the world. That’s how we learn. If you’re not taking risks, but yet as principals I feel like if you make a mistake, you’re in trouble and so it makes you gun shy to make those mistakes; just get out on a limb and do that . . . . And so I think it makes me more cautious and sometimes less willing to confront hard situations honestly, which is unfortunate, but I don’t feel necessarily that I would have that support.

The other principal responded more positively to her increased awareness:

I do know that I have been more aware of how I communicate decisions . . . the timelines that I give, and the way I respond to situations based on the feedback from my superintendent last year. . . . And so I do think that it did help me make some changes in who I am and how I lead . . . and mostly how I communicate, because I can be kind of a straightforward answer, and sometimes I guess some people need you to kind of flower it up a little bit more than I was doing.

The researcher inquired if professional development was tied to the outcomes of any of the principals’ evaluations. All 10 said no. Four principals mentioned that
professional development could be available to them; however, two felt they needed to ask for it and two offered that all professional development for principals was the same. “PD for principals is tied to the expectations that they have for us,” stated one of the participants. Another explained, “We’re getting some professional development in how to collaborate with teachers and things that some of us have been doing for years and others haven’t.”

**Success.** The researcher probed further and asked participants how successful they thought their district’s principal evaluation process was. Three felt strongly that their district’s process was not at all successful. “I don’t know that it’s the piece that makes or breaks a principal. So, in that regard, I don’t think it’s that successful,” conveyed one of the principals. Four other principals felt their processes had some successful components and might even be successful for some principals. One principal contemplated:

I think the process is good for giving feedback and for opening dialogue. I don’t think the process is good for what, if there’s belief that the purpose of the process is to determine if a principal should remain in the profession, which is what the teacher evaluation process is supposed to be, then I’m not sure it’s successful in that sense.

A comment from another principal offered specific feedback of the process:

I would say if you were to give it a 1 to 5, I’d put it probably at about somewhere between a 2 and a 3. I’d say maybe a 2 plus or a 3 minus. I would say it has some good parts—you know the goal writing—and I think that data piece was good this
year going in and discussing that. I just think that, I don’t think that the focus and the goals are clear to us necessarily.

Another principal felt that the new process adopted by the district had the potential to be successful, but being the first year couldn’t really comment. The final two principals stated that they just didn’t know if the process was successful or not.

Principals were asked how their district determined if the principal evaluation system was successful. Some believed that it was probably something discussed in the superintendent’s cabinet meetings or between the superintendent and the school board. One mentioned that they may assume success if the test data are successful. Nonetheless, the ultimate answer given by all of the participants interviewed was that they really had no idea.

After each interview, the researcher recorded an overall vibe of the participant. Participant vibe, as recorded in researcher notes, revealed that principals fell into roughly three categories: 3 out of 10 motivated no matter what, 4 out of 10 had a lack of confidence in the process or evaluator, and 3 out of 10 saw their evaluation as just a hoop to jump through (Figure 2). Yet, despite an overall feeling that current evaluation practices fail to influence leadership practices, principals enthusiastically offered suggestions for improvement as described in the next section of this chapter.

**Suggestions for Improvement**

To improve principal evaluation systems and better measure leadership effectiveness, participating principals offered approximately 16 different suggestions for evaluators, such as surveying stakeholders; reviewing expectations and standards with principals; and providing professional development aligned to principal responsibilities
Three predominant themes emerging from principal suggestions were trust, visibility, and increased feedback.

**Trust**

Principals discussed trust in various ways. Three participants specifically referred to consistency, being able to count on receiving feedback from their supervisor. One quote illuminated this point:

I think my boss conceivably could get this done for every principal, but my biggest recommendation would be getting it done. And really, honestly, I’ve been thinking about it for a month, and I almost asked at our principal meeting yesterday, like so what’s up with the evaluations last year. I thought you don’t want to be that person to say it in front of everybody else . . . you don’t want to be that person. But I honestly would like that feedback.
Similarly another principal pondered,

Well, I think that there needs to be a little bit more consistency of how . . . are implemented. I’m not saying that I need to do a full binder; put evidence for my goals; and make it into a dissertation at the end of the year, but I do think I should have been expected to have something in writing at the beginning of the year and have some sort of reflection whether it was verbally articulated or a summary in writing or a combination of the two because I do see that, that year was more beneficial in many ways.

Principals believed that evaluations could be improved if they had trust in the process and faith that their evaluators would actually follow through.

Leading by example was another way that principals referenced trust. Many participants concurred that evaluators should be modeling the things they are asking principals to do. Illustrating this view, a principal made the following comment:

I think it makes sense that we all want our students to achieve, but I think that it might be helpful if you want, for example if you want PLCs in place, then we should be actually doing PLCs. If we’re going to be held accountable for having instructional conversations with our teachers, and we should be looking for this because this is our focus, well then there should be some real specific PD to me to ensure that everybody knows what it’s supposed to look like.

Another principal made a comparable observation:

Lead by example. . . . And I think when I say that like I strongly believe there are many, many times that, as a principal, I should be the one who brings a directive or an initiative to my staff, and it should come from me, and it should be out of
my mouth to them and so forth. But there are also times where when it’s overarching, if it’s coming from above, and I felt this way in my last district, and I feel this way here. There are times when a superintendent or an assistant superintendent should be at a staff meeting and it should be their directive.

Furthermore, principals suggested that, to lead by example, it would be helpful if evaluators had experience as principals. “The biggest problem I have also is that the person that evaluates me doesn’t really know how to give me that help that I would need anyway because of the experience or lack thereof,” offered one participant. Another expressed, “Be principals themselves. . . . Our assistant superintendent was never a principal, so some of the expectations are unrealistic so motivation gets affected.” If possible, principals suggested considering evaluators’ experience would help strengthen evaluation processes.

Participants also suggested building relationships between supervisors and evaluatees to increase trust. By building relationships, principals noted that there would be an increased comfort with the evaluator and enable principals to take greater risks. A principal described the positive rapport she has with her evaluator:

Being an elementary principal, it’s like the loneliest job in the world. You don’t talk to anybody. Nobody’s even there with you. . . . You make all your decisions by yourself, so you think that your boss and the person that evaluates you would be the best person to know when you have a good rapport with them, which do I do; I would feel comfortable saying I’m failing in this area . . . I need help.
Conversely, another principal had this to say about her relationship with her evaluator:

I think in our situation, I think that there is not a trust. There isn’t a true relationship where I could raise my hand and say, hey, why are we doing that, because I think that, you know because there’s a fear there of well you know what might happen. I don’t want to get on somebody’s bad side.

Participants expressed a desire for evaluators to be open, supportive, and get to know principals better to more effectively evaluate their leadership effectiveness.

Visibility

Being more visible surfaced as another way that principals believed supervisors could improve evaluation practices. “I think just making school visits for one and seeing what’s going on,” said one principal. Participants suggested that supervisors visit schools more often, attend more functions, and meet more often with principals. One principal proposed that doing these things might provide better insight into her leadership effectiveness:

I think in talking to me and talking to my staff and engaging with me in conversations about instruction and being with me at my site more. Kind of seeing me in that role. I think that that would probably be a better measure because I think right now . . . it’s a bit removed.

Observing principals on the job was a popular suggestion made by at least half of the principals interviewed. One principal remarked, “The more aware that the supervisors are of what the principals are actually doing . . . The same thing like a principal evaluates a teacher more effectively if they know what it is that they’re doing in the classroom every day.” Another principal elaborated on this point,
I think that the person evaluating you should be around you more than once a year. And I mean around you meaning at the school site, walking through the classrooms with you, talking through what’s going on at your school, and be aware of the good things happening without you having to tell them.

Several principals also added that they would like their evaluators to solicit input from stakeholders. “I think getting that feedback is more effective when you pull a teacher aside in the staff lounge and say, ‘Hey, do you have five minutes, can I talk to you?’” suggested a participant. Other examples of this included the following statements:

- Talk to my staff. Things like how many people transfer? How many people ask to transfer? How many people are willing to participate in a leadership role?
- Things like that. . . . What’s going on with your staff? What’s morale like?
- What’s the school culture like and do kids come to school every day?
- I think talking directly to staff members and I know, I know my boss is good at that . . . I know that one of the reasons that my boss has a very strong opinion of me is because my staff has such a strong opinion of me and she heard that regularly from them and so she took a lot of stock in that for good or for bad.

Two principals also discussed visiting each other’s sites with their evaluator as a way of increasing visibility. “I would actually like to see more opportunities for some coaching, some peer coaching or more opportunities for us to be able to visit each others’ sites and really have those conversations around what we’re doing as colleagues,” offered one of the principals. These visits could also spark conversations about common goals.
Feedback

Principals’ desire for feedback also arose as a theme during data analysis. As stated by one principal:

It would be nice to come and observe a day and see how I handle things and then give me feedback and ideas of how maybe I could have handled some things differently and would have had a different outcome. Like the way you’re coaching or the way I evaluate even teachers, I go and observe them and then I give them feedback and ideas. I don’t just say to them work on this. I tell them how.

Principals commented on a need for more formative feedback as well as more one on one conversations with their evaluators. Over half of the participants interviewed suggested that in addition to praising positive behaviors, it would be helpful for supervisors to also give suggestions on how to improve. One principal highlighted:

All of the good job feels good, but good job only gets you so far, and so I think there needs to be that positively critical feedback as well to really say here’s the next step. I think there’s a fine balance between, yes, you achieved your goal and you’re doing a good job and you are failing at this job.

Likewise a second participant put forward:

I don’t mind critical feedback. I mean I think like any other human being, yeah, sure feelings get hurt for a couple of minutes, but overall I get over it and so I can, I totally, I think it’s important to get critical feedback, but along with that, I want some support on how to improve. I want some possible solutions.
Still a third principal suggested having conversations to help determine next steps:

I think specific feedback that would help me define next steps. If there’s something that they’re seeing that I’m not addressing, tell me. I guess be really forward about it. I would want to know if I’m doing too much of one thing, not focused, if I don’t have my eye on the right ball, that kind of thing.

What is more, several participants noted that it would be helpful for their supervisors to help refine and regularly review their goals as part of the feedback process. In some cases the principals suggested a need for more autonomy with goals, while others suggested having goals determined by their supervisor. One principal explained the need for personalization:

I think the goals are too broad. I think the district goal is so broad that anything could fall into it, and so that dilutes it so much that with such a big school, you have to refine it. You have to pick a group to really put under the microscope. So I think I would give more autonomy to us to really develop our goals.

A different principal proposed the opposite view:

I think it would be nice if at least one of the goals was actually written and directed by that superintendent instead of it being completely self-generated. And then there’s guidance within there, but sometimes I feel like I strive to get better, but it would be nice to have the input from somebody else to say this is what I see.

In either case, the principals suggested more of a collaborative approach to goal setting.

**Summary**

The results presented in this chapter reported the status of current elementary school principal evaluation procedures. In addition, it described the processes and tools
used in principal evaluation systems; reported the degree to which current principal
evaluation practices align across districts; informed how practices align to research on
effective leadership practices; and highlighted the degree to which elementary school
principals perceived evaluation procedures influenced or supported improvement of their
leadership effectiveness. A third section presented suggestions for improvement made by
the principals interviewed.

Data from interviews, evaluation instruments, documents, websites, demographic
data, and field notes comprised the data corpus. Findings in each section were organized
under distinct themes including processes, tools, alignment, influence, trust, visibility,
and feedback. The next chapter will present a more detailed discussion of the results,
discuss study limitations, and offer implications for future research.
CHAPTER 5—SUMMARY AND DISCUSSION

This final chapter summarizes and reports the general findings of the study. It offers a concise overview of previous chapters by restating the research problem and methodology. The major emphasis of this chapter provides a summary and interpretation of key findings; recommendations and implications for educators; study limitations; and suggestions for additional research.

Overview of the Problem

Studies have proven that strong principal leadership is a critical factor in the achievement of schools (Edmonds, 1980; Elmore, 2000; Ginsberg & Thompson, 1992; Lezotte, 1992; Mace-Matluck, 1987). In addition, the literature review established that there is a vast amount of research identifying leadership behaviors and practices of successful principals (Cotton, 2003; Marzano et al., 2005; Murphy, Goldring, Cravens, Elliott, & Porter, 2010; Stiggins & Duke, 2008). Therefore, thorough evaluation of principals is necessary to ensure outstanding leadership and effective schools (Thomas et al., 2000).

Regrettably, research is limited regarding principal evaluation systems and processes (Albanese, 2003; Davis et al., 2011; Ginsberg & Berry, 1990; Murphy et al., 1985; Studebaker, 2000). The literature review for this study indicated that investigating various evaluation methods and tools might unveil areas of strength while uncovering potential ways to maximize the effects of principal evaluation. This study attempted to add to the paucity of theoretical and empirical research to better inform administration responsible for supervising and evaluating principals of practices most supportive of leadership effectiveness. More specifically, the study attempted to better describe the
status of current elementary school principal evaluation procedures and identify how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness.

Two main research questions emphasized this focus:

1. What is the status of current elementary school principal evaluation procedures?
2. How do elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness?

The following inquiries helped to further deconstruct the research questions:

1. What do processes and/or practices used in principal evaluation systems entail?
2. How consistent are principal evaluation practices across districts?
3. Do current principal evaluation practices align to research on effective leadership practices?
4. How do principals perceive their evaluation systems influence their leadership behaviors?

Methodology

The study encompassed a comprehensive literature review including an historical look at the principal role, key leadership behaviors of today’s principals, an overview of evaluation, current principal evaluation practices, and factors influencing principal evaluation. Results indicated a significant lack of empirical research describing effective principal evaluation practices and their impact on leadership behaviors.

A general study was designed to collect data missing in the literature. Merriam (1998) described a basic or generic qualitative study as encompassing three main
attributes: includes description, interpretation, and understanding; identifies recurrent patterns in the form of themes or categories; and delineates a process. Using this basic qualitative approach, the study produced descriptive findings potentially useful to principal evaluation and leadership.

A purposeful and convenience sample of 10 elementary school principals representing different school districts within Southern California participated in this study. Interviews conducted between June and August 2011 served as the primary source of data collection. The researcher not only asked principals to describe current processes and tools, but also elicited their perceptions of the influence their evaluation systems have on their leadership. Describing and understanding principal evaluation in terms of the meaning people bring to them exemplified this research (Boeije, 2010).

Principal evaluation instruments, documents, websites, demographic data, and field notes added additional components to the data analysis and served as a means of triangulation of the data. A constant comparative method helped to organize and analyze data collected. The researcher categorized the data into the themes of processes, tools, alignment, influence, trust, visibility, and feedback. NVivo, a qualitative data analysis software, also aided in the analysis process.

Data analysis was used to compare and draw conclusions about principal evaluation practices compared across districts. The next section summarizes and discusses these findings in greater detail.

**Summary and Discussion of Findings**

The summary and discussion of findings are reported in three parts. Sections one and two align with the research questions of the study. First, what is the status of current
elementary school principal evaluation procedures? And second, how do elementary
school principals perceive evaluation procedures support improvement of their leadership
effectiveness? A third section details principal suggestions for improvements of current
practices.

**Status of Principal Evaluation Practices**

The researcher began by asking what processes and/or practices used in principal
evaluation systems entail and examined the consistency of these practices across districts.
The data corpus illuminated a wide range of both processes and tools. Overall, processes
were found to be inconsistent both within and across districts. Principal evaluation tools
proved incongruent to evaluation practices and principal leadership practices.

Approximately a decade ago, scholars concluded that principal evaluation
practices were varied in approach and often not consistent between districts or states
(Albanese, 2003; Studebaker, 2000). Results of this study confirmed that this statement
is still true in Southern California in 2011. Though most participants interviewed
described a principal evaluation process that occurs annually or biannually and consisted
of written and face-to-face components, these are essentially where the similarities end.
Interestingly, out of the 10 principals interviewed, no two processes were the same.

As outlined in the results section, variations in principals’ descriptions included
processes that were formal to informal, vague to clear, formative to summative, and
consistent to nonexistent. Among the differences explained by study participants were
who evaluates principals, how principals were monitored, the type of feedback principals
received, and discreet components of the principal evaluation process.
Yet, variances existed even within common practices. For instance, although multiple principals brought up goal setting, there were variations on whether goals were prescribed for them or whether they could be self-selected by the principals. Likewise, meetings between the principals and supervisors were common to most evaluation processes, however meeting frequency varied greatly. Not only did meeting frequency vary between districts, it also varied between many intended and actual processes. These variances underscore the inconsistency of approaches.

Guerra-Lopez and Leigh (2009) advocated that evaluation processes were designed to measure results, identify strengths and barriers, and make recommendations for improvement. Of the principals interviewed, most seemed uncertain about the purpose of their evaluation. Participants offered assumptions split between accountability, fostering growth, and determining support needed. None of the principals supported the idea that their evaluations were intended to measure leadership. One principal felt that her district’s instrument had potential; however, the district did not implement the evaluation as designed. Other principals felt there was some potential for measurement of goals or to receive feedback on leadership, but not specifically to measure leadership itself.

Many participants reported receiving some sort of formative feedback either within or outside of the evaluation process; however, the general consensus was that evaluation practices tended to be predominantly summative in nature. This aligns directly with Albanese’s (2003) conclusion that even when practices for principal evaluation exist, methods are often summative or punitive. During interviews, principals stressed a
need for more formative feedback from their evaluators. This will be discussed in more detail during the discussion of principal suggestions for improvement of practices.

In addition, the majority of principals described their evaluation as changing from year to year, or changing depending on the evaluator completing the tool. Many explained that principal evaluations practices were not implemented with fidelity, or as fully as intended. Principals were unaware of district board policies on principal evaluation, or even how their evaluators had been trained to perform these evaluations. Concomitantly, few principals felt they knew their role in the process and did not feel they had much flexibility in the evaluation process. Most shockingly, most principals felt little or no accountability to their evaluations. These factors demonstrate a lack of communication, understanding, and “buy-in” of the overall process.

As referenced in the literature review, many studies have confirmed that principal evaluation tools tend to be locally developed, not well aligned with professional standards or the literature on leadership effectiveness, and offer little feedback to principals (Davis et al., 2011). Findings of this study support this assertion. Not only were a preponderance of the tools self developed, or perceived by principals to be, none of the principals were involved in the development of the tool. Only one principal reported having any input into the process. Alignment of tools to standards or the literature on leadership effectiveness will be discussed further in the following section.

Overall, tools varied in design. Some were more narrative and focused on goal setting, others were rating scales (most of which were aligned to standards or indicators), and some were a combination of both. Though principals described the principal evaluation processes in their districts very differently, several tools were actually similar
to one another. This indicated to the researcher a discrepancy in implementation between processes and tools.

Of the tools analyzed, four were very tightly aligned to the ISLLCS/CPSEL standards, one aligned to the Six Dimensions of Leadership, one promoted behaviors aligned to the theory of instructional leadership, and four did not seem tied to any identified standards or research in educational leadership. Surprisingly, only one principal was aware of any standards attached to their evaluations, and specifically referenced the California Professional Standards for Educational Leaders (WestEd, 2004). Unfortunately, as pointed out in the literature review, instruments tied to ISSLC and CPSEL standards remain weak, and more empirical data are needed to assure instrument validity and generalizability for practical application (Babo, 2009; Brooks & Voss, 2008).

Bossi (2008), on the other hand, advocated for the Six Dimensions of Leadership. He referred to the ISSLC and CPSEL standards as incomplete and unchanging, and Marzano et al. (2005) and Cotton’s (2003) as too cumbersome. Remarkably, the one instrument based on Bossi’s (2008) Six Dimensions of Leadership added an additional 65 indicators of success to the six proficiencies for instructional leaders, more than three times Marzano et al.’s leadership responsibilities. The final tool that was somewhat aligned to the theory of instructional leadership provided only a loose correlation, not appearing to be tied to any specific research or standards.

Research from the literature review also indicated that few evaluation instruments are published, most lack rigor, have low reliability, are outdated, and have not been field tested or vetted (Condon & Clifford, 2009). Only one of the tools put forth in this study appeared to be published. That particular tool, however, seemed to be outdated and not
aligned to current research on principal standards or behaviors. There was no evidence that any of the remaining instruments had been field tested or vetted.

Principals described many of the tools as “bits and pieces” of others, almost as if thrown together without much thought. To take this one step further, study participants were unaware of any research tied to their respective evaluation instruments. These findings support those of Ginsberg and Berry (1990) who discovered that, of the wide array of principal assessment practices being used, little systematic research exists in support of one over another.

**Perceived Effects on Leadership**

This study also sought to understand if current principal evaluation practices aligned to research on effective leadership practices and whether principals perceived their evaluation systems influenced their leadership behaviors. Researchers have conducted many studies linking effective leadership with student achievement, and scholars agree that the duties of the elementary school principal are quite complex (Catano & Stronge, 2006; Marzano et al., 2005). Data collected in this study centered on activities principals engaged in on a regular basis that they perceived had the greatest impact on student achievement, how principals spend the majority of their time on a typical school day, and other tasks they are required to do as principals.

Because the quality of principal leadership constitutes a primary indicator of school effectiveness, many scholars have attempted to better understand the complexities and needs of the role of the school principal (Elmore, 2000; Ginsberg & Thompson, 1992; Lezotte, 1992). Therefore, the researcher attempted to isolate the key behaviors that principals felt they engaged in on a regular basis. When study participants described
the myriad of duties they perform, an incongruent alignment emerged between what principals reported and the evaluation systems used to evaluate their work. The researcher also compared principal duties described to Marzano et al.’s (2005) 21 leadership responsibilities.

Most principals concurred that they did not believe the principal evaluation system in their district aligned to the leadership practices they deemed necessary and important. Comparing principal duties to evaluation systems, there was no overlap between what the principals identified as key behaviors and the emphasis placed on district tools. This finding alone suggests one reason why principals do not feel that their evaluations are meaningful or useful. Likewise, there were no commonalities between the emphasis of district tools and the top 21 leadership behaviors with the largest correlations to student achievement.

There was overlap, however, in two areas between the most common responsibilities reported by principals and the top leadership behaviors identified in the research. These two areas were monitoring and evaluating, and outreach. It cannot be determined if this overlap is due to principals choosing to engage in these responsibilities or whether they are overall necessary functions of the job. Neither distinction is as important as why they are also not represented on district evaluation tools. Also, why are the other most highly correlated behaviors that impact student achievement not represented or measured as indicators of principal success?

Principal evaluation tools should measure the performance of an administrator by aligning principal behaviors to indicators that demonstrate an effective principal, yet many are not (Brooks & Voss, 2008). Researchers once thought that the complexity of
the principal role and the context in which they serve makes it challenging to standardize assessment practices (Glasman, 1992; Heck & Glasman, 1993; Heck & Marcoulides, 1992). With the meta-analysis by Marzano et al. (2005), district supervisors now know which principal behaviors can lead to the greatest impact on student achievement, yet only one of the evaluation tools in this study is closely aligned to this research. Unfortunately, the instrument that was aligned was not implemented in the evaluation of the participating principal in this study and, because it was not used, it failed to measure anything.

Evaluation systems that do not consider the day-to-day reality of a principal role, even ignoring certain aspects of the job, are not a new phenomenon (Albanese, 2003). Therefore, principal evaluation systems continue to have limited influence on principal behavior or performance. In this study, participating principals also did not find their district’s evaluation systems to be important to their leadership. In addition, no upward academic achievement trends could be associated with effective evaluation practices.

Marzano et al. (2005) identified 21 specific leadership responsibilities that have the greatest impact on student achievement. Therefore, districts should consider revising principal evaluation practices and tools to more closely align with that research. With a more effective tool, districts could more accurately identify areas in need of development for principal leaders. Knowing areas of need would help enable districts to design appropriate professional development and offer specific support to principals. Evidence in this study demonstrated that, in the districts studied, none of the principals are receiving professional development based on the outcomes of their evaluation.
Districts should also consider investigating ways to measure the success of principal evaluation systems. None of the participants in this study could definitively state that they felt their district had a successful system. In addition, most agreed that they had no idea how the district determined the success of its system. Without a successful system, evaluation practices will continue to have little effect on principal behaviors, just as none of the participating principals felt that current evaluation practices influence their leadership practices.

**Suggestions for Improvement**

Seeing their evaluation process as basically a broken system, principals offered many suggestions for improvement. Three themes of trust, visibility, and feedback arose as most universal. Principals described trust as supervisors needing to be consistent, lead by example, and build relationships. Participants also felt that their supervisors should be more visible. Visiting schools, observing principals, soliciting input from stakeholders, and providing opportunities for coaching were ways that participants felt their supervisors could increase visibility. Finally, principals expressed a need for increased feedback from supervisors through formative feedback opportunities, individual conversations, and a more collaborative approach to goal setting.

Participants described the need to trust in their evaluator and to feel supported. They also needed to believe that the process would be implemented and that they could count on feedback. For many principals in this study, this did not happen. Thus, the principals lost faith in the process and the process lost the ability to influence the principals.
Overwhelmingly, principals felt that their supervisors were not around them enough to really understand their performance. Most agreed that supervisors could gain a deeper knowledge of how they work by observing them on the job and attending school functions more than just once or twice a year. Principals also believed that talking to stakeholders such as teachers and parents could provide their supervisor a more objective view of their abilities.

Most of all, principals seemed to crave feedback. Many mentioned the desire for positive affirmation; however, over half also expressed the importance of receiving critical feedback. Based on the comments of the participants in this study, principals want their supervisors to give them specific suggestions on how to improve and support them in doing so.

**Limitations**

To determine applicability of these findings, it is important to discuss the study limitations (Marshall & Rossman, 2006). A small sample size of only 10 principals and the localized geographic region of Southern California could both be considered limitations to the generalizability of the research. Also, using a convenience sample may have limited the scope of potential participants. The study sample consisted of all women. However, since many of the findings echo what has been previously reported in the literature review, it is unlikely that the gender of the participant affected study results.

Another consideration is the perspective of the participant. Evaluation is a sensitive topic, and some principals may not have divulged true feelings. Evaluation is also a private process. Principals may not have information about others and therefore can only speak for themselves and their own perspectives. Perspectives may have been
exaggerated or downplayed and could potentially have influenced study findings. Triangulation between other data sources such as evaluation instruments, documents, websites, demographic data, and field notes attempted to mitigate this factor and provide true insight into practices.

Finally, researcher bias should always be considered as a potential study limitation. To limit internal bias, the literature review provided a framework in the development of the interview questions. Principals outside of the study were also asked to provide input on interview question clarity and objectivity. This feedback was not included in the findings of the study.

**Implications**

There is sufficient research proving effective principal leadership is a significant aspect in the achievement of schools (Edmonds, 1980; Elmore, 2000; Ginsberg & Thompson, 1992; Lezotte, 1992; Mace-Matluck, 1987). Therefore, districts should place more emphasis on how principal leadership is measured and evaluated. As highlighted in this study, and confirmed by the literature, principal evaluation systems are plagued by inconsistent processes and ineffective instruments. Principals have limited understanding of the purpose of their evaluations, and in result the evaluations lose effect.

To remedy this, superintendents and other supervisors of principals should consider revisiting their district’s principal evaluation processes. Principals should be asked to offer input and perceptions of the strengths and weaknesses of the system. A comparison between research on effective leadership identified by Marzano et al. (2005) and district practices would also identify how systems could be modified for maximum results.
As emphasized throughout this study, there is a considerable amount of research identifying leadership behaviors and practices of successful principals (Cotton, 2003; Marzano et al., 2005; Murphy, Elliott, et al., 2010; Stiggins & Duke, 2008). To increase alignment of principal evaluation processes, superintendents should start with this research. Principals are desperate for feedback, but if evaluations are not focused on the correct leadership responsibilities, supervisor suggestions could lead principals astray.

Developing a list of expected behaviors and standards for principals would be time well spent for districts. Although some of this work has been done through the development of standards for leaders, such as the ISSLC and CPSEL standards, some scholars would argue that the these standards leave out or underemphasize important components of effective leadership (Bossi, 2008; Davis et al., 2005). Predetermined standards may be used as a starting point; however, it would behoove districts working to align their practices to also consider the research on effective leadership when developing expectations.

As found in this study, most principals were unaware of the standards or research behind their evaluations. In addition, none of the participating principals were involved in the development of evaluation instruments used by their districts. However, Roberts (2003) found that involving stakeholders in the process of creating evaluation tools could provide increased benefits such as better employee attitude, increased understanding of process, and perceived reduction of bias/favoritism. Thus, involving principals in the process would increase principals’ understanding of the expectations of them, create more buy in, and perhaps result in more alignment to their leadership practices.
Still, it is not only important to look at the quality of processes and instruments, but also at how consistently practices are implemented. Longnecker and Nykodym (1996) recommended that since performance appraisals are an important function for employee growth and development, organizations should continually look for ways to keep performance appraisals at the forefront of management responsibilities. Principals need to feel valued and be able to count on receiving beneficial feedback about their performance. If not, principals, like those in this study, will lose trust in the system and evaluations will not be effective. Moreover, districts need to identify barriers to consistency and take steps to remedy them.

Lastly, and probably most importantly, processes not only need to be consistent, they also need to influence principal behaviors. Principals should have confidence in the training and ability of their supervisor to accurately and fairly evaluate them. Next, steps should be identified and principals should be held accountable to work on evaluator suggestions. Principals should also see that professional development and support is tied to the outcomes of their evaluation.

To best understand the effect of individual evaluation systems, superintendents should listen to their principals. In this study, participants interviewed offered the following suggestions to improve evaluation practices of supervisors:

1. Build relationships and trust.
2. Provide more formative feedback.
3. Engage in regular conversations about leadership.
4. Visit schools more often.
5. Be consistent in completion of evaluations.
6. Align professional development to expectations.

7. Include suggestions for principal improvement.

By understanding the needs and motivations of principals, districts can discover how to best leverage principal evaluations.

**Recommendations for Future Research**

The literature review in this study indicated a deficit in research of principal evaluation systems and processes (Albanese, 2003; Davis et al., 2011; Ginsberg & Berry, 1990; Murphy et al., 1985; Studebaker, 2000). This study does not come close to filling the gap in the research and is only one of many studies that still need to occur on the topic of principal evaluation. Empirical studies such as this one could help educators better understand the limitations of current practices and increase transferability of findings. Furthermore, some states have begun to make changes in their evaluation systems to better align to professional standards; however, due to increasing debates over evaluation policies, more studies need to be done, specifically in California (Morse, 2011).

Additional studies would also increase the potential to identify exemplars in the field worthy of deeper study. A few studies have revealed some promising tools and practices such as Vanderbilt’s VAL-ED instrument, portfolio based evaluation systems, and the work of West Ed (Davis et al., 2011; Marcoux et al., 2003; Porter et al., 2010a). There is an acute need to study these and other various methods and tools that might uncover potential ways to maximize the effect of principal evaluation. Unfortunately, finding progressive districts using good measures of principal performance to study remains a challenge.
Results of this study compared principal reported leadership behaviors with those identified in the research and those found in principal evaluations. Many scholars also believe that supplementary valid studies are needed to better understand how principals truly allocate their time across the various leadership practices (Camburn et al., 2010; Hallinger & Heck, 1996). Increasing accountability of principals augments the importance of having accurate, effective, and fair evaluations of principal performance (Cravens et al., n.d.; Davis et al., 2011; Polikoff et al., 2009). Thus, it is imperative to study current principal evaluation practices and consider how to improve practices to better align with current leadership expectations for improving student achievement. How evaluations influence principal behavior should also be studied further.

Of course, barriers to effective principal evaluation warrants further investigation. These can be explored not only through principal perceptions, as in this case, but also the perceptions of supervisors. For instance, how well do superintendents feel evaluations in their district measure principal leadership? Are superintendents aware of how principals feel regarding the effect, or lack thereof, of the evaluation system? Why are evaluations stagnant? Why is not more emphasis placed on principal evaluations? These critical questions only scratch the surface of possibility for additional exploration.

Conclusion

This chapter provided a summary and interpretation of key findings as well as recommendations and implications for educators. Study limitations were acknowledged and suggestions for additional research were suggested. The following are final thoughts on the research overall.
By conducting this study, the researcher gained a better understanding of the status of current elementary school principal evaluation procedures in Southern California, and how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness. Evidence collected from participants clarified that policies and practices across districts are varied, and principals do not feel that current evaluation practices influence their leadership. Predominant themes emerging from data were evaluation processes, evaluation tools, alignment of evaluation processes with leadership research, influence of evaluation practices, trust between principals and supervisors, visibility and feedback of supervisors.

Now more than ever, there is a strong need to revise and align our current practices to the established research on effective leadership practice. Schools are being held accountable for high levels of achievement, and principals are at the center of that accountability. As of October 2011, the Senate Education Committee of the United States approved legislation to rewrite the No Child Left Behind Act (Duncan & Martin, 2010). If this Education and Secondary Education Act (ESEA) cannot be reauthorized by the start of 2012, the President’s administration will implement an NCLB waiver plan for states. One of the key provisions of this comprehensive bill examines how states and districts choose to evaluate both teachers and principals (Association for Supervision and Curriculum Development [ASCD], 2011).

The time is now for states and districts to critically look at principal evaluation practices and create systems that accurately measure principal performance and support principals in their professional growth. Unfortunately, principal evaluation is still a much underdeveloped topic in educational research.
REFERENCES


Bippus, S. (2007). When a performance review is seen as criticism. *School Administrator, 64*(2), 43-44.


record?group=003019&resultSet=044127&startRecord=6

CERTIFICATE OF TRANSCRIPT

I, the undersigned, Laureen Minnich, do hereby certify:

That the audio recordings of interviews for Leighangela Brady's dissertation were transcribed by me to the best of my ability and were a true and faithful transcript of the interviews taken with a digital recorder.

In witness whereof, I have subscribed my name this 13th day of November, 2011.

Laureen Minnich
Owner
Southern California Transcription Services
4354 Avon Drive
La Mesa, CA 91941
Phone: (619) 464-2553
Email: LMIN@COX.NET
Appendix B

Email Script to Potential Participants

Dear Elementary School Principal,

My name is Leighangela Brady and I am a second year doctoral student with San Diego State University and the principal of La Costa Heights Elementary School in the Encinitas Union School District. As part of my degree, I will complete a dissertation. The intent of my study will be to describe the status of current elementary school principal evaluation procedures, and identify how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness.

To do this, I plan to interview 10 different principals from 10 different school districts here in Southern California. I am seeking current elementary school principals with a minimum of 3 years experience to be interviewed for my study.

I know how busy you must be. I too am an elementary school principal. My goal is to add to the field of research in the area of leadership that may benefit all principals in our position. Results of this study may add both theoretical and empirical research to inform superintendents, districts, and other administration responsible for supervising and evaluating principals of practices that principals perceive most support leadership effectiveness.

Please let me know if you would be willing to participate in my study. The interview should take no more than 45-60 minutes of your time at a time and location most beneficial to you. All data collected will be absolutely anonymous and cannot be traced back to individual participants, and any links between codes, names, or districts will be permanently deleted upon completion of the dissertation.

Attached you will find a consent form with more information. Should you choose to participate, you will be asked to sign the consent form at the time of the interview. You will also have an opportunity to ask any questions before giving your consent. Of course, participation is voluntary and you may opt out at any time during the study.

Please email me at leighangel.brady@eusd.net or call me at (760) 420-6010 for more information or to arrange a time for an interview. Thank you in advance for your consideration of participation.

Sincerely,

Leighangela Brady
Principal
La Costa Heights Elementary School
Encinitas Union School District
Doctoral Student, San Diego State University
Appendix C

Consent to Act as a Research Subject for a Study on Principal Evaluation:

A Description of Current Practices

Dear Elementary School Principal,

You are volunteering to participate in a research study exploring the current status of principal evaluation procedures and what ways elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness. It is important that you provide informed consent prior to participating. To ensure this, please carefully read the information below and ask any necessary questions to be sure you understand the information and what you will be asked to do.

The principal researcher in this study is Leighangela Brady, a doctoral candidate in the Educational Leadership program at San Diego State University. Leighangela Brady is also an elementary school principal in Southern California. This research study will be supervised by Dr. Margaret Basom, a professor in the College of Education at San Diego State University.

Purpose of the Study
The intent of this study will be to explore current elementary principal evaluation practices by studying methods of principal evaluation in 10 different school districts. The study seeks to describe the status of current elementary school principal evaluation procedures, and identify how elementary school principals perceive evaluation procedures support improvement of their leadership effectiveness.

Description of the Study
If you decide to participate, you will answer several interview questions developed by the researcher. The questions will attempt to elicit a description of your district’s current evaluation system and provide details regarding how you perceive evaluation procedures support improvement of your leadership effectiveness. Interview questions will concentrate specifically on what salient actions, processes, and themes exist in the principal evaluation process, how evaluation processes measure principal effectiveness, what feedback, if any, evaluation methods provide principals, and how principal evaluation practices improve leadership effectiveness.

The interview will take approximately 45-60 minutes of your time. For your convenience, you can decide both the location and time of the interview. You may also be asked to provide any documents relating to the principal evaluation process of your district. You should only provide documents that you feel comfortable sharing and that you feel would benefit the study. If you are unable to provide copies of documents at the time of the interview, but wish to send to the researcher at a later date, a self-addressed envelope will be provided to you.
Confidentiality:
All data collected will remain private and confidential. No names, school, or district will be specifically cited in the findings. Any interview data will be stored on the researcher’s password-protected computer, accessible only to the researcher and an SDSU professor overseeing the study. Any documents provided, as well as this consent form, will be kept in a locked filing cabinet in the researcher’s home. Upon completion of the dissertation, any possible link to your identity will be destroyed.

Benefits of the Study:
Data collected from this study may add both theoretical and empirical research to inform superintendents, districts, and other administration responsible for supervising and evaluating principals of practices that principals perceive most support leadership effectiveness.

Costs and/or Compensation:
There are no costs to you for participation in this study. You will not be paid to participate in this study.

Voluntary Nature of Participation:
Participation in this study is voluntary. If you decide to participate, you are free to withdraw your consent and stop your participation at any time without penalty.

Questions About the Study:
If you would like any additional information before participating in the interview, please contact me, Leighangela Brady, at (760) 420-6010 or email me at Leihangela.brady@eusd.net. You may also contact my faculty advisor, Dr. Margaret Basom at mrb@mail.sdsu.edu.

Consent to Participate:
The San Diego State University Institutional Review Board has approved this consent form. The consent form must be reviewed annually and expires 1 year after it has been approved. Your signature below indicates that you have read the information in this document and have had a chance to ask any questions you have about the study. Your signature also indicates that you agree to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time. You have been told that by signing this consent form you are not giving up any of your legal rights. Please return a copy of this consent form with your completed survey and print a copy of this consent form for your records.

Name of Participant (please print)

___________________________________ __________________
Signature of Participant Date

Signature of Investigator Date
Appendix D

Interview Questions

Principal Evaluation: A Description of Current Practices

Leighangela Brady SDSU 2011

Interview Questions:

1. There is a large variance of principal evaluation practices in the literature. This study aims to describe current principal evaluation procedures and practices and identify how elementary school principals perceive evaluation practices influence their leadership effectiveness. Please describe your district’s overall principal evaluation process.

   Probing questions:
   1a. How often are you evaluated?
   1b. When are you evaluated?
   1c. Who evaluates you?
   1d. Where does the evaluation take place?
   1e. Are you evaluated in writing, in person, or both?
   1f. Would you describe the process as formative or summative? Please explain.
   1g. Is the process the same for both new and veteran principals?

2. Please describe in detail the tool used for your evaluation.

   Probing questions:
   2a. Is it self-developed or purchased from outside?
   2b. If self-developed, how was it developed? Is it tied to specific research?
   2c. Is data a component of the tool? If so, what type or types?
   2d. Is the tool aligned to any state or national standards? If so, which ones?
   2e. Is there a scale or a rubric used in conjunction with the tool?
   2f. Do you have any forms or artifacts that you would be willing to share?
3. In the historical context of evaluation, evaluation was designed to measure results, identify strengths and barriers, and make recommendations for improvement. What do you believe is the **purpose** of your district’s principal evaluation system?

**Probing questions:**
3a. Provide feedback? Measure growth? Determine support needed? Improve teaching?
3b. Does your district have a board policy or regulation on principal evaluation?
3c. How well do you feel your district’s principal evaluation process measures leadership?
3d. How beneficial is your evaluation in helping you improve your practice?
3e. Please describe any of the ways you feel the process supports or helps you.
3f. How successful do you feel your district’s principal evaluation process is?
3g. How does the district determine if the principal evaluation system is successful?

4. What do you perceive your **role** is in the principal evaluation process?

**Probing questions:**
4a. What flexibility do you have in goal setting?
4b. Do you have any input into the process? What is your level of participation?
4c. Are expectations of you clearly communicated?
4d. How adaptable is the evaluation system to your own growth needs?

5. The literature review in this study revealed that in addition to a variance in methods of evaluation, there is also a wide-range in **implementation** of practices. To what degree is what you describe to be the process implemented in your district?
Probing questions:
5a. How well do you feel principal evaluation practices are implemented?
5b. What factors do you believe influence this process?
5c. How would you describe the level of accountability for reaching your goals?
5d. To your knowledge, how is your evaluator trained to perform principal evaluations?
5e. Do you receive feedback from your supervisor about your performance in other ways?

6. The duties of the elementary school principal are far reaching. Scholars have conducted many studies linking effective leadership with student achievement. What activities, practices, planning, or decisions do you feel you regularly engage in that have the greatest effect on students?

   Probing questions:
6a. How do you feel you spend the majority of your time during a typical school day?
6b. What else are you required to do as a principal?
6c. Are your supervisor’s expectations of you clearly articulated?

7. In what ways do you feel your district’s evaluation system aligns to the leadership practices you just described?

8. In what ways do you feel your district’s evaluation system impacts your leadership practices?

   Probing questions:
8a. What kinds of decisions do you make based on outcomes of your evaluation?
8b. Is your professional development tied to outcomes of your evaluation?
9. What suggestions would you make for improvement of the current practices?

**Probing questions:**

9a. What type of feedback would be most important to you and why?

9b. Is there any opportunity for you to seek feedback as needed? Please describe.

9c. How could supervisors best support principals to continually improve effectiveness?

9d. How do you feel your supervisor could best measure your effectiveness as a leader?

10. Is there anything else you would like to add that you believe would benefit this study?

- Researcher will thank the participant for their time and remind them that they can request a copy of the findings at any time.
- Researcher will remind participants that a follow-up interview may be necessary if the data collected warrants it.