

SUPERVISORY PERCEPTIONS OF TEACHER SUPERVISION AND EFFECTS ON
STUDENT ACHIEVEMENT IN MISSOURI

SHANNON SNOW
2014

I, Shannon Snow, declare the paper presented to the Southwest Baptist University faculty, as a requirement for my Doctor of Educational Leadership Degree, is my original work. Others have been participants during the process but the end product is a result of my research and my efforts.

Signature of Researcher

Date

The undersigned, approved by the Department Chair of Graduate Studies in Education, have examined a dissertation entitled:

SUPERVISORY PERCEPTIONS OF TEACHER SUPERVISION AND EFFECTS ON
STUDENT ACHIEVEMENT IN MISSOURI

Presented by Shannon Snow a candidate for the degree of Doctor of Education and hereby certify that in their opinion it is worthy of acceptance.

Dr. Pamela Hedgpeth, Advisor/Chair
Graduate Education

Dr. Robert Perry, Committee Member
Professor of Graduate Studies in Educational Administration

Dr. Vicki Raney Sandberg, Committee Member

SUPERVISORY PERCEPTIONS OF TEACHER SUPERVISION AND EFFECTS ON
STUDENT ACHIEVEMENT IN MISSOURI

A Dissertation
Presented to
The Faculty of the Graduate Education Department
Southwest Baptist University

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Education

By

Shannon Snow, (B.S., M.S., Ed.Spec.)

Dr. Pamela Hedgpeth, Dissertation Advisor

December 2014

ACKNOWLEDGMENTS

I would like to thank my family for their support and patience on this journey. You have always believed in me. I love you all so much! To Shelby and Griffin, I am so proud of you for pursuing your own education. You are both going to accomplish great things because you have determination and dreams.

A special thank you goes to my accountability partner, Dr. Chris McClimans. I appreciate the motivation and multiple phone calls and texts telling me to get to work. You kept me going, and I appreciate it.

I would also like to thank my committee members who invested much of their own time in this process for me. I want to thank my advisor, Dr. Pamela Hedgpeth, for all that I have learned from you over the past 3 years. You are an amazing lady! I also thank Dr. Robert Perry, who provided technical support for methodology and statistical work on this project. You have a way of making even complicated research methods seem easy to understand. Dr. Vicki Sandberg has mentored me well beyond the scope of this project. Her support over the last ten years has been invaluable.

SUPERVISORY PERCEPTIONS OF TEACHER SUPERVISION AND EFFECTS ON STUDENT ACHIEVEMENT IN MISSOURI

Shannon Snow

ABSTRACT

The purpose of this study is to examine the practices of supervisors of instruction, specifically in the area of teacher evaluation and supervision, in high performing schools to determine if these perceptions and practices affected student learning. This study described principal perceptions of teacher evaluation practices in high performing and non high performing districts in Missouri. Data was collected via a survey asking the respondent to rate their perception of their current teacher supervision and evaluation process and a teacher supervision and evaluation system they considered ideal. Data was separated into practices, beliefs, effectiveness and professional development practices of the supervisor and analyzed using ANOVA and t-tests. The data from this study does show a significant difference in practices of principals in high performing and non-high performing districts.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	II
ABSTRACT.....	III
TABLE OF CONTENTS.....	IV
CHAPTER ONE.....	1
Problem Statement	3
Rationale for the Study	3
Main Research Question.....	5
Subset Research Questions	5
Theoretical Framework.....	6
Rationale/Purpose of Study.....	7
Limitations of Study	8
Delimitations of Study	8
Definition of Terms.....	8
Evaluation	9
High performing districts	9
Instructional supervision.....	9
Missouri Annual Performance Report	9
Non-high performing districts.....	10
Professional development	10
Supervisor	10
Study Overview	10
CHAPTER TWO.....	12
Introduction.....	12
What is an Effective Teacher?	14
Principles of Teacher Supervision	15

Evolution of Teacher Supervision	21
Clinical supervision	22
Hunter model	24
Developmental/reflective models	26
RAND Study.....	30
Danielson model	31
Missouri evaluation model.....	31
Current Practices for Teacher Supervision	35
Student achievement and growth models	35
Performance continuum models	37
Marzano teacher evaluation system	38
Missouri educator evaluation system.....	39
Network for educator effectiveness	43
Highly Qualified Teachers vs. Highly Effective Teachers	44
Additional Educational Practices that Support Teacher Supervision	46
Professional development	46
Peer coaching	48
Mini-observations/Walk-throughs	48
Teacher portfolios	49
Student surveys	49
Principal Professional Development to Promote Effective Teacher Supervision	50
Conclusion	51
CHAPTER THREE	53
Participants.....	53
Sampling Procedure	55
Research Design.....	58
Instrumentation	61
Reliability of Instrument	63
Guidelines for Administration	64
Data Analysis	65
Rationale for Selected Statistical Treatment.....	66
Likert scaled queries	67
Open-ended queries	67

Interview Questions	68
Summary	69
CHAPTER FOUR.....	70
Introduction.....	70
Analysis of Data.....	70
Descriptive Data.....	71
Inferential Statistics.	85
Qualitative Results	94
Interview Data.....	101
Summary of Findings.....	107
Summary	116
CHAPTER FIVE	117
Introduction.....	117
Summary of Methods.....	118
Summary of Findings.....	119
Limitations of the Study.....	121
Significant Differences of the Study with Previous Research	122
Recommendations for Future Research	123
Implications for Practice	125
REFERENCE LIST	127
APPENDIX A: SURVEY INSTRUMENT	140
APPENDIX B: INTERVIEW QUESTIONS.....	143
APPENDIX C: CONSENT EMAIL.....	144

LIST OF TABLES

Table 1: Demographic Characteristics of Respondents.....	72
Table 2: Practices Subscales-Item Mean & Standard Deviations.....	75
Table 3: Practices Subscales-Subscale Means & Standard Deviations	77
Table 4: Beliefs Subscales-Item Means & Standard Deviations	79
Table 5: Beliefs Subscales-Subscale Means & Standard Deviations	81
Table 6: Effectiveness Subscales-Item Means & Standard Deviations	84
Table 7: Effectiveness Subscales-Subscale Means & Standard Deviations	85
Table 8: <i>t</i> -test results for High Performing Schools Practices Scale and System	86
Table 9: <i>t</i> -test results for Non-high Performing Schools Practices Scale and System	86
Table 10: <i>t</i> -test results for High Performing Schools Beliefs Scale and System.....	87
Table 11: <i>t</i> -test results for Non-high Performing Schools Beliefs Scale and System	87
Table 12: ANOVA for Present Practices Scale and School Level	88
Table 13: ANOVA for Present Beliefs and School Level	90
Table 14: ANOVA for Item 19	90
Table 15: ANOVA for Present Effectiveness and School Level	91
Table 16: Practices Subscales-Item Mean & Standard Deviations	92
Table 17: ANOVA for High Performing Schools Professional Development	93
Table 18: ANOVA for Non-high Performing Schools Professional Development	93
Table 19: ANOVA for Present System Professional Development	94
Table 20: Perceived Strengths of District Supervision-Evaluation System	96
Table 21: Perceived Weaknesses of District Supervision-Evaluation System	98
Table 22: Professional Development	100

LIST OF FIGURES

Figure 1: Practices Subscales-Means of Subscales.....77

Figure 2: Beliefs Subscales-Means of Subscales.....82

CHAPTER ONE

INTRODUCTION

Teacher evaluation and supervision practices have developed over the life of education. However, recent changes in the political climate have caused parents and legislators to demand more accountability of teachers. Evaluation of an individual or group is completed to make sure that those being evaluated are accountable to established standards detailing what the individual ought to be doing. Yet, while the overall purpose of evaluation is to provide a basis for sound decision-making and increased effectiveness, practice does not necessarily match intent (Thomas, Holdaway, & Ward, 2000).

Teacher evaluation models have deemed teachers either satisfactory or unsatisfactory based on models designed to comply with state and local evaluation regulations. Evaluations have been conducted with little feedback to help the teacher improve and few have improved the quality of instruction given by the teachers evaluated (Coulter, 2013). New teacher evaluation models use multi-level tier rating systems instead of merely satisfactory/unsatisfactory ratings. These new systems use indicators that represent effective teaching. In fact, the foundational principle of teacher supervision is that the purpose of supervision should be the enhancement of teachers' pedagogical skills with the ultimate goal of enhancing student achievement (Marzano, Fronter, Livingston, 2011).

With such importance placed on student achievement by the legislation of No Child Left Behind (NCLB), teacher effectiveness is crucial to improving schools. The requirement of NCLB to test all students in grades three through eight in reading and mathematics and track student achievement in those grades has led to the creation of vast longitudinal collections of

student data. Using these data, researchers have created statistical methodologies (i.e., value-added and growth model or progress measures) that use achievement test data to measure the impact of instruction on student learning. Although the methodology continues to be refined, the capability exists within most value-added and growth model systems to identify the impact of instruction provided by individual teachers on the group of students in their classroom (Himmelein).

The Obama administration announced in 2011 it would award waivers under the No Child Left Behind Act to states that adopted certain education ideas, such as teacher evaluations tied to student test scores. In exchange, states would get flexibility from some of the core tenets of the law, such as the requirement of 100 percent of students being proficient in math and reading by 2014. In Missouri, the Department of Secondary and Elementary Education's flexibility request was approved on June 29, 2012. In Missouri's waiver, ambitious performance targets were adopted, a plan was developed to turn around the state's lowest performing schools, and there was increased emphasis on accountability along with support for districts. The Missouri waiver also included a new Educator Evaluation System based on Model Teacher and Leader Standards adopted by the Missouri State Board of Education in 2011. Because of these requirements, teacher evaluation began to receive unprecedented attention both in Missouri and in other states who sought to develop similar evaluation models to meet federal requirements. Because of similar waiver requirements in other states as well, teaching has become a major focus of policy attention, and teacher evaluation has become the primary tool to improve teaching. Much of the research in the area of teacher supervision focuses on teachers' perceptions with little emphasis on supervisors' perceptions (Colby, Bradshaw, & Joyner, 2002).

This study will look at supervision practices and the supervisory perceptions of principals of high achieving and non-high achieving school districts in Missouri. By using both quantitative and qualitative methods, the study will attempt to link principal perceptions about the teacher evaluation process to determine if fidelity of supervision practices results in higher student achievement (Marzano, Frontier, & Livingston, 2011; Marshall, 2009; Clotfelter, Ladd, & Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004).

Problem Statement

Administrator perceptions concerning teacher supervision may affect the fidelity of the evaluation process leading to higher student achievement. Perceptions and beliefs may have an impact on practices of administrators when performing teacher supervision and evaluation. Professional development of the administrator regarding the evaluation process may affect the administrator perceptions.

Rationale for the Study

This study identified the perceptions of principals of teacher supervision and evaluation processes in both high performing and non-high performing public schools in Missouri to determine if differences existed between principal perceptions in high performing school vs. non-high performing schools. The data analysis was expected to identify the common and unique perceptions, which indicated the aspects of these processes that were necessary, to develop an effective teaching staff and become a high performing school. This analysis also highlighted the efforts within the supervision and evaluation process found to be effective or in need of improvement for the educators to meet academic goals. To fulfill these objectives, the research was guided by four main purposes.

First, the study was proposed to identify and describe two groups of supervisor perceptions of the supervision and evaluation processes in high performing and non-high performing school districts. The Likert scale portion of the surveys examined the supervisors' perceptions of the (a) districts' foundational objectives and philosophies evident within the supervision and evaluation process, (b) participants' perceptions of the multiple phases of the supervision and evaluation process, (c) level of professional development/training of the supervisor regarding teacher supervision, and (d) perceptions of collaboration and communication between supervisors and teachers. These themes within the survey instrument helped elicit the supervisors' views about the supervision and evaluation practices in their schools. By describing supervisor perceptions, the study gave a clear view of the perceptions of the necessary characteristics of the supervision and evaluation processes within their schools.

A second purpose of this study involved comparison of the perceptions of both groups of supervisors of their present evaluation system and an evaluation system they would consider ideal. The comparison facilitated drawing conclusions about the type of supervision and evaluation practices the supervisors of both groups of schools currently use and those they believed were effective. This step was relevant for confirmation each group of supervisors had provided an accurate representation of the practices necessary for successful student academic success. This study sought to define to what extent do perceptions of both groups of supervisors of their present system of supervision and evaluation match their perceptions of their ideal system of supervision and evaluation.

A third purpose of this study involved comparisons of the perceptions of the professional development that each supervisor received in the area of teacher evaluation and supervision. The comparisons facilitated drawing conclusions about the type of supervision and evaluation

professional development that all the supervisors of high performing schools versus all the supervisors of non-high performing schools had received to help them be more effective.

Finally, the data was compiled and compared to describe any variation existing between the perceptions of supervisors in high performing schools and perceptions of supervisors in non-high performing schools. Describing the variation was necessary because the variation in these perceptions were anticipated to give the researcher the ability to draw conclusions about the areas of the supervision and evaluation processes that supervisors in high performing schools practice that lead to fidelity of the teacher evaluation process and greater student achievement.

Main Research Question

Are there differences in teacher supervision perceptions regarding the effectiveness of their present teacher supervision system of supervisors of teaching in high performing schools and supervisors of teaching in non-high performing schools?

Subset Research Questions

1. What are the characteristics and perceptions of supervisors of teaching regarding practices and characteristics they would consider ideal for teacher supervision?
2. Is there a difference between the *perceptions* of supervisors in high performing schools and *perceptions* of supervisors in non-high performing schools with respect to teacher supervision?
3. Is there a difference between the *practices* of supervisors in high performing schools and *practices* of supervisors in non-high performing schools with respect to teacher supervision?
4. Is there a difference between the *professional development for supervision practices* of supervisors in high performing schools and *professional development for supervision*

practices of supervisors in non-high performing schools with respect to teacher supervision?

Theoretical Framework

With increased accountability placed on school districts and teachers by No Child Left Behind (NCLB) legislation and, most recently, NCLB waivers granted to states by the Obama administration and Race to the Top, teacher evaluation practices have evolved. To date, President Obama's Race to the Top initiative has dedicated over \$4 billion to 19 states that have created robust plans that address the four key areas of reform:

- Development of rigorous standards and better assessments
- Adoption of better data systems to provide schools, teachers, and parents with information about student progress
- Support for teachers and school leaders to become more effective
- Increased emphasis and resources for the rigorous interventions needed to turn around the lowest-performing schools (United States Department of Education, 2014).

New evaluation models have been implemented recently in several states, including Missouri. The principal's role as an evaluator is more critical than ever as they make judgments during the evaluation process to develop and retain the most effective teachers. Teacher effectiveness is crucial to improving schools and increasing student achievement, thus leading to increased pressure for principals to make good decisions about teacher competency.

Missouri's No Child Left Behind waiver included a new Educator Evaluation System based on Model Teacher and Leader Standards adopted by the Missouri State Board of Education in 2011. Because of these requirements, teacher evaluation began to receive unprecedented attention both in Missouri and in other states who sought to develop similar

evaluation models to meet federal requirements. Because of similar wavier requirements in other states, as well, teaching has become a major focus of policy attention and teacher evaluation has become the primary tool to improve teaching. Much of the research in the area of teacher supervision focuses on teachers' perceptions with little emphasis on supervisors' perceptions (Colby, Branshaw, & Joyner, 2002).

This study seeks insight into the perceptions of principals towards their teacher evaluation process to determine if principal perceptions or attitudes have an impact on student achievement. This study is significant because teacher evaluation practices are clearly changing. What remains to be answered is whether or not current practices influence student achievement. Research shows a link between the effectiveness of a teacher and the achievement of students (Marzano, Frontier, & Livingston, 2011; Marshall, 2009; Clotfelter, Ladd, & Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004).

Rationale/Purpose of Study

High student achievement is crucial in today's educational system. More and more accountability for student achievement is placed on schools by parents, communities, business, higher education and even through legislation such as No Child Left Behind. The purpose of this study is to examine the practices of supervisors of instruction in high performing schools. The point of this study is to describe principal perceptions of teacher evaluation practices in Missouri districts and to link those perceptions to fidelity of the teacher evaluation process leading to higher student achievement. This study expands the body of knowledge on teacher evaluation and the impact it has on student achievement.

Limitations of Study

1. The data used in this study will be obtained through an attitudinal survey. Therefore, the researcher acknowledges the perceptual nature of the data and the honesty of responses by the respondents. The responses will offer a perception on teacher evaluation practices.
2. The accuracy of the data obtained is also dependent on the rate of return of the survey responses.
3. Data will be collected via a survey, but further interviews will be conducted with five school districts for additional information about perceptions of principals concerning teacher evaluation. Interviews will be conducted via phone due to budget and time constraints. The researcher realizes a telephone interview will yield a less dynamic interview than if the interview were face to face.

Delimitations of Study

1. The study will contain a purposive sample with random selection of 50 high performing school districts in Missouri and a random selection of 50 non-high performing school districts in Missouri.
2. Only principals and assistant principals who supervise and evaluate teachers and who have had at least three years in their current position will be considered for this study. Three years of experience in the district will be considered to ensure the principal surveyed has made the impact in the time studied by the researcher.

Definition of Terms

For the purpose of this study, the supervision process is defined as the progression of teaching and learning using various approaches (Harris, 1998). Zepeda (2007) combined

instructional supervision, professional development, and evaluation, stating that, when “woven together in a holistic way, learning opportunities follow their own course while contributing to the overall development of the faculty and the organization” (p. 13). The researcher will define for this research the terms teacher supervision and teacher evaluation synonymously. The terms will be defined as teacher professional growth and the process through which the supervisor makes evaluative judgments of the teachers’ performance (Clark, 1998).

Evaluation. Evaluation is defined as judging the quality of a teacher’s performance (Sergiovanni & Starratt, 2002). Evaluation is part of supervision (Zepeda, 2007).

High performing districts. Missouri districts were classified as high performing based on the number of “Mets” on the APR’s six MSIP 4 academic achievement standards and indicators for years 2010, 2011, and 2012. High performing districts were those who received a “Met” on each of the six academic indicators for each of the three years selected. Non-high performing school districts were those who received a “Met” on less than 80% of the eighteen (6 per year) academic indicators over the three year period.

Instructional supervision. Zepeda (2007) defined instructional supervision as that which “aims to promote growth, development, interaction, fault-free problem solving, and a commitment to build capacity in teachers” (p. 29).

Missouri Annual Performance Report. Fourteen measures of performance indicators were used during the 4th Missouri School Improvement Program (MSIP) Cycle to determine the accreditation level of a school district. Performance standards were evaluated using status and progress measures to determine if a standard was met. Status and progress points were combined to determine if a standard was met, unless no progress points were possible.

Non-high performing districts. Missouri districts were classified as non-high performing based on the number of “Mets” on the APR’s six MSIP 4 academic achievement standards and indicators for years 2010, 2011, and 2012. Non-high performing school districts were those who received a “Met” on less than 80% of the eighteen (6 per year) academic indicators over the three year period.

Professional development. For this study, professional development is defined as the teacher’s or supervisor’s focus on the development of professional expertise using problem solving and inquiry (Sergiovanni & Starratt, 2002). This is also categorized as 23 parts of supervision (Zepeda, 2007a).

Supervisor. In this study, the term supervisor referred to the principals, assistant principals, and instructional coaches who were involved with the supervision and evaluation processes.

Study Overview

The goal of Chapter One was to introduce the study by providing a brief overview. Beginning with an introduction, the chapter provided a background discussion of how teacher supervision practices have evolved and why teacher supervision has been brought to the forefront of education reforms. The research problem and purpose of the study explained why this study was worthwhile. The introductory chapter also included the significance of the study and who will potentially benefit from reading the study. The final section of Chapter One included an introduction of the research design. The contents of ensuing chapters are described below.

In Chapter Two, the reader found an overview of literature and evaluation models as it relates to the research questions and design of the study. The literature was divided in five

major components: Effective teachers, principles of teacher supervision, the evolution of supervision, current research in effective supervision practices, and principal professional development to promote effective teacher supervision.

Chapter Three described the qualitative and quantitative research methodologies, procedures, research design, and data analysis used in this study. Incorporated into Chapter Three was an explanation and rationale for the various methodologies employed in the study. Research ethics and limitations were also explained.

Chapter Four provided the findings for this study. Included in this chapter was an analysis and summary of the data in relationship to the purpose of this study. The findings were also matched in relationship to the research questions and what was learned through the principal interviews.

Chapter Five was the final chapter and provided a summary of how the findings of this study related to theoretical rationale as discussed in the literature review and through discussion of the research questions. Suggestions for further research and recommendations to current principals, teachers, district administrators, and university educational administration program personnel were also included in this chapter.

CHAPTER TWO

LITERATURE REVIEW

Introduction

In 2001, No Child Left Behind legislation was enacted requiring all students to perform at or above the proficient level on standards based assessments by 2014. Both policymakers and the public viewed student achievement being related to the quality of teaching. No Child Left Behind legislation held schools accountable for increasing student achievement, which, in turn, increased pressure to ensure the quality of teaching staff. In September 2011, President Barack Obama announced his administration would consider waivers from the No Child Left Behind law, in exchange for rigorous and comprehensive State-developed plans designed to improve educational outcomes for all students, close achievement gaps, increase equity, and improve the quality of instruction. Since 2011, 37 states and the District of Columbia have been approved for Elementary and Secondary Education Act (ESEA), better known as No Child Left Behind (NCLB) legislation flexibility waivers (U.S. Department of Education, 2013). Missouri's flexibility waiver was approved in June 2012 and contained a three-pronged approach to improving Missouri schools, college and career-ready expectations for all students, improved district and state accountability for all students, and supporting effective instruction and leadership. Because of these requirements, teacher evaluation began to receive unprecedented attention in Missouri and across the nation. Teaching has become a major focus of policy attention and teacher evaluation has become the primary tool to improve teaching. Federal requirements called for the use of multiple categories of teacher ratings, rather than satisfactory or unsatisfactory, and for ratings to be based on multiple observations, feedback, and the use of student test scores to assess teacher effectiveness (Darling-Hammond, 2013).

Teacher evaluation policies raise fundamental questions about what constitutes effective instruction and whether those practices can be fairly measured. These policies also tend to be highly politicized because they involve issues central to collective bargaining agreements between teachers' unions and school districts related to compensation, hiring and firing, and career advancement (Sartain, Stoelinga, & Brown, 2011).

Throughout the following literature review, "teacher supervision" will be defined as practices which promote teacher growth and development as well as practices which are used to make evaluative judgments about the teacher's performance (Clark, 1998). The argument for increased accountability from parents, taxpayers, and legislators is often in contrast with the views of many educational improvement theorists and educators who believe teacher supervision practices should be used for improving teaching practices and fostering teachers' professional growth (Clark, 1998, Danielson & McGreal, 2000, Sergiovanni, 2002). In fact, Marzano, Frontier, and Livingston (2011) conjectured improving teachers' pedagogical skills, thereby increasing student achievement, should be the foundational principle of supervision.

This comprehensive review of literature will explore three themes that are related to the research questions probed in this study. First, the review of literature discussed effective teaching and summarized current methods of teacher evaluation and from leading researchers in teacher evaluation. Second, the review was used to establish a framework for discussion of teacher evaluation practices with historical practices of teacher evaluation being examined. Of particular interest were practices in Missouri. Third, the review of literature explored research on what is known about effective teacher supervision.

Effective Teacher Characteristics

Ensuring each classroom had the most effective teachers in place and helping those teachers attain higher levels of effectiveness through professional growth was no simple task. In 2008, Toch and Rothman's report *Rush to Judgment* blamed lack of accountability for school performance, staffing practices of supervisors, union ambivalence, and the practice in public education of using teacher credentials as an indicator of teacher quality, for a superficial teacher supervision system which did not address quality of instruction and measure student learning (pg. 1). In 2009, *The Widget Effect* (Weisberg, Sexton, Mulhern, & Keeling, 2009) cited failure to provide accurate information about individual teacher instructional performance and a failure to recognize and respond to variation in the effectiveness of teachers. This failure to distinguish effectiveness among teachers made it difficult to identify areas for professional development which would ultimately improve teacher effectiveness.

In the United States, existing federal, state and local policies for defining and measuring teacher quality either rely almost exclusively on classroom observations by principals who differentiate little among teachers and offer little useful feedback, or focus on assessment information from tests of basic academic skills and subject matter knowledge which were poor predictors of later instructional effectiveness in the classroom (Darling-Hammond, 2010). To address this issue, twenty states joined together to form the Teacher Performance Assessment Consortium and created a common initial licensing assessment used nationwide to make preparation and licensing performance based, as well as predictive of teacher effectiveness. The assessment also included an advanced component used during the mentoring process and assessed the new teacher after the typical three-year probationary period and could be linked to additional compensation (Darling-Hammond, 2010).

Supervision of teachers or supervision of instruction inevitably leads to analysis of teacher effectiveness. Many researchers (Marzano, Frontier, Livingston, 2011; Marshall, 2009; Zepeda, 2009; Danielson, 2009; Valentine 2005) claimed supervision of teacher practices was an important tool in improving teacher effectiveness and one purpose of supervision was providing feedback to improve teacher effectiveness. It seems important then to start with a definition of an “effective teacher” (Himmelein, 2009, p. 13). Goe, Bell, & Little’s (2008) five-point definition of effective teachers consisted of the following:

- Effective teachers have high expectations for all students and help students learn, as measured by value-added or other test-based growth measures, or by alternative measures.
- Effective teachers contribute to positive academic, attitudinal, and social outcomes for students such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior.
- Effective teachers use diverse resources to plan and structure engaging learning opportunities; to monitor student progress formatively adapting instruction as needed; and to evaluate learning using multiple sources of evidence.
- Effective teachers contribute to the development of classrooms and schools valuing diversity and civic-mindedness.
- Effective teachers collaborate with other teachers, administrators, parents, and educational professionals to ensure student success, particularly the success of students with special needs and those at high risk for failure (p. 8).

Principles of Teacher Supervision

Research has shown the most effective teachers have the greatest impact on student achievement (Marzano, Frontier, & Livingston, 2011; Marshall, 2009; Clotfelter, Ladd, &

Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004). Darling-Hammond defined teacher quality as personal traits, skills, and understanding an individual brings to teaching and cited the following qualities to be important in defining teacher effectiveness:

- Strong content knowledge related to what should be taught,
- Knowledge of how to teach others in his or her teaching area (content pedagogy) and skill in implementing productive practices,
- Understanding of learners and their development, including how to support students who have learning differences and difficulties,
- General abilities to organize and explain ideas, as well as to observe and think diagnostically, and
- Adaptive expertise allowing teachers to make judgments about what is likely to work in a given context in response to student needs (Darling-Hammond, 2013).

Researchers (McMillan & Schumacher, 2001; Nolan & Hoover, 2008; Marshall, 2009; Marzano, Frontier, & Livingston, 2011) have acknowledged the purpose of teacher evaluation was to make personnel decisions and to measure teacher competency for the improvement of instruction. Danielson labeled these two essential purposes as quality assurance and improving teaching (Danielson, 2008). Researchers (Danielson & McGreal, 2000; Marzano, Frontier, & Livingston, 2011; Zepeda 2007) have found formative evaluation facilitates the developmental process of educational practitioners. School administrators engaged in formative evaluation activities direct, coach, and support the teachers they evaluate with the end goal of improving student learning. These two purposes could also be categorized as summative and formative, respectively. Summative purposes were aligned more with accountability and competence and used for decision making concerning hiring, terminating, and promoting, whereas formative

purposes were aligned with professional development, growth, and improvement of teacher performance (Hughes, 2006; Nolan & Hoover, 2008). Nolan and Hoover further defined evaluation as the summative function of making judgments concerning teacher performance for purposes of personnel decisions as evaluation. Nolan and Hoover also defined supervision as the formative process of improving teaching performance and greater student learning, thereby, clarifying the difference between supervision and evaluation (Nolan & Hoover, 2008). Marzano, Frontier, and Livingston (2011) defined the purpose of supervision as the enhancement of teachers' pedagogical skills with the ultimate goal of enhancing student achievement.

Zepeda's (2007) definition of teacher supervision has a cyclical, three-pronged approach to supervision: instructional supervision, professional development, and evaluation. The three aspects included in Zepeda's supervision approach are all integrated and each is part of the supervision process as a whole. All of the facets were essential threads necessary to complete the entire representation of teacher supervision. Leadership of the supervisor was critical to promote and champion the connections and to work proactively with teachers supporting professional learning (Zepeda, 2007). According to Zepeda (2007), supervisors first needed vision shaped by beliefs and values about students, learning, and leading.

Substantial research has shown the principal who exhibited effective instructional leadership impacted student achievement (Ebmeier, 2003; Colby, Bradshaw, & Joyner, 2002(b); DuFour & Marzano, 2011; Marzano, Waters, & McNulty, 2005). Ebmeier (2003) produced research which linked teacher efficacy to supervision and defined efficacy as an individual's belief about his or her capacity to achieve success in a given situation. Ebmeier (2003) cited teacher efficacy as a key to improving the quality of schooling and significantly related to many valued outcomes, including teachers' classroom behavior, student learning, and change in

teacher practice through staff development. Active principal supervision in the form of frequent classroom observations and conferencing activities did not directly influence a teacher's confidence, trust, or support of the principal (Ebmeier, 2003). Using scales to measure the extent a principal was engaged in activities associated with supervision, such as classroom observations, feedback to teachers, and goal setting, a principal's support of teaching, and teacher's satisfaction with working conditions and commitment to a school, Ebmeier's data determined principal supervisory behaviors and the beliefs of being able to affect student learning for 28 teachers in the principal's school were remarkably similar (Ebmeier, 2003). Ebmeier (2003) stated active supervision must be accompanied by other activities focused on classroom teaching, such as placing greater emphasis on teaching, rewarding sound teaching, and providing technical and symbolic leadership. These activities help convince teachers of the high value the principal places on the instructional process.

Colby, Bradshaw, and Joyner (2002) reviewed and synthesized data from 57 research studies. The researchers found five foundational criteria for effective teacher evaluation in the 1980's and trends developed after the mid 1990s focusing in school reform, school cultures, and the connection between teaching and learning processes. The criteria included: (1) purposes for evaluation matched the methods/procedures, (2) district commitment was evident in sound policies and practices and appropriate allocation of time and resources, (3) teacher evaluation was tightly connected to district priorities and school functions such as school improvement, professional development, and student learning, (4) educational leaders played a strong, positive role in evaluation, and (5) the evaluation environment was supportive of ongoing, professional learning.

Colby, Bradshaw, and Joyner (2002) used both quantitative and qualitative data to determine if educators perceived a stronger impact on school improvement, professional development, and student learning with locally developed or state-mandated teacher evaluation systems. Colby, Bradshaw, and Joyner's research indicated the locally developed evaluation systems were better able to support district school reform initiatives, guide professional development for teachers, and use student learning as a focus for teacher evaluation (2002). The researchers identified two implications from the study. First, teacher evaluation policies and practices could serve as a catalyst for creating connections in practice among school improvement, professional development, and student learning and could be used by districts for creating a greater impact in these areas. Secondly, Colby, Bradshaw, and Joyner (2002) found development of teacher evaluation systems at the local level as a viable strategy for strengthening teacher evaluation and its connections to school improvement, professional development, and student learning (p. 8). Their findings demonstrated that by making decisions at the district level, a structure could be created to channel district resources into support for the district's school improvement initiatives, the increased professional development of teachers, and student learning by making connections in practice.

Blasé and Blasé (2004) defined supervision as a combination of supervisory beliefs and philosophies with the purpose of building trust, empowering teachers, and fostering reflection. Blasé and Blasé (2004) maintained supervision should be inquiry oriented, encourage teacher participation, and acknowledge the context and complexity of teaching. Blasé and Blasé (2004) noted a connection between supervisory actions and professional growth of teachers, teacher commitment, involvement, innovativeness, and increased student learning.

According to Danielson and McGreal (2000) teacher evaluation systems must be able to answer two questions: (a) what do we believe good teaching looks like, and (b) what are the processes and procedures we want our system to accomplish? Shortcomings existed in teacher supervision because school staff lacked the time, training, and inclination to become knowledgeable about the best evidence from research on teaching and the difficulty of changing what had always been done or believed.

Teacher evaluation systems failed to develop a highly skilled workforce and accurately measure teacher quality because the evaluation systems did not discriminate between effective and ineffective teachers (Marzano, 2012). Marzano wrote measuring teachers and developing teachers were different purposes with different implications. Marzano suggested measuring teacher performance in terms of pedagogical skill required fewer criteria than developing teachers. For example, to measure a teacher, a principal would expect them to be able to provide clear learning goals, and most would agree failure to achieve a certain level of competence in providing clear learning goals would cause teaching to suffer. In contrast, developing teachers required ratings on more criteria in order to identify areas of strength or weakness and systematically begin improving the areas of weakness (Marzano, 2012). Marzano included many learning strategies used by teachers which lead to increased student achievement, however, he also suggested some strategies may not be used at all and gains in student learning may still be produced. For example, using physical movement has been shown to increase student achievement, however, many teachers produce significant gains in student learning without using physical movement strategies. In addition to a comprehensive system containing specificity about strategies and behaviors, research by Marzano suggested teacher supervision systems with the goal of developing teachers had to include a developmental rating scale used to guide and

track teacher skill development. Marzano described clear levels of performance with detailed descriptions for each of the following scales: 0) not using, 1) beginning, 2) developing, 3) applying, and 4) innovating (Marzano, Frontier, & Livingston, 2011). The five levels were designed to enable teachers, with the aid of a supervisor or instructional coach, to pinpoint his or her current level of performance and set goals for improvement (Marzano, 2012). The third characteristic of a system with the goal of developing teachers was acknowledging rewards and growth. Marzano suggested teachers identify elements for improvement and chart progress throughout the year. Supervisors would use the progress to score a teacher on the extent to which they reached growth goals on summative evaluations (Marzano, 2012).

Evolution of Teacher Supervision

Before the 1950s, teacher supervision was not a standardized process and there were no procedures for administrators to follow when they entered the classroom or interacted with teachers about professional development (Watkins, 2011). Therefore no system of measurement was in place for the administrator to use to evaluate teacher performance (Nolan & Hoover, 2008). Teacher evaluation was defined from a moralistic and ethical perspective, and thus, teachers were judged on their moral standing in the community rather than based on a knowledge base of effective teaching and learning practices (Ellett & Teddlie, 2003). During the 1920s-1940s, teaching and teacher evaluation focused on personal characteristics of teachers due to emerging theories of personality and personality characteristics in psychology. There were studies to identify factors contributing to the education and training of prospective teachers. In the early 1940s, several conceptual frameworks for evaluating teaching began to appear in the literature and texts (Ellett & Teddlie, 2003). During the 1950s and 1960s, scientific management and behaviorism in psychology and education spurred research in educational practice which

began to identify the linkage between teacher behavior and student outcomes (Ellett & Teddlie, 2003, p. 104). There were increased efforts to identify effective teaching methods.

Clinical supervision. As a result of competition between the United States and Russia and the launch of Sputnik in 1957, pressure to improve student results in science and mathematics led to developing a standardized approach for supervising teachers and observation instruments which accurately reflected classroom practice. Supervisory skills and classroom observation techniques also made advances leading to clinical supervision (Danielson & McGreal, 2000). Researchers developed clinical supervision, as a “systemic approach to teacher evaluation” (Watkins, 2011, p. 37). Clinical supervision relied on the development of a relationship built on mutual trust between a supervisor and teacher. Krajewski and Anderson (1980) cited the goal of clinical supervision as helping teachers better perform a job according to their capabilities with the role of the supervisor as a teacher of teachers.

In 1969, Robert Goldhammer published *Clinical Supervision*, which documented a five step process designed to involve teachers and supervisors in a reflective dialogue.

- Phase 1 – Preobservation Conference: This was a meeting between the supervisor and the teacher before the observation to establish rules and plan the specifics of the observation.
- Phase 2 – Classroom Observation: During this phase, the supervisor observed the teacher and collected the data agreed upon in Phase 1.
- Phase 3- Analysis: Data from the observation was organized by the supervisor into a format understandable to the teacher. Patterns and themes emerged and were discussed in Phase 4.

- Phase 4 – A Supervision Conference: The teacher and supervisor engaged in a dialogue about the data. The teacher was asked to reflect upon and explain his or her professional practice. This stage also could include providing didactic assistance to the teacher.
- Phase 5 – Analysis of the Analysis: The supervisor’s practice was examined and a plan of action is developed for the next cycle of supervision.

Morris Cogan conducted research published as the book *Clinical Supervision* in 1973. Cogan expanded Goldhammer’s stages into eight phases of clinical supervision. Cogan’s first phase, which he viewed as the most important, was building a relationship between the supervisor and the teacher with the common goal of improving teaching. The second phase was the supervisor and teacher planning the lesson together, developing expected outcomes, goals, objectives, activities, materials, and so on. Cogan’s third phase was equivalent to phase two of Goldhammer’s model; both parties developed a plan for the collection of data, and the rules for the observation were established. The fourth phase was, as with Goldhammer’s model, observing in the manner set forth in phase three. In phase five, the teaching-learning process was analyzed. This phase differed from Goldhammer in that Cogan believed the teacher and supervisor should analyze the data together. In the case of an inexperienced teacher, there may have been a need for some coaching in data analysis. Phase six consisted of planning the conferencing strategy. Both parties participated in this stage, which was essentially developing a plan for the post-observation conference. Phase seven was the conference, wherein the teacher and supervisor met and exchanged ideas about the observed data. The last phase was renewed planning. This eighth phase was to develop a plan for fixing problems and lay out a new plan for the teacher. The cycle then started over again (Minnear-Peplinski, 2000).

By the 1980s clinical supervision was used by about 90 percent of school administrators. Both Goldhammer and Cogan emphasized just going through the steps in a mechanical manner would not result in substantially improved teacher behavior. A genuine air of collegiality and mutuality in the relationship between supervisor and teacher had to exist. The supervisor must view his or her role as trying to help the teacher develop more effective methods rather than imposing theories (Reavis, 1978). The phases of clinical supervision were intended to be a vehicle to disclose effective instructional practices through the rich dialogue envisioned by Goldhammer. However, over time, the phases of clinical supervision became an end to themselves, and the dialogue was reduced to a ritualistic set of steps to be followed (Marzano, Frontier, & Livingston, 2011, p. 19).

Hunter model. In the late 1970s, Madeline Hunter, an educational psychologist, developed a theory-based way of looking at teaching which had its roots in a behavioristic view of basic learning theory (Danielson & McGreal, 2000). Although Hunter's model was called "clinical supervision," it had little similarity to Cogan and Goldhammer's work (Nolan & Hoover, 2008). Observation and scripting, where a supervisor recorded teaching behaviors and then later categorized them into those promoting learning; those using precious time and energy, yet contributed nothing to learning; and those, unintentionally, actually interfering with learning, were critical components of Hunter's process of supervision (Marzano, Frontier, & Livingston, 2011). Combined with clinical supervision, Hunter's seven step model became known as "mastery teaching" and was the topic of the preconference observation, classroom observation, and the supervision conference. Hunter's seven elements were:

- Anticipatory Set - Focusing the learning task, its importance, or the prior knowledge/experience of the learners.

- Objective/purpose - Select an objective at an appropriate level of difficulty and complexity, as determined through a task analysis, diagnostic testing, and/or congruence with Bloom's cognitive taxonomy.
- Input - Identify and teach main concepts and skills, emphasizing clear explanations, frequent use of examples and/or diagrams, and invite active student participation.
- Model – Show students several examples of the process or products students are expected to acquire or produce.
- Check for understanding - Observe with frequent formative evaluations and immediate feedback. Adjust instruction as needed and reteach if necessary.
- Guided practice – Students practice under direct teacher supervision so immediate feedback and reteaching may occur, if necessary.
- Independent practice - To solidify skills and knowledge after the teacher is reasonably sure students will not make serious errors (Marzano, Frontier, & Livingston, 2011)

Hunter's model encouraged teacher-centered, structured classrooms and supervision. Evaluation practices during this time were often driven by checklists and rating scales of teacher behaviors which presented a simplistic view of teaching (Danielson & McGreal, 2000; Nolan & Hoover, 2008). Wolfe maintained that Hunter described the seven elements and intended the elements should only be "considered" when designing a lesson (Wolfe, 1987). Teachers began to try to fit all elements into every lesson, and administrators, following suit, began to look for all elements in observations.

Developmental/reflective models. The 1980s also began a period marked by extensive research on teacher evaluation and a shift from local, district policies to state mandated, on-the-job assessments and evaluations for the purpose of licensure (Himmelein, 2009). Professional growth and development for teachers came to the forefront. While clinical supervision and mastery teaching was designed for the masses, the idea of individualized or differentiated teacher supervision began due to the work of William Glatthorn, Carl Glickman, and Thomas McGreal. Developmental supervision and reflective supervision were similar to the Cogan and Goldhammer models in terms of collaborative efforts to improve teaching and insisting evaluation and supervision be viewed as separate activities (Nolan & Hoover, 2008).

McGreal recommended new teachers be supervised intensely while more experienced staff were to be placed on a more self-directed regimen of professional development (1983). Spending several years working with over 75,000 instructional supervisors McGreal called for the purpose of supervision for tenured teachers primarily to be to help teachers improve instruction (1982). In developing commonalities associated with desirable supervision practices, McGreal cited districts often claimed the primary purpose of evaluation was improvement of instruction. However, in practice, actual supervision methods using procedures and instrumentation that required ratings on standardized criteria heavily loaded toward administrative concerns, resulted in unfocused classroom visitations (1982).

Glickman, Gordon, and Ross-Gordon (1998) defined developmental supervision as “the match of initial supervisory approach with the teacher or group’s developmental levels, expertise, and commitment” (p. 197). Ascertaining the conceptual level of teachers was central to the success of developmental supervision so supervisors could apply a supervisory approach to match the level of the teacher (Zepeda, 2007). Developmental supervision matched the

supervisory approach with the teacher's current developmental levels and immediate situation. Developmental supervisors also modified supervisory behaviors to promote and accommodate long-range teacher development toward high levels of reflection and problem solving ability (Glickman, Gordon, and Ross-Gordon, 1998). The instructional leader operating in developmental supervision gives three types of assistance: (a) directive, (b) collaborative, and (c) nondirective. Teachers with low developmental levels, expertise, and commitment required the structure provided by directive supervision. Teachers with a more moderate developmental level, expertise, and commitment could be best served by a collaborative supervisory approach. The collaborative supervisory approach allowed teachers to generate their own possible solutions to instructional problems, but the supervisor still played a role in offering assistance. Teachers who functioned at a higher developmental level, had high levels of expertise, and a strong commitment could be supervised in a nondirective approach. These higher functioning teachers could be more self-directed and explore their own methods to instructional improvement (Glickman, Gordon, and Ross-Gordon, 1998). The approach the supervisor took depended upon the developmental level of the teacher (Glickman et al., 2013).

Glatthorn's *Differentiated Supervision* (1997) explained teachers should have some input and sense of control over their own development. Differentiation allowed the supervisor to focus clinical supervisory practices on those teachers who would derive the greatest benefit from the clinical approach. Different opportunities for professional growth were to be provided for teachers based on individual needs. Glatthorn's model included three developmental options and two evaluative options (1997). Developmental options gave teachers choices in their own professional development but included intensive development for non-tenured teachers and tenured teachers with serious instructional problems. Intensive development focused solely on

improving teaching methods with the goal of improving student learning (Glatthorn, 1997). Cooperative development groups of teachers worked together to develop professionally using a variety of strategies. Self-directed development allowed teachers to work more independently by setting a growth goal and working to achieve the goal with the support of the principal. Glatthorn provided two evaluative options. First, Glatthorn provided intensive evaluation, a process which required many instructional leader observations which focus on learning outcomes instead of teaching methods, for teachers working in intensive development. Secondly, Glatthorn provided standard evaluation for those not using intensive development. Intensive development was used with a small number of teachers who experience difficulties with the teaching process. Glatthorn cites several factors contributing to classroom complexity. Organizational factors, such as school culture and work conditions; student factors, such as varying developmental levels and motivation; instructional supports in the form of curriculum guides, text-books, and tests; and teacher factors, including the teachers' knowledge and motivation level should all be examined by the supervisor to determine the cause for lack of learning (Glatthorn, 1997).

In *Supervision of Instruction: A Developmental Approach*, Glickman (1998) described supervision as the glue of a successful school by linking instruction and classroom management with professional development, direct assistance to teachers, group development, curriculum development, and action research under a common purpose to achieve school objectives. Glickman described supervision as the actions enabling teachers to improve instruction for students. Glickman called for a paradigm shift from conventional schools, characterized by dependency and hierarchy, and congenial schools, characterized by friendly social interaction, where teachers tend to be professionally isolated to a collegial model characterized by purposeful adult interactions about improving school-wide teaching and learning practices. Glickman's

supervision suggested a more a collegial rather than hierarchical relationship between supervisors and teachers. Where supervision became the responsibility of teachers and supervisors, the focus was on teacher growth rather than teacher compliance, and supervisors facilitated teachers collaborating with each other in instructional improvement efforts and teacher involvement in ongoing reflective inquiry (Glickman, 1988).

Concerns over effective school research utilizing standardized literacy and mathematics tests to identify effective schools and correlational rather than causal results, led researchers to focus on school improvement research. Effective school research takes statistical snapshots of performance measures and their correlates, but school improvement research is concerned with how a school is successful (Glickman, 2013). Research and understanding of cognitive learning theory led to a greater understanding of how children learn and have expanded research in effective teaching practices.

Even though the developmental and reflective models continued to be utilized, Nolan and Hoover (2008) blamed the “longstanding tradition of teacher evaluation as an important function of supervisors” for preventing the schools from utilizing these collegial supervisory models to the fullest (p. 5). The federal mandates from No Child Left Behind and Race to the Top increased pressure to focus more on teacher evaluation than supervision. Because of this pressure, Nolan and Hoover (2008) contend “many districts often embrace collegiality in rhetoric, but continue to practice inspectorial evaluation as the predominant mode of interaction between teachers and supervisors” (p. 5). Ellett and Teddlie (2003) contend statewide systems of classroom based teacher evaluation in the 1980s and 1990s contained a fundamental flaw of focus on teacher behavior and teacher performance with little attention on the connection between teaching and student learning.

RAND Study. The publication of *A Nation at Risk* in 1983 by the National Commission on Excellence in Education (NCEE) emphasized the need for higher public school accountability and the application of research-based practices to the supervision process (Henry Barton, 2010). The recommendations of the NCEE asserted an effective supervision system should be used to reward superior teachers, encourage average teachers, and improve or terminate poor teachers (NCEE, 1983). Along with concerns about the U.S. economy and a changing job market, the NCEE report also led to increased pressure for educators to help students attain complex skills such as critical thinking, problem solving, lifelong learning, collaborative learning, and deeper understanding (Danielson & McGreal, 2000). A report by Wise, Darling-Hammond, McLaughlin, & Bernstein in 1984, entitled *Teacher Evaluation: A Study of Effective Practices*, often referred to as the RAND Study in the literature, reported four consistent problems with supervision. Systems of supervision from 32 districts were studied and many systems were found to be prescribed or formulaic in nature. The study also found when supervisory approaches were more developmental and reflective, participants viewed them as not specific enough to enhance pedagogical development (Marzano, Frontier, & Livingston, 2011, 23). Nearly all respondents reported principals “lacked sufficient resolve and competence to evaluate accurately” (p. 22). Teacher resistance or apathy was the second most frequently cited problem. The study pointed out much of the discomfort for teachers came from the third problem area: lack of uniformity and consistency with a school system. Even though supervision instruments had become more uniform, many teachers still believed the process relied too much on judgment and bias of the principal which led to different ratings for similar practices among different schools. The fourth problem was inadequate training for supervisors. Supervisors were found either to have not had

enough training or the training received provided insufficient guidance on how to supervise (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984).

Danielson model. In the mid-1990s Charlotte Danielson's model sought to capture the complexity of teaching, provide a language for professional conversation, and provide structure for self-assessment and reflection (Marzano, Frontier, & Livingston, 2011). Danielson (2007) described her framework in four domains. Planning and preparation, classroom environment, instruction, and professional responsibility were integral to establishing an evaluation process encouraging ongoing development of teaching practices and improved student learning. Danielson sought to accomplish three goals with her framework. The framework sought to acknowledge the complexity of teaching and confirmed "teaching is also cognitively demanding" (Danielson, 2007, p. 2). Secondly, Danielson (2007) established a common language for professional conversation so educators could have rich conversations about exemplary practices. Finally, Danielson's framework provided a structure for self-assessment and reflection on practice to strengthen the educators' own practice (2007).

Many evaluation systems in place today are rooted in what educators believed about teaching in the 1970s and based on the work originally done by Madeline Hunter (Danielson & McGreal, p 3). The research on student learning available were norm-referenced, machine-scorable, and multiple choice items of low-level knowledge. Current goals for student achievement have expanded to more complex learning, problem solving, and application of knowledge. (Danielson & McGreal, p 3).

Missouri evaluation models. In 1983 Missouri legislators enacted a bill mandating "comprehensive performance evaluation for each teacher" (Valentine & Harting, 1988). A 24 person committee developed a model performance evaluation system for teachers called the

Missouri Performance-Based Teacher Evaluation (PBTE) system. The Missouri PBTE model was updated in 1999 with input from teachers, principals, superintendents, university personnel state department personnel, and state legislators from across Missouri. (Missouri Department of Elementary and Secondary Education Guidelines for Performance-Based Teacher Evaluation, 1999). Districts could choose to create their own evaluation model or use the PBTE model. Both of the state's versions of the Performance-Based Teacher Evaluation Model included evaluative and professional development criteria. The PBTE included six standards with 20 criteria. The Missouri PBTE called for non-tenured teachers to be formally evaluated on an annual basis. Tenured teachers were evaluated on a five-year cycle or as often as deemed necessary by the supervisor. For first-year teachers the PBTE called for a mentor to be provided to assist the teacher with developing his or her evaluation portfolio. Other non-tenured teachers were to develop a plan focused on professional development or enrichment of teaching strategies. The PBTE required for non-tenured teachers to have a minimum of one scheduled and two unscheduled observations annually during their first three years. After three years, non-tenured teachers were recommended to receive a minimum of one scheduled and one unscheduled observations annually. If the non-tenured teachers were not meeting expectations, a Professional Improvement Plan (PIP) was developed in response to an observed deficiency or concern regarding a specific criterion. Tenured teachers were encouraged to work with administration to select suitable goals to develop and implement a professional development plan. Evaluation of tenured teachers was on a five-year cycle unless the supervisor determined a more frequent schedule was needed. During the formal evaluation year tenured teachers had a minimum of one scheduled and one unscheduled observation. At the end of the evaluation process the administrator provided a summative evaluation of data gathered on the 20 criteria using a three

or four point rating scale. The three point rating scale measures were meets expectations, progressing toward meeting expectations, or does not meet expectations. The PBTE four-point rating scale included an additional rating of exceeds expectations. (Missouri Department of Elementary and Secondary Education Guidelines for Performance-Based Teacher Evaluation, 1999).

While the Missouri PBTE required tenured teachers to develop professional development plans, there was little emphasis on supervision of instruction in the PBTE, which required a formal evaluation of tenured teachers once every five years. The three or four point rating scale did little to give specific feedback.

In June 2010, the Missouri Legislature passed Senate Bill 291 requiring public school districts in Missouri to adopt teaching standards. While the districts are responsible for actually adopting standards, the Missouri Department of Elementary and Secondary Education defined model standards for districts to use. The standards developed were based upon theories of teaching and learning and expectations that educators should continuously acquire new knowledge and skills and constantly seek to improve their practice to provide high academic achievement for all students (Missouri Department of Elementary and Secondary Education, 2011). In 2010 the Missouri Legislature directed school districts to adopt teaching standards which were to include the following elements: (a) students actively participated and were successful in the learning process; (b) various forms of assessment were used to monitor and manage student learning; (c) the teacher was prepared and knowledgeable of the content and effectively maintained students' on-task behavior; (d) the teacher used professional communication and interaction with the school community; (e) the teacher kept current on instructional knowledge and sought and explored changes in teaching behaviors to improve

student performance; and (f) the teacher acted as a responsible professional in the overall mission of the school.

The Missouri Model Teacher and Leader Standards adopted by the Missouri State Board of Education in June 2011, served as a resource for educators by defining a professional continuum to demonstrate how educators' knowledge and skills develop over their career (Missouri Department of Elementary and Secondary Education, 2011). The Missouri Teacher and Leader Standards were aligned to the Interstate new Teacher Assessment and Support Consortium (InTASC) standards created and distributed by the Council of Chief State School Officers (CCSSO). The InTASC standards were developed upon research and were aligned to the Common Core State Standards for students in mathematics and English language arts, the National Board for Professional Teaching Standards (NBPTS), the National Council for Accreditation of Teacher Education (NCATE) accreditation standards, the National Staff Development Council (NSDC; now called Learning Forward) professional development standards, and the Interstate School Leaders Licensure Consortium (ISLLC) 2008 educational leadership policy standards. Additional research and consultation took place with Mid-continent Research for Education and Learning (McREL), Dr. Laura Goe and the Educational Testing Service (ETS), Dr. Cheryl King and the Education Development Center (EDC), and Dr. Michelle Young and Dr. Ann O'Doherty and the University Council for Educational Administration (UCEA) (Missouri Department of Elementary and Secondary Education, 2011).

The Missouri Model Teacher Standards defined nine standards, along with quality indicators for each. The basis for the development of these standards was geared to the philosophy of a teacher supervision system that was formative and led to a continuous improvement, was aligned to standards which reflected excellence, led to a culture of informing

practice and promoting learning, and used multiple, balanced measurement that was fair and ethical (Missouri Department of Elementary and Secondary Education, 2011, April 18). From research, Missouri leaders learned that assessing teachers on every standard through observations, analysis of artifacts, collection of additional evidence, etc., would present an enormous burden on schools and districts. Second, prioritizing of standards that are most important to determine teacher effectiveness in the classroom may be required, while other standards would be “guidance” standards. Finally, it may be necessary to prioritize different aspects of the standards for teachers at different stages in their careers. (Missouri Department of Elementary and Secondary Education, 2011, April 18).

Current Practices for Teacher Supervision

Student achievement and growth models. Tucker and Stronge advocated for the use of student achievement as a component of teacher evaluation (2005). Many states have begun to hold teachers accountable for student scores on high stakes tests by including these scores on teacher evaluations. The Race to the Top initiative under President Barack Obama’s administration allowed for states to apply for federal funding if teacher evaluations include student achievement results. As many states sought to develop databases which tracked individual student achievement over time, educational research has brought changes to how teachers are evaluated. These longitudinal databases have allowed researchers to measure changes in the achievement individual student level.

The Oregon Teacher Work Sample Methodology (TWSM) had a goal of finding better ways to assess the complexities of teaching and the connection between teaching and student learning. By using formative and summative teacher reflection the TWSM model required teachers to document and present samples of their work as evidence of student progress in

learning (Tucker & Stronge, 2005). Tucker and Stronge (2005) interviewed program developers, university faculty, and teacher candidates and cited several advantages and disadvantages to the project. Advantages included focusing a teacher's work on good instruction, encouragement of teacher reflection and action research, and alignment between the goals and practice of teaching. Disadvantages listed were difficulty in providing traditional measures of reliability and validity for applied performance measures, and significant time required to design, develop, deliver, and assess work samples required.

Efforts to include student achievement as part of teacher evaluation at the Thompson R2-J School District in Loveland, Colorado began as a performance pay design (Tucker & Stronge, 2005). The standards-based teacher evaluation system identified 10 teacher professional standards to measure teacher performance. With such standards, the school principal is responsible for increasing the probability of advancing student achievement requiring teachers to collect and analyze student data to drive instruction, use multiple measures to document student growth, and implement strategies based on student achievement data. The Thompson model was connected to student achievement. Tucker & Stronge asserted the model provided a tighter connection among instruction, student learning, and teacher professional growth focused on student achievement (2005).

In 1992, the Tennessee State Legislature enacted the Education Improvement Act, and Tennessee became the first state to mandate teacher evaluations include a value added system of measuring student learning gains. The Tennessee Value-Added Assessment System (TVAAS) measured student academic progress on state tests already in place in Tennessee. While achievement indicated student mastery of content on one test, growth measured student learning over a period of time taking into account the student's prior performance (State Collaborative on

Reforming Education, 2012). TVAAS was designed to use a sophisticated statistical methodology to measure the influence that school systems, schools, and teachers have on indicators of student learning (Tucker & Stronge, 2005). While the Education Improvement Act required TVAAS data be used in evaluating teachers for whom data was available, it could not be the sole source of information used to make evaluative decisions. Tennessee required that teachers also work with principals to develop a professional development plan linked to the school improvement plan and reflective of data from the TVAAS teacher report (Tucker & Stronge, 2005).

In Ohio, schools used value-added data from the state Ohio Achievement Assessment to provide data for teacher evaluation. Students in grades 3-8 were tested yearly in reading and mathematics. While Ohio has mandated that teacher's evaluations come from student growth measures, the state has allowed three ways to measure student growth. Value-added data from the state Ohio Achievement Assessments were used when available. If value-added data are not applicable for a given subject or grade, other assessments provided by national testing vendors and approved for use in Ohio was used. For subjects without state assessments or approved vendor assessments, such as art or music, schools had to establish a process to create Student Learning Objectives (SLOs) to measure student progress (Ohio Department of Education, 2013).

Performance continuum models. With increased emphasis on professional development and teacher growth, a new model of teacher evaluation became popular. Danielson (2008) called for meaningful conversations about teaching and supervision to be grounded in clear definitions of practice reflecting professional consensus of educators in a particular school or district. Feedback must be specific to improve teaching. A simple checklist approach to providing feedback indicating strategies used or not used does little to enhance teacher expertise (Marzano,

2011). Checking a box only provided feedback as to whether or not the strategy was used and did not give any information to the teacher about the level of skills exhibited by the teacher using the particular strategy (Marzano, Frontier, & Livingston, 2011). Performance continuum models were developed to allow feedback to be focused with clear descriptions of levels of performance.

Marzano teacher evaluation system. Robert Marzano developed the Marzano Teacher Evaluation System based on a synthesis of his own research and theory. The primary purpose of the Marzano Teacher Evaluation System was to enhance pedagogical skills of teachers (Marzano, 2011). Marzano's Teacher Evaluation System was comprised of four domains: classroom strategies and behaviors, planning and preparing, reflecting on teaching, and collegiality and professionalism. Marzano defined 60 elements in the four domains, of which 41 elements were classroom strategies and behavior. This emphasis on Domain 1 reflected the importance of classroom strategies and behavior, causal link with student achievement, and complexity of the domain (Marzano, Frontier, Livingston, 2011). For each element Marzano described clear levels of performance; from zero, indicating the strategy was called for but not used, to detailed descriptions for each of the following scales; 1) beginning, 2) developing, 3) applying, and 4) innovating. Marzano based the scale on his own research regarding skill development. Skill development typically progresses through three stages:

1. *The cognitive stage.* The individual is learning about the particular strategy but not actually able to perform or even attempt the strategy.
2. *The associative stage.* The individual is trying out and experimenting with the strategy in changes and adaptations are made to address preferences of the learner.
3. *The autonomous stage.* The individual can perform the strategy fluently, without error. Much practice is required to complete this stage (Marzano & Toth, 2013).

At the “not-using” level the teacher was not aware of a particular strategy or had not tried it in his or her classroom, making this level consistent with the cognitive stage of skill development. At the beginning level, a teacher used a strategy with errors and omissions and was in the early associative stage of development. At the developing level on the scale, the teacher used the strategy without error and relative fluency and would have completed the associative phase. The applying level would be the autonomous stage, and a teacher using a strategy at this level would start to produce positive returns in student learning (Marzano & Toth, 2013). Marzano’s model contrasted with previous systems designed to give teachers feedback. Rather than strategies being checked off as used, the Marzano model sought to give feedback to the teacher about the level of skill exhibited, which in return provided more direction for improving the practice of the teacher (Marzano, Frontier, Livingston, 2011).

Missouri educator evaluation system. Missouri defined growth and learning of students as the primary responsibility of teachers and school leaders. According to the Missouri Department of Elementary and Secondary Education, “effective educator evaluation systems promote the improvement of professional practice resulting in the improvement of student performance” (2013, p. 1). Missouri’s Educator Evaluation System was created in 2011 and field tested in 174 districts across the state and piloted in 2012-2013. The Educator Evaluation System was refined and revised based on the feedback received from the 2012-2013 pilot project, representing over 20,000 teachers statewide (Missouri Department of Elementary and Secondary Education, 2013). The Missouri model, approved by the Missouri State Board of Education in May 2013, included the following research-based essential principles:

- Measures educator performance against research-based, proven performance targets associated with the improvement of student performance.

- Uses multiple ratings to differentiate levels of performance.
- Highlights a probationary period of adequate duration to ensure sufficient induction and socialization support for new teachers and leaders.
- Uses measures of growth in student learning as a significant contributing factor in the evaluation of professional practice at all levels and ensures a proficient or a distinguished rating cannot be received in educator performance if student growth is low.
- Provides ongoing, timely, deliberate and meaningful feedback on performance relative to research-based targets.
- Requires standardized, initial and periodic training for evaluators to ensure reliability and accuracy.
- Utilizes the results and data to inform decisions regarding personnel, employment determinations and policy regarding employment (Missouri Department of Elementary and Secondary Education, 2013).

Missouri NCLB Wavier Teacher Supervision Requirements

No Child Left Behind legislation also required public reporting of student test scores. Researchers recognized the worst possible use of test data for public reporting was the presentation of simple test averages by districts and schools. It was documented these simple averages were complicated by socio-economic factors outside of the control of schools and teachers and determining the effectiveness of the school or teacher by these aggregate scores alone is impossible (Adcock, 1995; Sanders, 2000; Wang, Haertel, & Walberg, 1993).

Recognizing the flaw with this model, the United States Department of Education through the No Child Left Behind legislation, required student test scores to be disaggregated by various socio-socioeconomic categories. According to Sanders (2000), meaningful comparisons were not

directly available because the grouping of students by socio-economic identification still left many other variables. For example, if the means were reported by ethnic group, then the influence of the educational attainment of the parents was hidden, etc.

The NCLB Waiver required states to address three principles: College and career-ready expectations for all students; state developed recognition, accountability, and support; and supporting effective instruction and leadership. The third principle is the articulation of Missouri's Educator Evaluation System. Districts would have the option of either adopting the state's model system or aligning their local system to the seven essential principles of effective evaluation (Missouri Department of Elementary and Secondary Education, 2012). The Missouri State Board of Education adopted Model Teacher and Leader Standards in 2011 which were organized into three professional frames which were, professional commitment, professional practice, and professional impact. The professional commitment frame was cited as a source of evidence related to the professional agreements a teacher and leader make as a result of their role as educator. This included the planning and the intent behind their efforts as an educator, or their quality as teacher or leader. The professional practice frame was defined as a source of evidence related to specific effective actions or behaviors in which a teacher and leader engages, or the measure of their teaching and leading. This includes the reactions, responses, and processes to which the educator engages. The professional impact frame was defined as a source of evidence related to the effect, consequence or result that occurs due to the behaviors and commitments of the teacher and leader. This includes the outcomes the educator intends as a result of the quality of the teacher and their teaching (Missouri Department of Elementary and Secondary Education, 2013). Quality indicators for each standard and professional continuum articulating multiple performance levels for each standard were also developed. Recognizing the importance of pre-

service teacher education, Missouri also drafted the Missouri Standards for the Preparation of Educators (MoSPE) which went into effect in March of 2013. The plan included yearly report cards on the colleges' performance, a more uniform evaluation of student teachers, and higher grade-point average requirements for those in the program. "What we're looking to build is a way to capture that performance and score whether or not that teacher is doing a good job in the classroom before they get into their own," said Hap Hairston, director of educator preparation for the Missouri Department of Elementary and Secondary Education (Bock, 2012).

According to the Missouri Department of Elementary and Secondary Education (2013), the Missouri Educator Evaluation System was predicated on the improvement of educator practice and used formative assessment that led to continuous improvement. The primary purpose of the Educator Evaluation System was to promote growth in effective practices leading to increased student performance. The Educator Evaluation System was aligned to the Missouri Teacher Standards which was a response to the need to develop and adopt teaching standards. In July 2010, the Missouri Department Elementary and Secondary Education organized and initiated a working group of key stakeholders and developed the Missouri Model Teacher and Leaders Standards. The Missouri Teaching Standards include nine standards with 36 quality indicators:

Standard 1: Content knowledge aligned with appropriate instruction

Standard 2: Student learning, growth and development

Standard 3: Curriculum implementation

Standard 4: Critical thinking

Standard 5: Positive classroom environment

Standard 6: Effective communication

Standard 7: Student assessment and data analysis

Standard 8: Professionalism

Standard 9: Professional collaboration (Missouri Department of Elementary and Secondary Education, 2013).

The Missouri Educator Evaluation System included a protocol to be used. The identification of indicators is essential to establishing a particular focus based on performances articulated in the indicators. The baseline data serves as a starting point by establishing a current level of performance. Strategies for improvement are identified and practiced. Meaningful feedback is provided regarding the extent to which the new strategies are addressing the area of focus. A follow-up rating provides indication of the amount of growth in performance that occurred. Reflection on the process and the amount of growth that did or did not occur informs whether this particular indicator remains an area of focus or whether there is a new area of focus. This sequence is an important component to the growth in educational practice that occurs in the teacher evaluation process (Missouri Department of Elementary and Secondary Education, 2013).

Network for educator effectiveness. Utilizing the Missouri Teacher Standards, the University of Missouri developed the Network for Educator Effectiveness (NEE) and piloted it in nine Missouri districts in 2011-2012. The NEE system provided multiple measures including; classroom observations, student surveys, units of instruction, professional development plans and measures of student performance. According to Marc Doss from the University of Missouri, feedback from supervisors was positive and supervisors were pleased at how NEE integrated each of the components to give a total picture teaching ability (M. Doss, personal communication, September 6, 2013). One unique feature of the NEE program was that

professional development resources were provided for each of the indicators to be evaluated. Scores on individual indicators could be compared to a teacher in other similar size school districts across the state to give each teacher a point of reference about their own performance. If an indicator was scored low, online professional development resources were available to help the teacher. Doss claimed that one of the biggest benefits to the NEE program was that professional development drives improvement (M. Doss, personal communication, September 6, 2013).

Highly Qualified Teachers vs. Highly Effective Teachers

There has been much legislation passed which requires accountability of the educational system, which in turn, has increased accountability to ensure that individual teachers are highly qualified. The enactment of the federal No Child Left Behind Act of 2001 introduced concepts such as “Adequate Yearly Progress,” based on annual testing, and “highly qualified teachers,” based on teacher credentials, as a way to improve United States education. Highly qualified teachers were deemed those who held at least a bachelor’s degree, were fully licensed or certified by the state in the subjects he or she taught, and could demonstrate competence in the subjects taught (Tucker & Stronge, 2005).

However, there must be a distinction between highly qualified and highly effective teachers. Recent research has shown quality of instruction was the single most important factor in student achievement. (Marzano, Frontier, & Livingston, 2011; Marshall, 2009; Clotfelter, Ladd, & Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004). Measuring teacher qualification in terms of credentials did little to guarantee that the teacher could provide effective instruction. The reauthorization of No Child Left Behind Act as the Elementary and Secondary Education Act (ESEA), made teacher effectiveness and leadership the most important topic in education

reform discussions. The President Barack Obama administration's \$4.35 billion Race to the Top (RTTT) Fund targeted various levels of reform, with the most significant emphasis placed on improving teacher effectiveness, specifically by strengthening evaluation procedures (United States Department of Education, 2009).

One of the goals of No Child Left Behind (NCLB) was to ensure that every student had the opportunity to learn from a well-trained, qualified teacher. The mistaken assumption behind NCLB's highly qualified mandate implied that when teachers were highly qualified, they would also be highly effective in improving student learning (Phillips, 2010). Phillips' (2010) research demonstrated the NCLB characteristics of highly qualified teachers were not necessarily associated with larger student achievement gains. Phillips (2010) cited NCLB's failure to identify teacher qualities which were consistently related to positive achievement gains and failure to facilitate consistent implementation of these standards. As a result, NCLB's highly qualified teacher mandate offered little in terms of ensuring all students (and especially at-risk students) received opportunities to learn from effective teachers. However, in Phillips (2010) research with first grade teachers some teacher characteristics associated with student achievement gains were identified including: subject specific graduate degrees and additional child development courses which suggested at least some aspects of teacher training were consistently related to student outcome.

Finland has frequently been cited as an exemplar of school improvement after rapidly climbing to the top of nations in the Organization for Economic Cooperation and Development (OECD) in the Program for International Student Assessment (PISA) in mathematics, science, and reading (Darling-Hammond, 2013; Sahlberg, 2010; Hargreaves & Shirley, 2009). These gains were attributed to five-year teacher university programs. Finland competitively recruited

top candidates from college graduates and provided up to a three year graduate level teacher preparation program free of charge with a living stipend. Teacher education in Finland was research-based, supported by scientific knowledge, and focused on thinking processes and cognitive skills used in conducting research (Salhberg, 2010). Finland's teacher education programs focused attention building pedagogical thinking skills (Salhberg, 2010). In Finland there was very little emphasis placed on formal teacher evaluation and much more emphasis on both pre-service teacher learning, and supporting teachers' collective learning in the field (Darling-Hammond, 2013; Salhberg, 2010). A broad-based curriculum ensured that newly prepared Finnish teachers possessed balanced knowledge and skills in both theory and practice. It also meant that prospective teachers possess deep professional insight into education from several perspectives, including educational psychology and sociology, curriculum theories, assessment, special-needs education, and pedagogical content knowledge in selected subject areas. Based on Finland's teacher education practices leading to effective teachers, Darling-Hammond called for support for teacher learning and evaluation to be part of an integrated whole promoting effectiveness at every stage of a teacher's career (2013).

Additional Educational Practices that Support Teacher Supervision

The literature also pointed to several components of successful teacher supervision practices. Professional learning, also known as professional development, peer coaching, mini-observations or walk-throughs, teacher portfolios, and student surveys were used in teacher supervision by principals and are reviewed below by the researcher.

Professional learning. Professional learning plays a role in the supervision of instruction and is a natural complement. Supervisors who connect the dots between supervision teacher evaluation and professional learning assisted teachers in identifying professional development

targets, provided resources, empowering teachers to build their own learning plans, and monitored results by being in the classroom (Zepeda, 2007, p 316). Glickman, Gordon and Ross-Gordon stated that professional development should be the responsibility of those who supervise (1998). Danielson and McGreal (2000) call for supervisors to review teacher developed plans for professional development to ensure the scope of the plan is appropriate, share ideas, and identify resources to support and enhance the teacher plan. According to Danielson and McGreal (2000) supervisors should not be asked to play the role of summative evaluators and then have no role in the required formative experiences. While professional development has been utilized in schools for years, research on how teachers learn and the best ways to education them is a relatively young field (Grossman & McDonald, 2008). Ericsson asserted that teaching to improve with practice but only up to a point. Ericsson called for high fidelity feedback and deliberate practice to develop expert practice in the teaching field. Ericsson (2006) noted that deliberate practice presented performers with tasks that provided growth on critical aspects were able to be refined after repetitions and feedback (p. 694). Dunn and Shriner (1999) identified teaching activities that met the criteria for deliberate practice. One important criteria included planning and preparation. Marzano and Toth (2013) assert the more effectively a teacher plans and prepares the higher probability that the teacher will utilize effective classroom strategies and behaviors. Careful planning should include the elements of effective scaffolding, progression toward a deep understanding, and attention to content standards. All three elements may not be easily observable during a single visit to the classroom, so the supervisor can obtain the information from analyzing unit planning documents, having a brief discussion with the teacher, and examining a teacher's written analysis of the unit (Marzano & Toth, 2013). The supervisor must provide opportunity for practice, high fidelity feedback by focusing the teacher on critical

components, and the teacher must be given time to reflect on the feedback (Coggshall, Rasmussen, Colton, Milton & Jacques, 2012).

Peer coaching. A building principal does not have to be the only source of supervision in a school. Peer observations and coaching provide opportunities for teachers to learn from each other (Marshall, 2009; Zepeda, 2007; Glickman, Gordon, and Ross-Gordon, 1998). Peer coaching pairs of teachers of similar experience and competence who observe each other teach, establish improvement goals, develop strategies to implement goals, observe one another during the revised teaching, and provide specific feedback (Bruce & Ross, 2008). Principals often have other priorities than providing direct assistance to teachers. Teachers naturally turn to other teachers for help more often than to a supervisor, and since supervision is concerned primarily with improving instruction rather than summative evaluation, Glickman, Gordon, and Ross-Gordon stated teachers helping teachers has become of formalized in way of assuring direct assistance to every teacher (1998). Clarity of purpose and goals were first needed to make peer observations effective. Peer coaching as defined in the literature is more than mentoring programs which help novices or master teacher programs to help struggling teachers. Peer coaching includes a pre-observation conference, an extended classroom observation, and a post-observation conference. Coaching occurs at two levels: in the classroom with the coach observing, a teacher and in the feedback conference.

Mini-observations/Walk-throughs. Marshall (2009) proposes short, unannounced classroom visits lasting approximately 5 to 10 minutes each, often referred to as mini-observations or walk-throughs, as the best way for principals to get an accurate sense of the quality of instruction students are experiencing on a daily basis, to give focused feedback to teachers, to identify teachers who are having difficulty, and to gather data for summative teacher

evaluations. Marshall (2009) recommends frequent visits to all classrooms yielding approximately 12 observations per teacher per year, along with individual feedback to each teacher, and ongoing communication about teaching and learning based on the observations.

Teacher Portfolios. Long used as a form of student assessment the use of portfolios for teachers has emerged to document growth and development. Zepeda (2007) describes a teacher portfolio as work in process that allows teachers to chronicle teaching practices, attainment of short and long-term goals, and knowledge gained through constructing artifacts. The ongoing nature of a teacher portfolio provides an enhancement to clinical supervision, mentoring activities, and action research. Zepeda and Mayers (2000) indicate that a teaching portfolio might include artifacts such as small philosophy of education, sample lesson plans and assessments, student work samples, committee work and interdisciplinary lesson artifacts, and items that show professional growth such as journals and videotapes.

Student Surveys. Hattie (2009) recognized the importance of student surveys in determining the quality of teaching and found that students were accurate judges of excellence and could discriminate between teachers who were experienced and expert from those who were experienced and non-expert. Research cited by Hattie, confirmed that students were adept at rating teachers based on what learning the student achieved rather than merely teacher behaviors. Arguments that student are capricious and were likely to award teachers high ratings was not supported by the research. The Bill and Melinda Gates Foundation Student (2012) reported that student feedback promises greater reliability because students have increased perspectives based on many more hours in the classroom. The Gates Foundation also found when students reported positive classroom experiences, those classrooms tended to achieve greater learning gains. The most predictive aspects of student perceptions are related to a teacher's ability to control a

classroom and to challenge students with rigorous work (Bill and Melinda Gates Foundation, 2011).

Principal Professional Development to Promote Effective Teacher Supervision

Research has shown a link between educational leadership behaviors and increased student achievement (Glanz, Shulman, and Sullivan, 2007; Waters, Marzano, McNulty, 2003; Owings, Kaplan, and Nunnery, 2005). The Wallace Foundation's Learning from Leadership Project provided evidence substantiating the contribution of school leadership was second in strength only to classroom instruction, and researchers cited successful leadership played a highly significant role in improving student learning (Louis, Leithwood, Wahlstrom, & Anderson, 2010). In these studies, supervisory practices were not examined Glanz, Shulman, and Sullivan, 2007; Waters, Marzano, McNulty, 2003; Owings, Kaplan, and Nunnery, 2005; Louis, Leithwood, Wahlstrom, & Anderson, 2010). In a case study on instructional supervision in New York City public schools, Glanz, Shulman, and Sullivan (2007) cited difficulty in distinguishing instructional leadership from supervision of instruction. Although the researchers' study did not provide definitive answers on how supervision made a difference in terms of promoting student achievement, Glanz, Shulman, and Sullivan (2007) called for future research into how instructional supervision works to influence teachers in classroom behavior and attitudes toward student learning which may affect student achievement levels. Glanz, Shulman, and Sullivan (2007) stated that interviews and observations conducted for the study did show that supervision was viewed as critical for enhancing teacher growth and that supervision was all encompassing from building a culture of reflection, collaboration, and improvement to encouraging leadership at all levels to offering faculty flexible and differentiated professional development on specific teaching strategies aimed to promote learning.

Lezotte (2011) proclaimed that a school principal must be a well-informed student of teaching and learning in order to be the instructional leader, yet also pointed to typical administration preparation programs which did not emphasize mastery of teaching and learning, curriculum, and instructional knowledge. In a survey of 385 practitioners, Lezotte asked for changes recommended for school leadership preparation programs. While Lezotte (2011) found few of the respondents had identified formal programs they felt were effective in attracting and maintaining effective school leaders, one strategy to develop principals into effective instructional leaders cited often by the respondents was to form collaborative groups in order to encourage principals to learn together and support one another.

Conclusion

Federal mandates and pressure from legislators and parents have led to increased accountability requirements for student achievement. Research has shown that effective teachers have the most impact on student achievement. Principals and district leaders are under even more pressure to ensure that teachers are highly effective and continue to improve. Supervision practices which lead to improvement of teaching are more critical than ever.

The literature review provided in this chapter, synthesizes a large amount of research related to the evolution of supervision, a look into supervision practices in Missouri, effective teaching, and current research in effective supervision practices. The results from this study will attempt to link principal perceptions regarding supervision of teachers with increased student achievement in their building.

Chapter Three presents the methodology, participants, survey instrument, and interview questions involved in this mixed descriptive, qualitative and quantitative study. Chapter Four

will provide the findings of this study with an analysis of the results and possibilities for research to follow in Chapter Five.

CHAPTER THREE

METHODS

The purpose of this mixed method, descriptive study was to describe the perceptions and practices of supervisors of instruction in both high performing and non-high performing Missouri schools. Of specific interest was the exploration of the perceptions and practices of supervisors of instruction in high performing Missouri schools. Tenets of Missouri's No Child Left Behind wavier and President Obama's Race to the Top have placed greater emphasis on developing effective staff to improve student achievement. With these higher stakes, teacher supervision and evaluation have become increasingly important for school leaders. The ultimate goal of this research project was to identify aspects of the supervision and evaluation process that are likely to lead to higher student achievement. The review of available literature revealed the purpose of supervision and evaluation was to enable teachers to become more effective and develop confidence in their teaching to successfully help students achieve (Ebmeier, 2003). This study was offered to help fill a gap in existing research regarding specific characteristics of supervision and evaluation practices as implemented in high performing school districts in Missouri.

This chapter describes the components of the research design and methodology applied to exploring the perceptions of principals regarding the supervision process. An explanation of the framework is provided including a summary of the research questions, a description of the research methodology, data collection procedures, interview questions, and data analysis.

Participants

Participants of this study consisted of principals and assistant principals who supervise teachers in public school districts in Missouri. Districts were classified as either high performing or non-high performing. The Missouri Department of Elementary and Secondary Education

relies on the Missouri School Improvement Program (MSIP) as the state accountability system for reviewing and accrediting public schools in the state. High performing districts were those who received a “Met” on each of the six MSIP 4 academic achievement standards and indicators on the Missouri Annual Performance Report (APR) or three consecutive years in 2010, 2011, and 2012. Districts who did not meet at least eighty percent of the total number of “Mets” were classified as non-high performing. There were 524 school districts in Missouri with APR scores for 2010, 2011, and 2012. The Missouri school district list was extracted from the Missouri Department of Elementary and Secondary Education Directory (DESE) available on the Missouri DESE website. This extracted list provided the name of the school district, the name of buildings within the school district, the name of the principal of each building, mailing address of the school, and email address of the principal. The surveys were limited to principals and/or assistant principals who had been in their current position for three or more years. The longevity of a principal in a building can have an impact on building performance (McDonald 2014, Bruggink, 2001). Hart (1993) found three or more years of supervision was necessary to allow time for the principal to influence instruction in the building. This study will focus on those principals in their current position for three or more years to ensure a minimum amount of time for a principal to influence instruction in the building. The Missouri Annual Performance Reports (APR) for years 2010, 2011, and 2012 for each district were emailed to the researcher from the Missouri Department of Elementary and Secondary Education Accountability Data and Accreditation department in an Excel spreadsheet.

The survey, informed email consent, ethics certificate and Research Review Board (RRB) application were sent to the RRB electronically and as a paper copy with the appropriate signatures in April 2014 for approval. Participants will give consent to be involved in the study

by completing the online survey. The submitted forms outlined participant confidentiality, the ability to withdraw at any time without penalty, lack of any foreseen harm to respondents and a brief overview of the purpose of the study to examine principals' perceptions of their current and ideal teacher evaluation process.

Sampling Procedure

This study employed a purposive sample with random selection. The researcher has used data from the Missouri Department of Elementary and Secondary Education's Missouri School Improvement Program, Cycle 4 (MSIP 4). The researcher classified districts as either high performing or non-high performing based on the number of "Mets" on the Missouri Annual Performance Report (APR)'s six MSIP 4 academic achievement standards and indicators for years 2010, 2011, and 2012. According to Missouri Commissioner of Education, Dr. Chris Nicastro, "Missouri is unique in our accountability system in that we insist on using multiple data points over multiple. Districts that meet academic standards over three consecutive years clearly show that they are on the right track, and the trend predicts sustainability" (C. Nicastro, personal communication, March 14, 2014). High achieving districts were those who received a "Met" on each of the six academic indicators for each of the three years selected. Non-high performing school districts were those who received a "Met" on less than 80% of the eighteen (6 per year) academic indicators over the three year period. Of the 524 Missouri school districts, 187 districts received a "Met" on each of the six academic indicators for 2010, 2011, and 2012. Of those 187 high performing districts, 50 districts were selected at random by computer random number generation to receive surveys. The remaining 337 were examined for the number of total "Mets" over the 3 year period. Districts who did not meet at least eighty percent of the total number of "Mets" were classified as non-high performing. There were 108 districts with less

than eighty percent of “Mets” on the 18 academic indicators over the three year period. Of these non-high performing districts, 50 districts were selected at random by computer random number generation to also receive surveys. The survey responses were limited to Missouri public school principals and/or assistant principals who had been in their current position for three or more years in 50 high performing districts and 50 non-high performing districts which were selected at random to send surveys.

The Missouri Department of Elementary and Secondary Education used a performance index to evaluate performance on the Missouri Assessment Program (MAP) state assessment. The index approach calculated the movement of students throughout all MAP achievement levels. Five years of data were analyzed by grade span (3-5, 6-8, and 9-11) for each subject area using status and progress measures. The status and progress methods were applied to each subject in each grade span. The progress method could only be applied when the same assessment is administered for two or more consecutive years. The method awarding the maximum total points from status (High 1, High 2, Average, Below Average, and Floor) and from progress (Annual, Rolling Average, and 3 over 2) was used for each subject area. The subject area/grade span standard was considered “Met” for grade spans 3-5 and 6-8 if the grade level data total was 40 status points, 50 status plus progress points, or 40 status plus progress points and the bonus gap was met. The subject area/grade span standard was considered “Met” for grade span 9-11 if the grade level and end-of-course test data combined total was 40 status points, 50 status plus progress points, or 40 status plus progress points and the bonus gap was met (Missouri Department of Elementary and Secondary Education, 2011).

High performing districts were further examined by looking at the criteria on the “Met” academic indicators. Status measured the district’s level of achievement and was based upon a five-year average of performance data. Status was divided into five levels as follows:

High 1 – 1 standard deviation above the mean for the state

High 2 – 1/3 of 1 standard deviation above the mean for the state

Average – Mean for the state

Below Average – 1/3 of 1 standard deviation below the mean for the state

Floor – 1 standard deviation below the mean for the state (Missouri Department of Elementary and Secondary Education, 2011).

Of the 50 high performing districts selected to receive a survey, five districts which received High 1 status on at least three of their 2012 APR six academic indicators were identified for interviews to gather further information about teacher supervision practices within the district.

The principals participated in the study by completing a brief survey about the supervision and evaluation process. The survey was developed by Dr. Pamela Clark in 1998. The volunteers were reminded that no personal or identifying information should be included on the survey. The survey cover letter, contained in an email, also informed participants of the right not to complete the survey (See Appendix C). It was explained to the participants that consent to use the data on the completed surveys was implied by the return of the survey. Because the educators were adults, this was considered minimal risk research and did not require a written signature to consent for inclusion.

Research Design

Both qualitative and quantitative research was used in this study to capture data on perceptions of principals who supervise teachers and to determine if the principals' perceptions of supervision influenced student achievement. This research began with an analysis of the literature to determine the current state of teacher evaluation and what had led to this current state. The literature review consisted of five areas of focus:

- I. Highly effective teachers
- II. Principles of teacher supervision
- III. Evolution of teacher supervision
- IV. Current practices for teacher supervision
- V. Principal professional development to promote effective teacher supervision.

The research questions associated with this study sought to gather information describing the perceptions of supervisors in high performing school districts and non-high performing school districts in Missouri regarding the supervision and evaluation process. To further delve down into the perceptions of high performing districts, the researcher chose to perform interviews with five of the districts who were found to be high performing. According to Gay, Mills, and Airasian (2012), mixed methods research combined quantitative and qualitative approaches by including both quantitative and qualitative data in a single study. The purpose of mixed methods research was to build on the synergy and strength existing between quantitative and qualitative research methods to understand a phenomenon more fully than would be possible using either quantitative or qualitative methods alone (Creswell, 2014; Gay, Mills, & Airasian, 2012). Therefore a mixed methods approach to conducting this study was most appropriate for answering the research question.

This study employed a random, cross-sectional sampling to develop a representative group of survey respondents as supervisors of instruction in school districts in Missouri. In order to get the maximum amount of information in the least amount of time, a cross-sectional survey design enabled the researcher to gather data about supervisors' perceptions of the supervision and evaluation processes used within their school in both high performing and non-high performing districts. The cross-sectional design facilitated the documentation of the supervisors' perceptions of the supervision and evaluation processes at the specific point in time when they completed the survey, providing a snapshot of the current behaviors, attitudes and beliefs of the supervisors (Gay, Mills, & Airasian, 2012).

The use of the Likert scales, open-ended inquiries from the survey, and additional interviews enabled the report of perceptions of the educators. These data collection methods also provided the triangulation of results of the descriptive quantitative analysis. Triangulation is used to get a more complete picture of the topic and to cross check information (Gay, Mills, & Airasian, 2012) The data were reported as Likert scores, percent agreement, and the percent mentioned related to predefined themes addressing the phases of the supervision and evaluation processes. Direct quotes from participants responses to open-ended questions and interviews were provided to further clarify the research findings.

The use of descriptive quantitative analysis methods enabled the researcher to compare the data from the two groups of supervisors (i.e. supervisors from high performing districts and supervisors from non-high performing districts). As a result, the researcher was able to identify commonalities and uniqueness reported among the perceptions indicated by the surveys. After describing and comparing data from the sample, inferences were made about the population (Gay, Mills, & Airasian, 2012). More specifically, within this study, the inferences reflected

approaches to be recommended for implementation or modification for the successful use of the adopted supervision and evaluation process. The data collected using the described methods were statistically assessed to analyze the variance and determine whether it could be concluded that the educators' perceptions about the adopted processes were shared. Validating the research hypothesis was expected to facilitate the identification of characteristics of the supervision and evaluation processes that were used by successful districts in Missouri. This study was conducted in 100 school districts in the state of Missouri during a two-month period.

The researcher also used qualitative research in this study to further attempt to gain a better understanding of human perceptions and experiences. The researcher used qualitative methods to compare teacher evaluation systems used in five of the high performing school districts which were identified and surveyed. This study offered an opportunity to examine a real life situation presenting a dilemma or an unsure outcome. The qualitative methodology was based on active participation by members of a group such as principals selected in this study. A key goal of this research was to find meaning. It was consistent with qualitative research to be interested in how different people make sense of their lives (Creswell, 2014; Bogdan & Biklen, 2007). The researcher sought to increase knowledge, which in turn could have contributed to the pool of knowledge, regarding the perceptions of principals in regard to teacher evaluation.

Within qualitative research, the constant comparative method was used. By using multiple participants and a systematic approach to data collection through interviews and open-ended questioning, study participants were allowed to articulate their perceptions and experiences freely and spontaneously. In analyzing data generated in this format, responses were not grouped according to predefined categories, but rather categories of focus developed from meaning and relationships derived from the data itself through a process of inductive reasoning

(Bogdan & Biklen, 2007). This method allowed the researcher to group relationships and themes found within the data.

Instrumentation

The Supervisor Survey (Appendix A) that was used for this study was developed by Dr. Pamela Clark (1998). Clark granted the researcher permission to use the survey via email. Clark (1998) developed the supervisor survey, along with a similar survey for teachers, to assess teacher and supervisor perceptions toward their present teacher supervision system and a supervision system the participants considered ideal. The survey for supervisors only, developed by Clark, was used for this study. Clark (1998) cited her literature review as her primary source for developing the survey, along with *Personnel Evaluation Standards* from the Joint Committee on Standards for Educational Evaluation (JCSEE). The final version of the survey used in this study consisted of 38 questions to be rated on a four-point Likert scale. Thirty one of the items (response items 1-19) were presented to the respondent to answer twice. The principals were asked to rate their perceptions on their present evaluation system currently used within the principal's district with one Likert rating scale, and the principals were also asked to do the same regarding their perceptions on an evaluation system they considered ideal. The last seven items (response items 20a-20g) asked the principal to rate effectiveness characteristics of their current system of supervision and evaluation over the last two years and were only presented to the respondent to answer once. Items asked the principal to rate their present evaluation system facilitated school improvement efforts, improved teacher classroom performance, and encouraged teachers to self-reflect about their teaching were part of the effectiveness scale. The seven questions from the effectiveness scale asked respondents specifically about the supervision and evaluation process used by the administrator for the past two years, so only one rating scale

was necessary. These 38 questions were classified into a practices scale, a beliefs scale, and an effectiveness scale for analysis.

The six-point scale used by Clark (1998) was converted by the researcher to a four-point scale for greater clarity. The four-point scale forced the participant to choose a direction for their response and did not offer a neutral response option. The Likert choices available to participants for this study were (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. The survey also included four open-ended questions regarding strengths, weakness, and recommendations for the current supervision/evaluation system and nine questions to gather demographic information. Survey questions were presented in four sections:

1. Section One: Demographic Information;
2. Section Two: Teacher Supervision/Evaluation Practices, Beliefs, and Values;
3. Section Three: Effectiveness of the Teacher Supervision/Evaluation System;
4. Section Four: General Perceptions (Clark, 1998).

Each of the four sections included multiple survey items identified as themes. The Likert scale portion of the survey encompassed multiple sub-themes including the district's foundational philosophies, phases of the supervision-evaluation process (i.e., pre-observation conference through professional development), and views of the collaboration between supervisor and teacher. The themes of each group of queries, in each section of the instrument, were used for the organization, presentation, and discussion of the data within this study. These themes addressed very specific concepts relevant to and with the evaluation of the use of the supervision and evaluation processes used in Missouri school districts studied herein.

Reliability of Instrument

To ensure validity of the surveys, Clark (1998) piloted the surveys with a panel of educators and also submitted the instrument to the New Hampshire Joint Education Council Executive Board for review and feedback. After obtaining feedback from the panel and the New Hampshire Joint Education Council Executive Board, Clark (1998) revised the surveys and had the surveys examined by additional educators. Clark (1998) stated the educators reported the items on the survey were worded clearly but the length of the survey was a concern. Clark (1998) continued to refine the number of questions and sought input from educators until the survey was able to be completed within 15 to 20 minutes. The final product was a 20 minute survey which diminished the threat to validity caused by incomplete surveys. For this study, only the supervisor survey developed by Clark was used.

Two threats to the validity of responses were bias and fear of identification. Clark (1998) reported these were not issues for her study because of the sample size chosen and administration procedures. For this study participants were informed not to leave any identifying information on the survey to reduce the opportunity or potential for bias. This method of anonymity ensured the researcher would be unable to identify the individual who completed each survey.

The survey required participants include demographic information. In order to encourage the participants to respond truthfully to the survey, the informed consent and cover letter explained the information was to be maintained confidentially. Responses from principals and assistant principals who had not been in their current position for three or more years were not considered in this study. The researcher committed to ensuring the findings were confidential by reporting summatively and using only identifiers such as District A, District B, and so on for reporting of data.

The survey instrument was assessed for its reliability by Clark (1998). Clark (1998) explained that a test of the reliability of the survey, along with subscales, was conducted by calculating the Chronbach's Alpha coefficient. Chronbach's Alpha estimates reliability by determining how all items on a test relate to all other test items and to the total test, to ensure the internal consistency of the elements (Gay, Mills, & Airasian, 2012). According to Clark (1998), the supervisor survey subscales had an alpha coefficient of 0.81. Because the reliability coefficient was greater than 0.50, it was concluded that the instruments were reliable. Specifically, the closer the alpha was to 1.00, the greater the internal consistency of items in the instrument being assessed (Gay, Mills, and Airasian, 2012). Clark found her instruments to be both reliable and valid as indicated by the statistical assessment of homogeneity. In addition, Clark was able to statistically compare the perspectives of the participant groups and logically address guiding research questions of Clark's study. The construction of similar format research questions for this study ensured the reliability and validity of the instrument was maintained.

Guidelines for Administration

For this study, surveys were emailed to principals of Missouri public school districts identified as high performing. Email addresses were obtained from the Missouri Department of Elementary and Secondary Education directory. A cover letter (See Appendix C) was included with the email and identified the purpose and significance of the project and the importance of the principal's participation. The letter also addressed the amount of time required for survey completion, the timeline for return of the surveys, and assurances of confidentiality. The email also contained a link to the QuestionPro survey tool site for the instrument. Each survey was coded as Group A or Group B along with an identifying number to allow the researcher to identify either high performing or non-high performing district surveys and sort results based on

these research categories. Survey participants were not able to access whether their survey was in Group A or Group B, as the information was only pertinent to the researcher. In order to reduce the potential for bias and encourage truthful responses without fear of identification of individual results, the letter contained assurances that information would be kept confidential and only aggregate survey responses would be reported. Principals who had not completed the survey within one week were sent a reminder email reiterating the purpose of the research and providing a link to the QuestionPro survey tool site. Individual interview responses that were reported were carefully selected to not include any identifiable information.

Data Analysis

The data was imported into a computer file in Microsoft Excel to facilitate statistical analysis and SPSS software was used to calculate statistics. Survey results were organized into current practices responses, ideal practices responses, and effectiveness of the current system. The principals were asked to rate their perceptions on their present evaluation system currently used within the principal's district with one Likert rating scale, and the principals were also asked to do the same regarding their perceptions on an evaluation system they considered ideal. Subgroups of data identified were current practices of high performing districts, current practices of non-high performing districts, perceptions of ideal practices of high performing district principals, and perceptions of ideal practices of non-high performing districts principals. Subgroups of data also identified professional development received by principals of high performing districts and professional development received by principals of non-high performing districts. Supervisor survey results were analyzed to identify perceptions related to their district's present teacher supervision system and to identify differences in their perceptions of an ideal teacher supervision system. Statistical analysis was done on survey results and used to determine

four main themes: 1) identify and describe perceptions of principals in high performing schools and principals in non-high performing schools, 2) if principals' practices of teacher supervision differed from the principals' ideal perception of teacher supervision, 3) if professional development in supervision and evaluation of principals in high performing districts differed from professional development in supervision and evaluation of principals in of non-high performing districts, and 4) to describe any variation between perceptions of supervisors in high performing schools and perceptions of supervisors in non-high performing schools regarding what?.

Rationale for Selected Statistical Treatment

The descriptive statistical analysis processes used to analyze the data in this study mirrored the ones used by Clark (1998) throughout the analysis of the data obtained from her initial use of the surveys. The researcher calculated the means and standard deviations for each subscale and the items within the subscale (Clark, 1998). Instances of non-response were coded as missing to eliminate errors in the calculation of the means. Rather than perform a repeated series of *t*-tests as Clark originally had done, the researcher chose to conduct a series of analyses of variance (ANOVA) to assess the presence of statistically significant differences or uniqueness between perceptions of principals in high performing districts and perceptions of principals of not high performing districts. The researcher chose the (Gay, Mills, & Airasian, 2012). A second series of ANOVA was conducted to assess the presence of statistically significant differences or uniqueness between supervisory practices of principals of high performing districts and supervisory practices of principals of non-high performing districts. An ANOVA test can only assess the presence or absence of statistically significant differences. This

information was reported using tables and graphs. The ANOVA does not assess the nature or degree of the similarities or differences (Gay, Mills, & Airasian, 2012).

When comparing high performing and non-high performing district principal perceptions regarding present teacher supervision and evaluation system and an ideal teacher supervision and evaluation system, the researcher used a *t*-test for non-independent samples. According to Gay, Mills, and Airasian (2012), the *t*-test for non-independent samples is used to compare groups that are formed by some type of matching.

Likert scaled queries. The data was preserved in a Microsoft Excel data file. Separate worksheets were generated for each school. The data from both files were merged as needed to conduct the comparative analyses of the data obtained from the two research categories of high performing districts and non-high performing districts. The data files were organized based on the query sub-sets. The response options included six-point Likert scale ratings, which were converted to Likert scores for each inquiry and mean theme-based Likert scores. The response options and values were: (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. The final version of the survey used in this study consisted of 38 questions to be rated on a four-point Likert scale. The six-point scale used by Clark (1998) was converted by the researcher to a four-point scale for greater clarity. The scale forced the participant to choose a direction for their response and did not offer a neutral response option.

Open-ended queries. The survey also included three open-ended questions regarding strengths, weakness, and recommendations for the current supervision/evaluation system. The data collected through the open response questions were assessed for common and unique themes regarding the perceived strengths, weaknesses, and recommendations for the improvement of the supervision and evaluation processes. The open response data was

summarized combining all respondent groups. The information gathered from the open response questions was documented as a theme. A multi-column table was generated for reporting the identified themes, percent mentioned, and the example responses associated with each theme. The table provided detailed insights that complimented the Likert scaled data.

The data gathered from the returned surveys were entered into the electronic database and sorted by respondent group. The survey items were sorted by themes and sub-themes. Each data set was compared in order to draw accurate conclusions that helped answer the research questions. A combination of tables and graphs were used to present the data.

Interview Questions

To further explore five districts receiving High 1 status on at least three of their 2012 Annual Progress Report (APR), six academic indicators were identified for interviews to gather further information about teacher supervision practices within the district.

The interview included a list of open-ended questions that was audio taped and transcribed. The purpose of conducting the interview was to provide the participants an opportunity to voice their experiences and opinions about teacher evaluation independently of the perspectives of the researcher or past research findings (Creswell, 2008). These interviews were conducted by phone. Although interviews conducted in person may have gained more in depth responses and additional data may have been obtained by reading the body language and visual cues of the participants, because cost and time of travel, the researcher chose to conduct interviews via phone. The interview questions aligned with the research questions. Questions one through five supported the first research question. Questions six supported the second and third research question. Question seven supported the fourth research question. A teacher from a local school participated in the piloting of the interview questions. Interview questions are listed in

Appendix D. Interviews were scheduled and conducted by phone and were audio recorded and transcribed by the researcher. In addition, the researcher also wrote descriptive field notes during the interviews related to the researchers direct observations about the interview and to chronicle the researcher own thinking, feeling, experiences, and perceptions throughout the research process (Creswell, 2014). Transcripts of the audio taped interviews and the researcher's descriptive field notes were used in a constant comparative method to code and analyze the interview data (Bogdan & Biklen, 2007).

Summary

The purpose of the proposed study was to describe and identify differences in the perceptions of principals in high performing school districts and non-high performing school districts in Missouri. The data for this study were collected using supervisor surveys that were created by Clark in 1998. The methodologies outlined herein were used for the documentation and analysis of the findings. Subsequently, the concepts were synthesized to develop professional discourse regarding the themes identified in the findings as well as the predetermined themes related to the components of the supervision and evaluation processes adopted at the study sites. The guiding research questions of the study were addressed in the research conclusions.

CHAPTER FOUR

RESULTS

Introduction

This study was proposed to gather information about supervisor perceptions of the teacher evaluation process and to determine if the supervisor perceptions had an effect on student achievement. This chapter was used to organize and present the gathered data. Statistical analysis was done on survey results and used to determine four main themes: 1) identify and describe perceptions of principals in high performing schools and principals in non-high performing schools, 2) determine if principals' practices of teacher supervision differed from principals' ideal of teacher supervision, 3) determine if professional development in supervision and evaluation of principals in high performing districts differed from professional development in supervision and evaluation of principals in of non-high performing districts, and 4) describe any variation between perceptions of supervisors in high performing schools and perceptions of supervisors in non-high performing schools.

Analysis of Data

A survey was distributed to principals of high and non high performing schools. The survey included four key sections. Section One contained nine general demographic questions. The second and third sections were used to collect descriptive quantitative data used to address the main research questions and the four subset research questions. The data collected from these sections of the survey included Likert rating scales with the following response choices: 1-Strongly Disagree, 2-Disagree, 3-Agree, and 4-Strongly Agree. Section Four contained four open ended response items about general perceptions of the respondent's current system of

supervision and evaluation, including strengths and weaknesses, recommendations for improvement, and professional development.

Descriptive Data.

The first section of the survey consisted of nine general demographic questions and helped define the population of this study consisting of principals and assistant principals who supervised teachers in public school districts in Missouri. Districts were classified as either high performing or non-high performing. The Missouri Department of Elementary and Secondary Education relies on the Missouri School Improvement Program (MSIP) as the state accountability system for reviewing and accrediting public schools. High performing districts were those who received a “Met” on each of the six MSIP 4 academic achievement standards and indicators on the Missouri Annual Performance Report (APR) for three consecutive years in 2010, 2011, and 2012. Both descriptive and inferential statistics were utilized to analyze the data and was presented in this chapter to give insight into answers to the research questions in this study. Demographic data collected was presented in Table 1. The data includes: total years of experience, years in their present district, school level, gender, highest degree earned, and number of teachers supervised.

Table 1: Demographic Characteristics of Respondents

<u>Total Years</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
3-9	45.1%	65.6%
10+	54.9%	34.4%
<u>Total Years in District</u>		
<u>Total Years in District</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
3-9	39.2%	75.0%
10+	60.1%	25.0%
<u>School Level</u>		
<u>School Level</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
Elementary	51.0%	46.9%
Middle	17.6%	21.9%
High	23.5%	9.4%
K-8	5.9%	9.4%
K-12	0.0%	12.5%
No Response	2.0%	0.0%
<u>Gender</u>		
<u>Gender</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
Male	47.1%	25.0%
Female	52.9%	75.0%
<u>Highest Degree</u>		
<u>Highest Degree</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
Bachelors	3.9%	25.0%
Masters	52.9%	50.0%
Specialist	19.6%	18.8%
Doctorate	21.6%	6.2%
No Response	2.0%	0.00%
<u>Number of Teachers Supervised</u>		
<u>Number of Teachers Supervised</u>	<u>Supervisors-High Performing (N=51)</u>	<u>Supervisors-Non-high Performing (N=32)</u>
<10	5.9%	0.0%
11-20	5.9%	34.4%
21-30	23.5%	31.3%
31-40	29.4%	9.4%
40+	29.4%	12.5%
No Response	5.9%	12.5%

In high performing districts, the number of supervisors with three-to-nine years of supervisory experience was 45.1 percent with 54.9 percent reporting more than ten years. For non-high performing schools, 65.6 percent of respondents had three-to-nine years of experience

and 34.4 percent reported more than ten years of supervisory experience. Supervisors with less than three years of experience were not considered for this study. In high performing districts, the number of supervisors with three-to-nine years of supervisory experience within their current district was 39.2 percent, with 60.1 percent reporting more than 10 years of supervisory experience. For non-high performing schools, 75 percent of respondents had three-to-nine years supervisory experience within their current district, while only 25 percent reported more than ten years of supervisory experience within their current district. For this study, only supervisors who had been in their current district three or more years were considered because the tenure for supervision was necessary to allow time for the principal to influence instruction in the building (Hart, 1993). High performing districts in this study tended to have more experienced supervisors with greater longevity in their current position than non-high schools from this study.

Supervisors from both high performing and non-high performing schools represented all levels of education from elementary through high school. The largest portion of representation was from elementary schools followed by middle and then high schools. There were four respondents from K-8 schools and four principals who had the task of evaluating teachers from grades K-12.

Among supervisors there were more female respondents with 52.9 percent for high performing schools and 75.0 percent for non-high performing schools. For high performing schools 47.1 percent were male, while only 25 percent were male at non-high performing schools.

Among supervisors of both high performing and non-high performing schools, the highest degree earned was a doctorate. A doctorate degree was earned by 21.6 percent of high performing school respondents with only 6.2 percent of respondents from non-high performing

schools reported an educational level of doctorate. Only 3.9 percent of supervisors from high performing schools reported their highest level of education as a bachelor's degree, while 25 percent of non-high performing districts held only a bachelor's degree.

The most common number of teachers a supervisor was responsible for in high performing districts was 31-40 (29.4 percent) and 40+ (29.4 percent). The most common number of teachers supervised in non-high performing districts was 11-20 (34.4 percent).

Section Two of the survey addressed the area of Teacher Supervision and Evaluation Practices, Beliefs, and Values. Section Two contained 31 items divided into a two scales, a practices scale and a beliefs scale. The items on the practices scale asked the principal to rate 18 statements (response items 1-16) based on practices the principal currently used in supervising and evaluating teachers from his/her district, such as frequency of observation, observation methods, post-observation meetings, and input gathered from the teacher. The practices scale inquired into the principal's interactions with teachers during the evaluation process. The practices scale also examined artifacts and data analyzed by principals during the evaluation process. The 18 items on the practices scale from Section Two (response items 1-16) were presented to the respondent to answer twice. The principals were asked to rate their perceptions on their present evaluation system currently used within the principal's district with one Likert rating scale, and the principals were also asked to do the same regarding their perceptions on an evaluation system they considered ideal.

Data from the practices scale of Section Two was organized into the following categories: high-present, high-ideal, non-high-present, and non-high-ideal. The high-present category contained the responses from high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation in their buildings. The high-

ideal category contained the responses from high performing schools pertaining to the principals perceptions of the system of supervision and evaluation the principal considered *ideal*. The non-high-present category contained the responses from non-high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation in their current building. The non-high-ideal category contained the responses from non-high performing schools pertaining to the principal’s perceptions of the system of supervision and evaluation the principal considered *ideal*. The mean and standard deviation for each item of each of the 4 categories (high-present, high-ideal, non-high-present, and non-high-ideal) on the practices subscales were presented in Table 2.

Table 2: Practices Subscales-Item Mean & Standard Deviations

Item	High Present				High Ideal				Non-high Present		Non-high Ideal	
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
q1	52	3.10	0.75	51	3.73	0.57	37	3.00	.85	36	3.89	0.32
q2	52	2.98	0.98	51	3.41	1.02	37	2.95	.94	37	3.51	0.93
q3	52	3.69	0.51	50	2.90	1.04	37	2.43	1.07	37	2.68	1.03
q4	50	3.30	0.93	49	3.47	0.89	36	3.31	.92	36	3.56	0.73
q5	52	3.27	0.79	51	3.71	0.58	37	3.16	.76	37	3.73	0.61
q6a	52	3.06	0.75	51	3.53	0.67	37	2.81	.97	37	3.68	0.53
q6b	51	3.39	0.72	50	3.80	0.53	37	3.32	.75	37	3.89	0.31
q7	51	2.96	1.09	50	3.32	1.02	37	2.89	1.17	37	3.35	0.98
q8	51	2.94	1.22	50	3.02	1.27	37	2.49	1.22	37	2.84	1.21
q9	51	3.12	0.95	50	3.64	0.63	37	3.03	.87	37	3.62	0.59
q10	51	3.06	0.90	50	3.46	0.81	37	2.95	.70	37	3.46	0.69
q11	51	3.10	0.73	50	3.80	0.53	37	2.73	.84	36	3.42	0.73
q12a	51	3.20	0.85	50	3.42	0.78	37	3.08	1.06	37	3.68	0.53
q12b	51	2.98	0.97	50	3.34	0.89	37	3.30	.70	37	3.65	0.48
q13	51	2.96	1.06	49	3.82	0.53	37	2.92	1.01	37	3.78	0.42
q14	50	3.32	0.82	49	3.86	0.50	37	3.19	.88	37	3.81	0.40
q15	50	3.32	0.94	49	3.63	0.73	37	3.19	.91	37	3.59	0.76
q16	51	2.69	1.16	50	3.46	0.73	37	2.08	1.09	37	3.30	0.85
Valid N (listwise)	49			45			36			34		

The means on the practices subscales for high performing school supervisors-present ranged from 2.69 to 4.00 indicating variation in levels of agreements from disagree to strongly agree. The means for high performing school supervisors-ideal ranged from 2.90 to 3.86 indicating a range of levels of agreement from agree to strongly agree. Levels of agreement differed more for supervisors in non-high performing districts with levels of agreement from 2.08 to 3.32 for the present system and 2.68 to 3.89 for the ideal system. Levels of agreement with the survey items reflected the extent supervisors perceived the presence of practices such as planning the observation, frequency of observation, observation methods, and teacher participation in their present system of teacher supervision and evaluations compared to the systems they considered ideal. The general pattern of these responses suggested supervisors in high performing districts perceived less variation between their present system of supervision and evaluation and the system they considered ideal than those perceptions of principals from non-high performing districts.

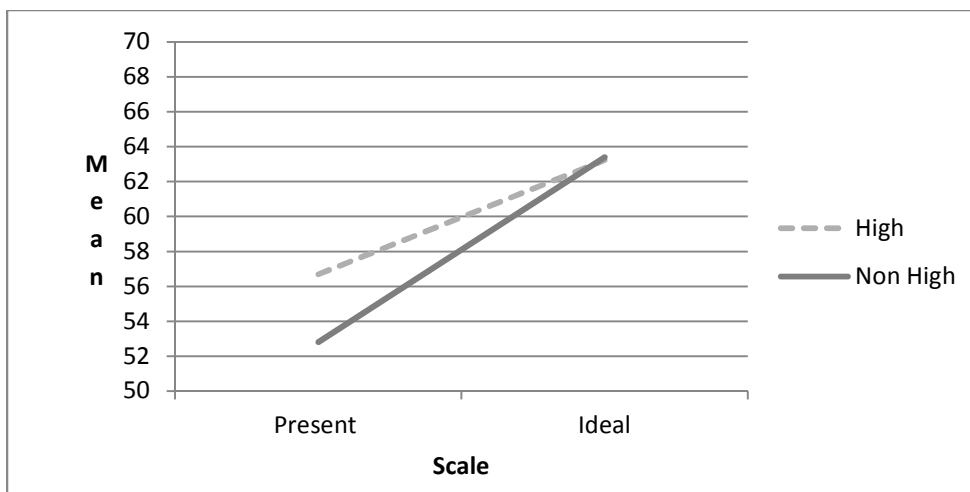
Reliability of each of the practices subscales was calculated using Cronbach's alpha coefficient. The alpha coefficient for the high performing school supervisors-present was 0.856. The alpha coefficient for the high performing school supervisors-ideal was 0.907. The alpha coefficient for the non-high performing school supervisors-present was 0.770. The alpha coefficient for the non-high performing school supervisors-ideal was 0.769. These four alpha levels reflect a high degree of internal consistency for each of the practices subscales.

The mean and standard deviation of the total subscale for each of the practices subscales was displayed in Table 3 and graphically presented in Figure 1. The means of the high performing school subscale was higher than the non-high performing school subscale on the present supervision system.

Table 3: Practices Subscales-Subscale Means & Standard Deviations

Group	N	Present Mean	Present Variance	Present SD	N	Ideal Mean	Ideal Variance	Ideal SD
High	52	56.38	77.19	8.78	51	63.23	74.87	8.65
Non	37	52.80	56.87	7.54	37	63.40	35.74	5.98

Figure 1: Practices Subscales-Means of Subscales



The mean of the high-present scale indicated supervisors in high performing districts in Missouri were favorable to current teacher supervision practices. The 56.38 mean translates to an overall 'agree' on the likert-type scale for high performing schools. According to the mean on the non-high-present scale, it appeared supervisors of teachers in non-high districts were slightly less favorable to the current teacher supervision practices in their buildings with a mean of 52.80. The means on the ideal scale for both groups of schools, high and non-high, were closer together (63.23 for high and 63.40 for non-high), but both types of schools' means showed higher favorability toward the ideal teacher supervision on the practices scale. The standard deviations on both present teacher evaluation scales were fairly close together, but the standard deviations on the ideal teacher

evaluation scales were more widely distributed which indicated more varied responses and stronger perceptions. The standard deviation on the ideal scales for both high performing and non-high performing schools indicated there was a wider diversity of opinion of what an ideal teacher supervision system means from principals within the high performing group and principals within the non-high performing group. Figure 1 showed a comparison of high and non-high schools on the practices scale for present and ideal. On the practices scale, the mean of present perceptions and the mean of ideal perceptions for high performing schools was closer together (present was 56.38 and ideal was 63.23) than the present and ideal means for non high performing schools (present was 52.80 and ideal was 63.40). This was an indication of more alignment of practice scale perceptions for present and ideal supervision systems for high performing schools than non-high performing schools. This suggested more fidelity in implementation for high performing district administrators.

The second scale on Section Two of the survey contained 13 items which asked the principal to rate 13 statements (response items 17-19) based on what the principal believed the evaluation system was intended to do and if the supervision and evaluation system reflected district beliefs and values. These statements were labeled the beliefs scale. The beliefs scale items were designed to gather perceptions about the purpose of supervision and evaluation, determine if the supervision system reflected district beliefs about teaching, and gather perceptions of the administrator about the level of trust in the relationship with the teacher being evaluated. The 13 items on the beliefs scale in Section Two (response items 1-16) were presented to the respondent to answer twice. The principals were asked to rate their perceptions on their present evaluation system currently used within the principal's district with one Likert rating scale, and the principals were then asked to do the same regarding their perceptions on an evaluation system they considered ideal. Data from the beliefs scale of Section Two was organized into the following categories: high-present, high-ideal, non-high-present, and non-

high-ideal. The high-present category contained the responses from high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation. The high-ideal category contained the responses from high performing schools pertaining to the principals perceptions of the system of supervision and evaluation the principal considered *ideal*. The non-high-present category contained the responses from non-high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation. The non-high-ideal category contained the responses from non-high performing schools pertaining to the principals perceptions of the system of supervision and evaluation the principal considered *ideal*. The mean and standard deviation for each item of each of the 4 categories (high-present, high-ideal, non-high-present, and non-high-ideal) on the beliefs subscales were presented in Table 4.

Table 4: Beliefs Subscales-Item Mean & Standard Deviations

Item	N	High Present Mean	High Present SD	N	High Ideal Mean	High Ideal SD	N	Non-high Present Mean	Non-high Present SD	N	Non-high Ideal Mean	Non-high Ideal SD
q17a	51	3.24	0.89	50	3.66	0.72	37	2.97	.87	36	3.78	0.48
q17b	50	3.56	0.61	49	3.84	0.37	36	3.14	.76	36	3.86	0.42
q17c	50	3.44	0.73	49	3.86	0.35	37	3.11	.84	37	3.81	0.52
q17d	50	3.56	0.58	49	3.88	0.33	37	3.14	.82	37	3.81	0.46
q17e	50	2.80	0.95	49	3.61	0.49	37	2.92	.80	37	3.65	0.54
q17f	50	2.92	0.78	49	3.76	0.52	36	2.89	.85	35	3.77	0.49
q17g	50	2.64	1.03	49	3.55	0.82	37	2.73	.87	36	3.33	0.83
q18a	49	3.04	0.89	48	3.65	0.60	36	2.67	.96	35	3.69	0.53
q18b	50	2.98	0.82	49	3.59	0.61	37	2.73	.93	37	3.54	0.56
q18c	50	2.90	0.86	49	3.65	0.56	37	2.89	.84	37	3.70	0.46
q18d	49	3.12	0.78	48	3.71	0.46	37	3.05	.85	37	3.81	0.40
q18e	50	3.06	0.87	49	3.71	0.46	37	2.89	.91	36	3.86	0.35
q19	49	3.35	0.63	48	3.90	0.31	37	3.08	.89	37	3.73	0.56
Valid N (listwise)	47			46			34			31		

The means on the present beliefs subscales for high performing school supervisors ranged from 2.64 to 3.56 indicating variation in levels of agreement from disagree to strongly agree. The ideal means for high performing school supervisors ranged from 3.55 to 3.90 indicating a range of levels of agreement from agree to strongly agree. Levels of agreement differed less for supervisors in non-high performing districts with levels of agreement from 2.67 to 3.14 for the present system. Levels of agreement for non-high performing districts on the ideal system (3.33 to 3.86) differed more than high performing districts on the ideal system. Levels of agreement with the survey items reflected the extent to which supervisors in high performing districts and non-high performing districts reported their beliefs and values about their present system of teacher supervision and evaluation and the teacher supervision and evaluations systems they considered ideal. However, when examining the differences in means for the present system and means of the ideal system for high performing schools and non-high performing schools, on questions 17a to 19, the present system means of high performing schools were closer to the ideal means than non-high performing schools, with the exception of question 17e and 17g. Items 17a-17g asked the principal questions about their beliefs on what their supervision and evaluation system was intended to do, such as; facilitate school improvement, improve teacher classroom performance, and resulted in improved student learning. Item 18 asked the principal to rate their perceptions about how the system of supervision and evaluation reflected the district beliefs of teachers being adult learners with varied experiences, teacher professional growth as a priority, and teachers learning from reflection on their own experiences. Item 19 asked the principal's perception of their relationship with the teacher throughout the supervision and evaluation process, and if the relationship was characterized by collaboration, honesty, trust, openness, and shared commitment to the teacher's professional growth. The general pattern of

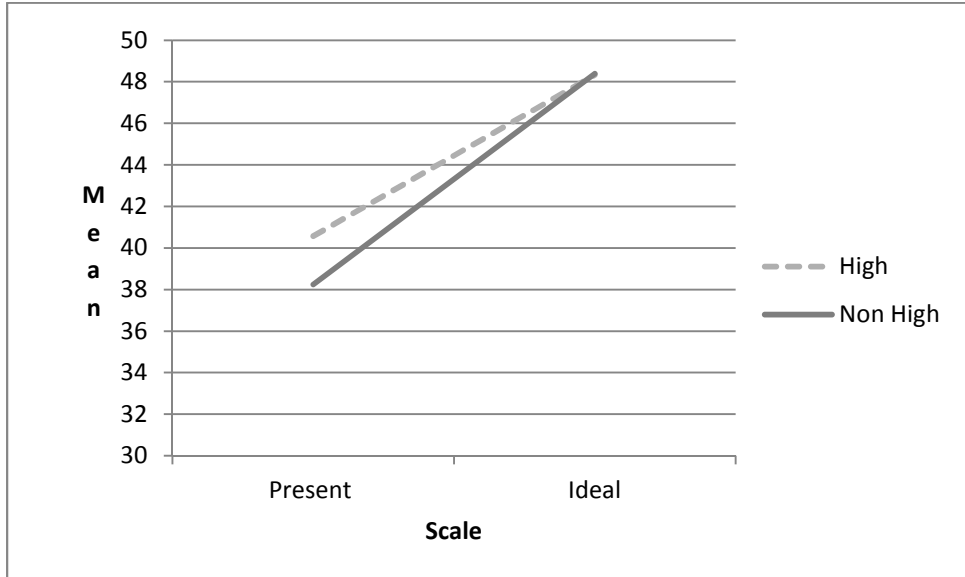
these belief subscale responses suggested supervisors in high performing districts perceived less variation between their present system of supervision and evaluation and the system they considered ideal than did the non-high performing districts. Reliability of each of the beliefs subscales was calculated using Cronbach’s alpha coefficient. The alpha coefficient for the high performing school supervisors-present was 0.929. The alpha coefficient for the high performing school supervisors-ideal was 0.909. The alpha coefficient for the non-high performing school supervisors-present was 0.938. The alpha coefficient for the non-high performing school supervisors-ideal was 0.863. These four alpha levels reflect a high degree of internal consistency for each of the beliefs subscales.

The mean and standard deviation of the total subscale for each of the beliefs subscales were displayed in Table 5 and graphically presented in Figure 2. The mean of the high performing school subscale was higher than the non-high performing school beliefs subscale on the present supervision system, but ideal subscales for beliefs was very similar for both high performing and non-high performing schools further noting principals in both high performing districts and non-high performing districts shared similar views of what the supervision and evaluation system was intended to do.

Table 5: Beliefs Subscales-Subscale Means & Standard Deviations

Group	N	Present Mean	Present Variance	Present SD	N	Ideal Mean	Ideal Variance	Ideal SD
High	50	40.57	59.83	7.74	49	48.33	22.21	4.71
Non	37	38.24	70.63	8.40	37	48.39	16.93	4.11

Figure 2: Beliefs Subscales-Means of Subscales



The mean of the high-present scale indicated supervisors in high performing districts in Missouri agreed with the beliefs and values presented in the survey, such as the supervision and evaluation system was intended to facilitate school improvement, improve teacher classroom performance, and resulted in improved student learning. The 40.57 mean translated to an overall ‘agree’ on the Likert-type scale. According to the mean on the non-high-present scale as well, it appeared supervisors of teachers in non-high schools were slightly less favorable to teacher supervision practices with a mean of 38.24. Both groups of schools, high and non-high, had ideal means that were closer together, but both higher than the present scale. The standard deviation on both present scales were fairly close together, as was the standard deviation on both ideal scales indicating responses were not widely distributed. This indicated supervisors in high and non-high schools in the sample population were in relative agreement and held similar beliefs regarding the statements under the beliefs scales.

Section Three (Effectiveness) of the survey asked administrators from high performing schools and non-high performing schools to rate their perceptions on the effectiveness of their present evaluation system. Section Three contained seven items (response items 20a-20g) which asked the principal to rate effectiveness characteristics of their current system of supervision and evaluation over the last two years. Items on the effectiveness scale asked the principal to rate their present evaluation systems' opportunity toward school improvement efforts, improved teacher classroom performance, and teachers' ability to self-reflect about their teaching were part of the effectiveness scale. Since the seven response items on Section Three pertained specifically to the present system of supervision and evaluation during the past two years, only one Likert score was necessary. Data from the effectiveness scale of Section Three was organized into the following categories: high-present and non-high-present. The high-present category contained the responses from high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation in their buildings. The non-high-present category contained the responses from non-high performing schools pertaining to the principals perceptions of the *present* system of supervision and evaluation. The mean and standard deviation for each item of high-present and non-high-present on the effectiveness subscales were presented in Table 6.

Table 6: Effectiveness Subscales-Item Mean & Standard Deviations

Item	N	High Present Mean	High Present SD	N	Non-high Present Mean	Non-high Present SD
q20a	51	2.86	0.63	37	2.81	.78
q20b	51	3.02	0.55	36	3.06	.67
q20c	51	3.06	0.54	37	2.97	.64
q20d	51	3.12	0.55	37	3.00	.58
q20e	51	2.69	0.62	37	2.89	.77
q20f	51	2.96	0.75	37	2.92	.76
q20g	51	2.57	0.73	36	2.75	.77
Valid N (listwise)	51			35		

Responses for these seven items reflected perceptions for the effectiveness of the respondent's current teacher supervision system. Responses on the effectiveness scale for high performing schools ranged from 2.57 to 3.12 and indicated a range of levels of response from disagree to agree. Responses on the effectiveness scale for non-high performing schools ranged from 2.75 to 3.00 and indicated a range of levels of response from disagree to agree. These scores indicated principals from both high performing and non-high performing districts gave responses which only slightly agreed their present system of supervision and evaluation was effective.

Reliability of each of the effectiveness subscales was calculated using Cronbach's alpha coefficient. The alpha coefficient for the high performing school supervisors-present was 0.765. The alpha coefficient for the non-high performing school supervisors-present was 0.874. These two alpha levels reflect a high degree of internal consistency for each of the effectiveness subscales.

Table 7: Effectiveness Subscales-Subscale Means & Standard Deviations

Group	N	Present Mean	Present Variance	Present SD
High	51	20.27	8.04	2.84
Non	37	20.39	13.93	3.73

The mean and standard deviation of the total subscale for each of the effectiveness subscales were displayed in Table 7. The means of high and non-high school supervisors on the effectiveness scale would indicate supervisors in both school levels agree their present systems of evaluation were effective. The 20.27 (high) and 20.39 (non-high) means translated to a slight “agree” on the Likert-type scale. Means for both high performing and non-high performing supervisory responses on the effectiveness subscales were very similar. However, when the variance and standard deviation of the responses were examined, supervisors in high performing districts had less variation between their perceptions of the effectiveness of their present system than non-high performing districts. This data suggested while there was only slight overall agreement about effectiveness of the supervision system, there was disagreement within the group of principals in non-high performing districts about the effectiveness of their present supervision and evaluation system.

Inferential Statistics.

The inferential statistics to follow were computed using the questions from Section Two and Section Three of the survey. First, the results were evaluated using a series of *t*-tests for non-independent samples determine differences between perceptions regarding the present teacher supervision system and a teacher supervision system the respondent considered ideal. Table 8

and Table 9 compared the practices scale. The first *t*-test for non-independent samples was used for high performing schools on the practices subscale to compare the principal's perceptions regarding the present teacher supervision system and a teacher supervision system the respondent considered ideal (Table 8). Table 9 was a *t*-test for non-independent samples for non-high performing schools on the practices subscale to compare the principal's perceptions regarding the present teacher supervision system and a teacher supervision system the respondent considered ideal.

Table 8: *t*-test results for High Performing Schools Practices Scale and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
Practices Scale	58.284	9.16	.903	56.49	60.07	64.57	102	.000*

Note. *= $p < .001$.

Table 9: *t*-test results for Non-high Performing Schools Practices Scale and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
Practices Scale	56.60	8.31	.966	54.68	58.53	44.83	73	.000*

Note. *= $p < .001$.

Table 10 and Table 11 compared the beliefs scale. The first *t*-test for non-independent samples was used for high performing schools on the beliefs subscale to compare the perceptions of principals regarding the present teacher supervision system and a teacher supervision system the principal considered ideal (Table 10). Table 11 was a *t*-test for non-independent samples for

non-high performing schools on the beliefs subscale to compare the principals' perceptions regarding the present teacher supervision system and a teacher supervision system the respondent considered ideal.

Table 10: *t*-test results for High Performing Schools Beliefs Scale and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
Beliefs Scale	42.91	7.24	.727	41.47	44.36	59.01	98	.000*

Note. $*=p<.001$.

Table 11: *t*-test results for Non-high Performing Schools Beliefs Scale and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
Beliefs Scale	41.82	8.02	.93	39.96	43.67	44.83	73	.000*

Note. $*=p<.001$.

For the practices scale, significant differences were found between the perceptions of high performing district supervisors on their present and ideal systems of teacher supervision (Table 8). Likewise, for the beliefs scale, significant differences were found between the perceptions of high performing district supervisors on their present and ideal systems of teacher supervision (Table 10). When examining the *t*-test for non-independent samples for non-high performing schools, significant differences were found between the perceptions of supervisors on their present and ideal systems of teacher supervision (Table 9) on the practices scale. For the beliefs scale, significant differences were also found between perceptions of supervisors on their present and ideal systems of teacher supervision (Table 11). This information indicated

differences existed on both the practices and beliefs scales consistently between the present and ideal evaluation system for both high performing and non-high performing schools. The significant differences on the practices scale suggested both groups of principals did not implement their current evaluation (present) with fidelity (ideal system). The significant differences on the beliefs scale suggested both groups of principals did not believe teacher supervision and evaluation (present system) accomplished the purposes it was intended and did not reflect district beliefs about teaching.

An analysis of variance (ANOVA) was used to evaluate differences for high and non-high schools on the practices subscale for the present supervision system (Table 12).

Table 12: ANOVA for Present Practices Scale and School Level (High vs. Non-high)

	Sum of Squares	Df	Mean Square	F	<i>p</i> value
Between Groups	277.58	1	277.58	4.04	.048
Within Groups	5983.63	87	68.78		
Total	6261.21	88			

Note. $*=p<.05$

For the practices scale, statistically significant differences were found between high performing and non-high performing schools (Table 12) on the respondents' present teacher supervision system. It is accepted at the 5 percent level. This is an acceptable significant level for the purposes of this study since .05 is widely known to be acceptable within the educational community. The data from this study did show a significant difference in practices of principals in high performing and non-high performing districts. These practices included assessing a

teacher's performance against a clearly articulated and written set of teacher performance standards, meeting with the teacher each year to establish goals for teacher professional growth and performance improvement, meeting prior to the observation to plan the observation, evaluating the total teacher performance with a variety of methods, and actively involving the teacher in the development of the supervision and evaluation system. In particular differences were noted in the following key areas:

- High performing district administrators reported meeting with teachers at the beginning of the evaluation process to plan the observation more often than did principals in non-high performing districts.
- High performing district administrators believed their evaluation of a teacher was a more accurate description of the teacher's performance than did non-high performing district administrators believed.
- High performing administrators were more likely to evaluate teacher's total performance each year, rather than on a three year cycle which was a process more favored by non-high performing district administrators.
- High performing district administrators also favored a rating scale rather than a narrative style evaluation. However, there was clear evidence the administrators preferred clear, descriptive rating scales with concrete examples.
- High performing administrators were more likely actively involve teachers in the development of the teacher supervision system.
- High performing district principals were more likely to have relationship with the teacher characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher's professional growth.

Table 13: ANOVA for Present Beliefs and School Level (High vs. Non-high)

	Sum of Squares	df	Mean Square	F	<i>p</i> value
Between Groups	114.85	1	114.85	1.78	.185
Within Groups	5474.45	85	64.41		
Total	5589.30	86			

Note. *P*=n.s.

For the beliefs scale, there were no significant differences found between high and non-high schools (Table 13). This lack of significant difference suggests administrators in both high performing and non-high performing schools had similar perceptions towards the beliefs and values of their present teacher supervision system. When individual beliefs items were analyzed, however, one item (item 19) did indicate a significant difference between perceptions of high performing district principals and non-high performing district principals.

Table 14: ANOVA for Item 19 (High/Non-high)

	Sum of Squares	df	Mean Square	F	<i>p</i> value
Between Groups	.58	1	.58	3.03	.085*
Within Groups	15.78	83	.19		
Total	16.35	84			

Note. *=*p*<.1.

An analysis of variance (ANOVA) was used to evaluate differences for high and non-high schools on item 19 “Throughout the supervision/evaluation process, my relationship with the teacher is characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher’s professional growth” (Table 14). While not highly significant since it was accepted at the 10 percent level, this was an acceptable significant level for the purposes of this study since .1 is widely known to be acceptable within the educational arena. This data indicated a difference between high performing district principals and non-high performing district principals’ perceptions about their relationship with the teacher. High performing district principals were more likely to have relationship with the teacher characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher’s professional growth.

Table 15: ANOVA for Present Effectiveness and School Level (High vs. Non-high)

	Sum of Squares	df	Mean Square	F	<i>p</i> value
Between Groups	.30	1	.30	.03	.867
Within Groups	903.72	86	10.51		
Total	904.02	87			

Note. *P*=n.s.

For the effectiveness scale, there was not a significant difference found between high and non-high schools (Table 15). This lack of significant difference suggested administrators in both high performing and non-high performing schools had perceptions similar to each other towards

the effectiveness of their present teacher supervision system. Data from Table 6 and Table 7 indicated principals from both high performing and non-high performing schools only slightly agreed their supervision system was effective.

Respondents were asked to rate the following statement with regard to professional development: I received training in how to supervise/evaluate effectively in my school district’s teacher supervision/evaluation system. Results of this item were displayed in Table 16. High performing district respondents reported a mean score of 2.96 on their present supervision system, while they reported a score of 3.82 for a supervision system they considered ideal. Non-high performing district respondents reported similar means for present (2.92) and ideal (3.78) systems. This indicated a difference in the level principals reported receiving professional development (present system) and the level of professional development they felt they should have received (ideal). This data suggested more professional development on teacher supervision and evaluation was desired by both groups of principals.

Table 16: Practices Subscales-Item Mean & Standard Deviations

Item	High Present				Non-high Present							
	N	Mean	SD	N	Mean	SD	N	Mean	SD			
q13	51	2.96	1.06	49	3.82	0.53	37	2.92	1.01	37	3.78	0.42

A *t*-test for non-independent samples for high performing respondents on professional development found statistically significant differences between the present system and one the respondent would consider ideal (Table 17). Also, a *t*-test for non-independent samples for non-high performing respondents on professional development produced statistically significant differences between the present system and one the respondent would consider ideal (Table 18).

This suggested a fair amount of consistency of supervisors from high and non-high schools in reporting a difference in professional development provided by their current system and the professional development they thought they should receive in an ideal system. This further supports the idea that principals did not feel as if they had enough professional development in the area of teacher supervision and evaluation.

Table 17: *t*-test results for High Performing Schools Professional Development and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
q13	1.89	.84	.08	1.72	2.06	22.51	99	.000*

Note. $*=p<.001$.

Table 18: *t*-test results for Non-high Performing Schools Professional Development and System (Present/Ideal)

	Mean	Std. D.	Std. Error Mean	95% Conf. Lower	95% Conf. Upper	T	Df	Sig. (2 tailed)
q13	2.03	.77	.09	1.67	2.03	20.66	73	.000*

Note. $*=p<.001$.

An ANOVA was performed to compare the present level of professional development in high performing and non-high performing (Table 19) to determine if there was a difference.

Table 19: ANOVA for Present System Professional Development and System (High vs. Non-high)

	Sum of Squares	df	Mean Square	F	<i>P</i> value
Between Groups	.04	1	.04	.04	.85
Within Groups	92.68	86	1.08		
Total	92.72	87			

Note. *P*=n.s.

For the present system of supervision/evaluation, there was no significant difference found between high performing and non-high performing schools with respect to professional development. The difference found between professional development on the present system of supervision and evaluation and the ideal system for both groups of principals (Table 17 and Table 18) indicated both groups of principals have similar perceptions (Table 19). Both groups of principals (high and non-high) perceived present professional development needed to be improved.

Qualitative Results

This section provided qualitative evidence used to triangulate and support data gathered quantitatively. Table 20 documented responses to the open-ended inquiry about perceived strengths of high performing districts present teacher supervision system. Responses were grouped according to themes which emerged. When analyzing responses to strengths of the district supervision and evaluation system, the following themes emerged: descriptive criteria, collaborative, user friendliness, and flexibility. Descriptive criteria were perceived as a strength of the current district teacher supervision and evaluation system by 48.3 percent of the respondents who answered. Example responses were “Clear expectations” and “The criteria and

descriptors give a pretty good description of what quality teaching should look like”. Another 31 percent of principals cited collaboration as a strength of their current district teacher supervision and evaluation system. Sample responses included “It was developed in tandem with teachers, administrators and instructional coaches” and “Provides opportunities for coaching”. User friendliness made up 17.2 percent of responses to the strength of their current district teacher supervision and evaluation system. Sample responses included “Simplicity” and “It is user friendly”. Another 3.4 percent cited the flexibility of their current teacher supervision and evaluation system as a strength. An example response was “It allows the administrator a lot of flexibility to use the system how he/she feels is most beneficial”.

Table 20: Perceived Strengths of District Supervision-Evaluation System

Theme	N	%	Example Responses
Criteria/descriptive	14	48.3%	Clear expectations; Specific indicators for each performance level; The criteria and descriptors give a pretty good description of what quality teaching should look like.
Collaboration	9	31.0%	We are engaged in a Continuous Improvement Process which encourages and has resulted in significantly more administrator time in the classroom and more teacher administrator conferencing; It was developed in tandem with teachers, administrators and instructional coaches; Provides opportunities for coaching
User friendliness	5	17.2%	Simplicity; it is pretty basic; It's user friendly. We are updating our teacher eval. system.
Flexibility	1	3.4%	It allows the administrator a lot of flexibility to use the system how he/she feels is most beneficial.

Note. Twenty-three participants did not respond to this query.

Table 21 documented responses to the open-ended inquiry about perceived weaknesses of high performing districts present teacher supervision system. Responses were grouped according to themes which emerged. When analyzing responses to perceived weaknesses of the district supervision and evaluation system, the following themes emerged; unclear descriptors, lack of professional development for the principal, the scale and rating system, the system was not user friendly, concerns about validity, and the system was not collaborative. Unclear descriptors were perceived as a weakness of the current district teacher supervision and evaluation system by 48.4 percent of the respondents who answered. Example responses were “The descriptors do not

adequately reflect building goals”, “Each evaluator has a different working definition of the descriptors”, and “Focuses on teacher behavior above student learning. The criteria are a bit vague”. Another 25.8 percent cited lack of professional development of the principal or evaluator as a weakness of their current teacher supervision and evaluation system. Example responses included “Not enough training on the document”, “More training on administrators part to be effective”, and “Not enough training as to what quality teaching looks like”. Scale or rating system was declared by 9.7 percent of the respondents as a weakness of their current system of teacher supervision and evaluation. A sample response was “Vocabulary between proficient, advanced and basic. Generally it is one word and not differentiated enough to help a teacher grow; 3 point scale”. Another 6.5 percent cited not being user friendly as a weakness of the current teacher supervision and evaluation system. A sample response was “The number of hoops one must jump through to document incompetency”. Concerns about validity were cited as a weakness by 6.5 percent of respondents about their current teacher supervision and evaluation system. Example responses included “We developed our own system and I do wonder about validity” and “Too subjective”. Another 3.2 percent listed not collaborative as a weakness of their current system of teacher supervision and evaluation. A sample response was “Needs more collaboration from teachers”.

Table 21: Perceived Weaknesses of District Supervision-Evaluation System

Theme	N	%	Example Responses
Unclear Descriptors	15	48.4%	The descriptors do not adequately reflect building goals; Each evaluator has a different working definition of the descriptors; Focuses on teacher behavior above student learning. The criteria are a bit vague.
Lack of Professional Development for Principal	8	25.8%	Not enough training on the document; More training on administrators part to be effective; Not enough training as to what quality teaching looks like; With a new system in place this year, teachers are unsure of how to interpret results.
Scale/Rating System	3	9.7%	vocabulary between proficient, advanced and basic Generally it is one word and not differentiated enough to help a teacher grow; 3 point scale.
Not User Friendly	2	6.5%	The number of hoops one must jump through to document incompetency; the length of the evaluation tool.
Concerns about Validity	2	6.5%	We developed our own system and I do wonder about validity; Too subjective.
Not Collaborative	1	3.2%	Needs more collaboration from teachers.

Note. Twenty-one participants did not respond to this query.

Table 22 documented responses to the open-ended inquiry professional development in the area of teacher supervision/evaluation received by high performing district administrators. Responses were grouped according to the following themes which emerged; within the district experience, collaboration with other administrators, Network for Educator Effectiveness or

Missouri Department of Elementary and Secondary Education professional development, cognitive coaching, and no professional development. Professional development experiences received within the district was cited by 28 percent of the respondents who answered. Example responses were “On-going training”, “Training within my district and years of experience”, “District training”, and “PLC process has helped us build relationships and focus on data driven instruction, which has indirectly helped me be a better supervisor”. Another 20 percent cited collaboration with other administrators as professional development they had received on their teacher supervision and evaluation system. Example responses included “Internship with other districts using digital or online evaluation program”, “Working one on one with a superior” and “Networking”. Workshops held by Network for Educator Effectiveness from the University of Missouri was declared by 20 percent of the respondents as professional development they had received on their current system of teacher supervision and evaluation. Sample responses were “Summer workshops by NEE”, “First year it is intense, then we re-certify every year”, “The training attended which is put on through the University of Missouri has greatly helped”, and “NEE training - awesome”. Another 12 percent cited workshops held by the Missouri Department of Elementary and Secondary Education as professional development they had received on their current system of teacher supervision and evaluation. Sample responses were “RPDC training this year has been beneficial to better align observations to the state model” and “Going to the PBTE training this year sponsored by DESE has caused me to reflect on my practices”. Cognitive coaching training was listed as professional development for the principal on their current system of teacher supervision and evaluation by 12 percent of respondents. A sample response was “Cognitive Coaching has been helpful especially for the teacher evaluation process”. Two respondents (8 percent) replied they had received no

professional development in the area of their current system for teacher supervision and evaluation.

Table 22: Professional Development

Theme	N	%	Example Responses
Within the district- experience	7	28%	On-going training; Training within my district and years of experience; District training; a powerpoint on how to implement; PLC process has helped us build relationships and focus on data driven instruction, which has indirectly helped me be a better supervisor. 21 years in US Army, constant focus on leadership and common sense. Two masters, specialist and doctorate, RPDC Training, district PD, Mayor of city 8 years while teaching, currently on city council, Vietnam, Desert Storm, raising two great kids (one a high school teacher), building relationships, really listening to teachers and working with them focus on student learning, Exemplary DESE PLC school 3 years, PBS school-Gold award; Professor Jerry Valentine and Professor Todd Whitaker, years ago.
Collaboration with other administrators	5	20%	Internship with other districts using digital or online evaluation program; Working one on one with a superior; networking;
Network for Educator Effectiveness – University of Missouri	5	20%	Summer workshops by NEE; NEE training. First year it is intense, then we re-certify every year; NEE training; The training attended which is put on through the University of Missouri has greatly helped; NEE training - awesome!

MO Dept. of Elementary and Secondary Education	3	12%	RPDC training this year has been beneficial to better align observations to the state model; Going to the PBTE training this year sponsored by DESE has caused me to reflect on my practices.
Cognitive Coaching	3	12%	coaching, listening to others on look fors learning to calibrate the evaluation tool when you have multiple evaluators in the same building or district; Cognitive Coaching has been helpful especially for the teacher evaluation process.
None	2	8%	None

Note. Twenty-seven participants did not respond to this query.

Interview Data.

For this study, to further support quantitative data and to delve more in depth into the perceptions high performing district principals, five principals from high performing districts were interviewed regarding their teacher supervision and evaluation system and their professional development experiences. The interview data from principals was coded by question and examined for themes. The researcher then coded by topic and conducted a formal analysis (Bogdan & Biklen, 2007). The coded topics were examined to identify common practices and perceptions of the five principals about teacher supervision and evaluation practices in their district and the professional development they had received in the area of teacher supervision and evaluation.

Two themes emerged from the interview data collected from the five principals from high performing districts: practices of principals and their beliefs about their present system of

evaluation and the professional development of the principal in the area of teacher supervision and evaluation. These themes are presented in depth.

Practices and Beliefs. The quantitative data from this study indicated a significant difference in practices of principals in high performing and non-high performing districts. Qualitative data from interviews with five high performing principals in this study verified these practices were present in their systems of supervision and evaluation. This strengthens the quantitative results. The practices that emerged from the interviews were; using a clearly articulated set of performance standards, frequency of observation and frequency of formal evaluation (every year vs. every three years), using a variety of information to evaluate each teacher, considering the teacher's input and involving teachers in the creation of the evaluation process.

The evaluation system used by principals interviewed in this study had clearly articulated performance standards. The principals indicated having clear descriptors was an integral part of the evaluation process. "The model uses a four-point descriptive scoring guide (which is a strength" P(1) [P=Principal Number]. "Descriptors, we use eight on each [standard], and they are tied to each category. Having the descriptors rather than just a scale to check is really helpful" (P3). "One of the biggest advantages, I think, is that it gives reference points or descriptors for the evaluators to look for when doing an evaluation" P(5). Principals interviewed reported the descriptors were adopted as part of the new Missouri teacher evaluation model. This data supported the quantitative finding of high performing district administrators favored a rating scale rather than a narrative style evaluation. However, there was clear evidence the administrators preferred clear, descriptive rating scales with concrete examples.

The principals who were interviewed reported frequent observations leading up to a summative evaluation of teachers. The principals reported a range of four-to-eight mini observations or walk throughs during the school year for each teacher. While stating teacher observations were a priority, one principal stated “time, or lack thereof, is always a factor and we feel rushed to get our evaluations completed” (P1). “All teachers have a minimum of four walkthroughs (four for tenured, six for non-tenured teachers). This is a minimum, and can be increased at administrator discretion. Teachers are then evaluated using a summative evaluation form” (P2). “Each teacher has between four-to-six (or more if needed) short walk throughs each year and one summative evaluation in February or early March” P(3). “Four-to-six observations with at least two follow-up conferences” (P4). “Six-to-eight observations where data is tracked along with very informal drop-in visits should be more than enough” (P5). The principals also reported yearly summative evaluations supporting the quantitative finding of high performing administrators were more likely to evaluate a teacher’s total performance each year, rather than on a three year cycle which was a process more favored by non-high performing district administrators.

Qualitative data gathered from the interviews with five high performing school principals about their evaluation practices revealed the principals accessed a variety of information to make evaluative decisions about teacher performance. “Our evaluation system looks at categories; instruction and assessment (feedback, clear learning expectations, essential learning goals, student engagement); professional responsibility (professionalism, reliability, flexibility/adaptability, professional growth, reflection, policies/procedures, collaboration, parent/community relationships); planning and preparation (instructional planning/preparation, integration of technology, building collaboration, content knowledge, differentiated instruction);

learning environment (student rapport, classroom design, learning time, classroom procedures, student responsibility, student behavior)” P(1). Our teacher evaluation system includes learning environment, instruction and assessment, planning and preparation, and also professional responsibilities, all with descriptors” (P2). “I believe that lesson preparation, classroom management, and student achievement data are most commonly used pieces of information. Staff relations and communication are also vital parts even though they may not be observed in a classroom situation” (P3). “We try to focus on three-to-four areas per year and make those areas the focus of our walkthroughs. This past year we looked at differentiation, student engagement, and time management” (P4). “We evaluate teachers using the Network for Educator Effectiveness (NEE) program developed by the University of Missouri. In addition to classroom observations, there is a unit of instruction component, a teacher professional development piece aligned to the evaluation data, and a student survey can be used to gather data about the teacher’s performance.” (P5). The interview responses provided information consistent with quantitative data regarding the practice of using a variety of information to evaluate the teacher’s total performance favored by high performing school principals.

The principals from high performing districts also reported a high degree of collaboration with the teacher during the supervision and evaluation process and actively involving teachers in the development of the supervision and evaluation system used. “This year we piloted the system with a few teachers in all buildings. The teachers' responses were very positive” P(1). “Our system was created by a team of teachers, administrators and central office reps. This was definitely a strength” P(2). “A district committee was formed to look at sample evaluation tools and current research” P(4). The principals also reported yearly summative evaluations

supporting the quantitative finding of high performing administrators were more likely actively involve teachers in the development of the teacher supervision system.

This interview data was consistent with quantitative results from this study which indicated high performing district administrators were more likely to have positive, trusting relationships with teachers and view teacher evaluation as a way to improve teaching rather than dismiss teachers. “The primary purpose of teacher evaluation should be a means of helping the teacher” P(1). “I think the evaluation process serves two purposes; contract renewal and developing teachers. If done correctly, reflection and growth can be achieved through the process. It is also worthwhile to have a ‘rating’ or ‘score’ to know where teachers fall in terms of effectiveness and where they can grow” P(2). “One of the biggest advantages I think is that it gives reference points for the evaluators to look for when doing an evaluation. It also gives teachers a clearer understanding of what is being evaluated, which is very important if the process is about improvement. It also develops trust” P(3). “I feel as though I can 'help' and give resources and an expert voice. Ultimately, I think cognitive coaching leads to more ownership and growth. I understand evaluation and value it...but I am more passionate about growing better teachers and not limiting them to what they see in an evaluation. They have the ability to grow and improve way beyond what I can say and do in the evaluation structure.” P(4). “Ratings can be used to identify areas more objectively where the teacher needs assistance. Staff relations and communication are also vital parts even though they may not be observed in a classroom situation” (P5). These qualitative discoveries support quantitative data key findings such as high performing districts were more likely to view the teacher supervision and evaluation system as a means to improve teachers rather than dismiss them, and high performing district administrators

were more likely to have relationships with teachers characterized by collaboration, trust, honesty, openness, and shared commitment to the teacher's professional growth.

Professional Development. Through interviews, four of the five principals from high performing districts reported recent professional development experiences based on the new Missouri teacher evaluation model. These professional development experiences had been within the past two years and, had provided similar experiences with professional development discovered in the quantitative data. Principals alluded to the importance of both collaboration with other administrators and networking as a form of professional development in the area of teacher supervision and evaluation, which supported administrator comments from the open ended section of the survey instrument. "Our admin team attended the DESE-provided professional development sessions that were offered this year. There were 4 sessions that explained the reasoning of changing the PBTE [Performance Based Teacher Evaluation] and the required criteria for the new PBTE" (P1). "I attended all 4 of the workshops that DESE offered regarding evaluation" (P2). "We have had the DESE training on each of the elements required in evaluation. The PD helped us focus, reflect and implement with fidelity" (P3). "I have been through all required trainings of NEE [Network for Educator Effectiveness], but I still struggle with aspects of the system. Our team will attend more training this summer" P(5). "We performed walk-thrus as an administrative team and calibrated our observations using the summative form" (P3). "One of the best informal PD that I have had was discussion in our conference principal meetings. Many of the principals use NEE and we are able to share insights. Our administrative team also met twice last year to make sure our evaluations were aligned to each other" (P5). One principal reported having no professional development for evaluation of teachers. "Minimal - I don't really remember any training - and we haven't had

any training for the new evaluation yet” (P4). Qualitative data supported the quantitative finding of principals do not feel as if they have had enough professional development in the area of teacher supervision and evaluation. Qualitative data suggests, however, an increased importance of providing principals with professional development through the Missouri teacher evaluation model. These results are encouraging.

Summary of Findings

This section summarized findings related to themes addressed in the survey (as organized in the presentation of the findings in Chapter 4). Finally, the data were presented regarding accepting or rejecting the hypothesis within the study. Guiding this overall aim, were the following research questions:

- Research Questions 1: What are the perceptions of supervisors of teaching regarding practices and characteristics they would consider ideal for teacher supervision?
- Research Question 2: Is there a difference between the *perceptions* of supervisors in high performing schools and *perceptions* of supervisors in non-high performing schools with respect to teacher supervision?
- Research Question 3: Is there a difference between the *practices* of supervisors in high performing schools and *practices* of supervisors in non-high performing schools with respect to teacher supervision?
- Research Question 4: Is there a difference between the *professional development for supervision practices* of supervisors in high performing schools and *professional development for supervision practices* of supervisors in non-high performing schools with respect to teacher supervision?

Perceptions Regarding Present Supervision Practices and Ideal Teacher Supervision

Practices. In reviewing survey results for items related to perceptions of present practices and ideal practices and responses on open ended survey questions, the researcher was able to support findings connected to research question one. These findings were triangulated with interview responses from principals for high performing schools.

On both the practices and beliefs scale there was relative agreement for the ideal teacher supervision/evaluation system for both high performing and non-high performing districts. This indicated that supervisors of teachers in both high and non-high performing districts share common perceptions of the ideal supervision system. The ideal system included clearly articulated and written standards for teacher performance which are used as a rubric for assessing teacher performance. A rating scale was slightly preferred over a narrative description. Ideal supervision practices, reported by principals included frequent observations and was based on a variety of observation methods, such as classroom walk throughs, cognitive coaching, units of instruction, and professional development. Perceptions of principals also noted a preference for an ideal system of evaluation be differentiated to meet the needs of teachers in various stages of their career and with differing expertise levels. Principal perceptions indicated mentors, training, and other support needed to be provided for those teachers whose performance was judged to be unsatisfactory, and more intense supervision should be provided for first year and non-tenured teachers. In the ideal teacher supervision system, principals noted that teachers need to be actively involved in the development of the evaluation system.

On the beliefs scale, the ideal system of evaluation reflected principals' perceptions about facilitating school improvement, improving teacher classroom performance, providing competent

instruction and improved learning for students, providing direction for the district professional development, and encouraging teacher self-reflection. Principals reported perceptions which recognized teachers as adult learners with varied experiences, needs, and levels of development who learn from reflecting on their own teaching experiences. Ideal supervision practices, reported by principals reported a supervisor's relationship with the teacher should be characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher's professional growth.

Open-ended comments regarding the strength of respondent's present teacher supervision system affirmed characteristics of the ideal teacher supervision system. In particular, supervisors cited as strengths the clear expectations and the collaborative nature of their present teacher supervision system. Supervisors commented that their present system had "clear expectations" and "specific indicators for each performance level". One principal cited the need for clear reference points for evaluators, and he further stated the need for these clear indicators would give teachers a clearer understanding of what is being evaluated. Supervisors also commented that their teacher supervision system "was developed in tandem with teachers, administrators and instructional coaches". Interviews with administrators from high performing districts supported this same thinking. One response supporting this was "A district committee was formed to look at sample evaluation tools and current research".

For both groups of supervisors, the present supervision system fell short in ratings of the ideal system. Supervisors of high performing and non-high performing districts differed more in their perceptions of their present supervision system. However, there were some similarities. Similarities included meeting at least once per year to establish goals for the teacher's professional growth and performance improvement and using a variety of observation methods

to gather data about the teacher's classroom performance. Both groups reported formally observing the teacher in the classroom two or more times per year, but high performing districts were more likely to evaluate the teacher's total performance once each year rather than on a three-year cycle.

Differences in Perceptions of High and Non-High Performing School Principals About Teacher Supervision. To determine if differences existed between the perceptions of supervisors in high performing schools and perceptions of supervisors in non-high performing schools with respect to teacher supervision, the items from the beliefs scale and effectiveness scale and responses on open ended survey questions were analyzed. The researcher was able to support findings connected to research question two. These findings were triangulated with interview responses from principals for high performing schools.

Based on the data, both high performing and non-high performing district supervisors reported some differences in their perceptions regarding the present teacher supervision system and a teacher supervision system the respondent considered ideal. The differences identified included: an ideal teacher supervision system should facilitate school improvement, an ideal teacher supervision system should result in improved teacher classroom performance, an ideal teacher supervision system should provide competent instruction and improved learning for students, and an ideal teacher supervision system should provide direction for the district professional development. However, in comparing perceptions on the present system, the two groups (high and non-high) had closely aligned perceptions that their present system of evaluation performed as intended in facilitating school improvement, improving teacher classroom performance, improved student learning, provided direction for professional development, and encouraged teacher self-reflection. On question 18a regarding if the

supervision evaluation system reflected teachers as adult learners with varied experiences, needs, and levels of development, high performing districts were in more agreement than non-high performing districts. Almost 80 percent of high performing districts rated this question either agreed or strongly agreed slightly over 55 percent of non-high performing districts who rated the questions similarly. Similar findings were presented on question 17b with regard to the supervision system being intended to improve teacher classroom performance. While 94 percent of supervisors from high performing districts either agreed or strongly agreed with this statement, only 77 percent of non-high performing district supervisors agreed or strongly agreed with the statement. Perhaps most telling was the response on question number 19. Question 19 asked if the relationship with the teacher was characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher's professional growth. High performing district supervisors either agreed or strongly agreed with the statement almost 92 percent of the time, while only about 75 percent of non-high performing district supervisors felt the same.

When interviews were conducted with principals of high performing districts, a theme of communication and positive relationships with the teachers emerged. Even though principals felt these aspects were not rated on teacher evaluation, they felt strongly communication and positive relationships were vital to an atmosphere of collegiality. The principals interviewed viewed the teacher evaluation process as a way to improve teachers, one even responded "a means of helping the teacher", indicating a positive and trusting relationship. One principal thought adding a component where teachers rated themselves to increase self-reflection would be helpful.

On the effectiveness scale, beliefs for the present supervision/evaluation system were similar for high performing and non-high performing district supervisors, in that both groups

believed their current system of the past two years had facilitated school improvement, improved teacher classroom performance, resulted in improved learning for students and encouraged teachers to self-reflect about their teaching. However, both groups reported less agreement on question 20g. Item 20g asked if the present system had resulted in the removal of incompetent teachers from the district during the past two year.

When examining the responses on the present effectiveness scale, both high and non-high districts either agreed or strongly agreed with responses on the effectiveness scale, leading to the conclusion both groups shared views that their present system of evaluation had been effective over the past two years at facilitating school improvement efforts, improving teacher classroom performance, resulted in competent instruction and improved student learning, provided direction for teacher professional development activities and encouraged teachers to self-reflect about their teaching. High performing district supervisors had less confidence their current supervision system had resulted in the removal of incompetent teachers from the district than supervisors from non-high performing. Only 58 percent of high performing school administrators agreed or strongly agreed that the current supervision system had resulted in the removal of incompetent teachers from the district, with 72 percent of non-high performing school administrators agreed or strongly agreed with the same. This supported the interview theme of communication and positive relationships with the teachers in high performing districts. High performing district administrators relied on the teacher supervision and evaluation process as a means to improve teachers rather than dismiss them. Much of the data from interviews supported this idea, such as this quote from a principal in a high performing school, “The primary purpose of teacher evaluation should be a means of helping the teacher”.

Differences in Practices of High and Non-High Performing School Principals with Respect to Teacher Supervision. To determine if differences existed between the practices of supervisors in high performing schools and practices of supervisors in non-high performing schools with respect to teacher supervision, the items from the practices subscale was examined for the present system of evaluation. The researcher was able to support findings connected to research question three. These findings were triangulated with interview responses from principals for high performing schools.

Practices surveyed were items such as meeting with the teacher to plan the evaluation and establish goals, using a variety of observation methods, and differentiation of supervision based on the teacher's need or career stage. The mean values on the practices scale which had the most difference between high performing and non-high performing schools were questions 3, 8, 11, and 16. Question 3 focused on the teacher and supervisor meeting prior to observation to plan the observation. On the practices scale, over 98 percent of high performing district principals reported either agreed or strongly agreed on question 3, meeting with the teacher prior to classroom observation to plan the observation. In contrast, only 54 percent of non-high performing district principals reported either agreed or strongly agreed for this question, indicating that high performing district principals were more likely to meet with teachers to plan the observation. Question 8 was to determine if the supervisor evaluated the teacher's total performance each year. Results from question 9 which asked if the teacher was evaluated at least once in 3 years, were closer in agreement for high performing and non-high performing schools than the means for question 8 (evaluation of teachers once per year) and indicated high performing schools were more likely to evaluate teachers yearly than non-high performing schools. Question 11 asked if the teacher's total performance evaluation accurately reflected

his/her performance. Over 86 percent of high performing district supervisors responded either agreed or strongly agreed, while only about 65 percent of non-high performing district supervisors responded likewise. This indicated high performing district supervisors felt as if their teacher evaluation system was a more accurate measure of teacher performance. Question 16 asked if the teacher was actively involved in the development of the supervision/evaluation system. The mean for high performing schools was higher than the mean value for non-high performing schools. This indicated teachers in high performing schools were more likely to be involved in the development of the teacher supervision/evaluation system. However, since only about 53 percent of high performing districts and 35 percent of non-high performing districts reported they agreed or strongly agreed teachers were involved in the development of the supervision/evaluation system, there seemed to be little teacher input and the process.

Qualitative data supported these findings. High performing school principals interviewed in this study reported frequent observation of teachers and yearly summative evaluations for all teachers. When further interviews were conducted with high performing district principals, one school utilized a system created by a team of teachers, administrators, and central office representatives. Another principal reported, while their district did not create their own evaluation system, a committee of teachers was involved in selecting the evaluation system used by the district. These findings supported the quantitative findings of this study.

Differences in Professional Development for Teacher Supervision and Evaluation of High and Non-High Performing School Principals. Respondents were asked to rate the following statement with regard to professional development: I received training in how to supervise/evaluate effectively in my school district's teacher supervision/evaluation system. Results of this item are displayed in Table 15. High performing district respondents reported a

mean score of 2.96 on their present supervision system while they reported a score of 3.82 for an supervision system they considered ideal. Non-high performing district respondents reported similar means for present (2.92) and ideal (3.78) systems.

In the quantitative data, both groups of principals (high and non-high) reported a difference in the amount of professional development they had received (present system) and the amount of professional development they preferred to receive (ideal system). This indicated both groups were not satisfied with the level of professional development they had received. When categorizing open ended responses describing professional development received by the supervisors, responses failed to cite a clear method shared by the majority of the districts. Within the district training, including leadership experiences made up 28 percent of the responses. Other responses were collaboration with other administrators (20 percent), the Network for Educator Effectiveness by the University of Missouri (20 percent), the Missouri Department of Elementary and Secondary Education (12 percent), and cognitive coaching (12 percent). One supervisor reported “the PLC process has helped us build relationships and focus on data driven instruction, which has indirectly helped me be a better supervisor”.

In further interviews with high performing district principals, professional development two principals reported receiving three days of training with the University of Missouri’s Network for Educator Effectiveness model, and two principals had attended the four sessions of training for Missouri’s teacher evaluation model. Only one reported not receiving any professional development for the evaluation of teachers. One principal reported collaboration with other administrative team members in evaluating teachers together to calibrate their observations. The calibration activity and the discussion produced were credited as being particularly helpful. A promising aspect of professional development was uncovered during the

interview process. A majority of high performing school principals who were interviewed reported increased emphasis on professional development of the principal with the adoption of the new Missouri teacher evaluation model.

Summary

This chapter provided information from the perspectives of supervisors of teachers in high performing districts and non-high performing districts in Missouri who participated in a survey and provided their perceptions of the teacher supervision system used in their district. The survey results were entered into the SPSS data software to run factor analyses to analyze data with the intention of answering the 5 research questions in this study. Descriptive and inferential statistics were employed to give insight into the survey results data. In analyzing and presenting the data, the overall picture of perceptions of administrators toward their teacher supervision and evaluation system. The descriptive statistics along with ANOVA values were displayed and tables to illustrate the significant differences.

Finally, recognition was given to the respondent's comments on the open ended questions and additional interview data. The responses on open ended questions were categorized and recorded in tables along with example responses. The respondents gave valuable and numerous insights into their perceptions of the strengths, weaknesses, and recommendations of their present teacher supervision system along with responses indicating that professional development on their teacher evaluation system they have received. Interview data was organized into the three themes that emerged: frequency of teacher supervision and evaluation, the goal of teacher evaluation and supervision, and professional development of the principal in the area of teacher supervision and evaluation.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

Introduction

Public and parent scrutiny and changes in the political climate and legislation such as No Child Left Behind (NCLB) and most recently, the Missouri waiver to NCLB, have increased accountability for student learning, specifically in the area of teacher supervision and evaluation. Missouri's waiver was approved in 2012 and included a new Educator Evaluation System based on Model Teacher and Leader Standards adopted in 2011.

Teacher evaluation has two main objectives. The legislative objective is to make employment decisions, while the other objective is to improve teacher effectiveness. However, practice does not necessarily match intent (Thomas, Holdaway, & Ward, 2000). Research has indicated that good teaching practices positively impact student achievement (Marzano, Frontier, & Livingston, 2011; Marshall, 2009; Clotfelter, Ladd, & Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004).

There were several reasons for this study. The body of knowledge on teacher evaluation and the impact it has on both principals and teachers as expanded as a result of this study. Much of the research in the area of teacher supervision focused on teachers' perceptions with little emphasis on the supervisors' perceptions (Colby, Bradshaw, & Joyner, 2002). This study explored whether or not administrators perceived their current teacher evaluation system to be effective in ensuring the delivery of quality instruction to students. In today's educational setting, there is increased pressure to ensure accountability for student learning. Teacher supervision and evaluation is seen as a tool to accomplish the goal.

The main goal of this dissertation was to gather administrator perceptions concerning teacher supervision to determine if the perceptions have an effect on fidelity of the teacher evaluation process leading to higher student achievement. This final chapter analysis will make comparisons between the findings of this study, the literature review, and research related to the teacher supervision and evaluation process. However, conclusions regarding the practices, beliefs, effectiveness and professional development are based upon respondent data. This chapter includes the summary of the methodology and findings. The findings were used to synthesize the outcomes of the study and to draw conclusions to answer the research questions. A key component of this chapter is a discussion of the relationships to their findings to the literature. Unexpected findings in this study are highlighted followed by conclusions. As part of the conclusions, implications for action and recommendation for further research are presented. This chapter closes with a review of the key findings from the study that pertain to the greater body of educational research.

Summary of Methods

A mixed design incorporating qualitative and quantitative methods was employed in order to investigate the perceptions of principals regarding the teacher supervision and evaluation process. First, surveys were sent to two groups of supervisors. Fifty high performing districts were randomly selected from a grouping of Missouri districts who received a “Met” on all six of the academic indicators for the Missouri School Improvement Process for each of the years 2010, 2011, and 2012 which were selected for this study. Fifty non-high performing districts were randomly selected from a grouping of Missouri districts who met the non-high performing criteria of this study. Non-high performing school districts were those who received a “Met” on less than 80% of the eighteen (6 per year) academic indicators over the three year period. The

survey also included four open-ended questions regarding strengths, weakness, and recommendations for the current supervision/evaluation system and nine questions to gather demographic information. Survey questions were presented in four sections to gather demographic information, teacher supervision/evaluation practices, supervision/evaluation beliefs and values, and effectiveness of the present teacher supervision/evaluation system. The survey asked participants to respond to each item twice, once for their present teacher supervision/evaluation system, and once for a teacher supervision/evaluation system they considered ideal. Data collected from these surveys were described and analyzed using a series of ANOVA to determine significance for the research questions. In addition, five respondents from high performing districts were selected in a sample of convenience, for additional interviews regarding their teacher supervision/evaluation system. Transcripts of the audio taped interviews and the researcher's descriptive field notes were used in a constant comparative method to code and analyze the interview data (Bogdan & Biklen, 2007).

Summary of Findings

The findings were presented related to each of the themes addressed in the survey in Chapter 4. A summary of key findings follows:

1. Principals from both high performing and non-high performing districts only slightly agreed their present system of supervision and evaluation was effective.
2. While there was only slight overall agreement about effectiveness of the supervision system, there was disagreement within the group of principals in non-high performing districts about the effectiveness of their present supervision and evaluation system.
3. Significant differences on the beliefs scale suggested both groups of principals did not believe teacher supervision and evaluation (present system) accomplished the

- purposes it was intended, did not reflect district beliefs about teaching, and did not reflect a level of trust in the relationship with the teacher being evaluated (ideal).
4. The data from this study does show a significant difference in practices of principals in high performing and non-high performing districts. These practices included the following;
 - a. High performing district administrators reported meeting with teachers at the beginning of the evaluation process to plan the observation more often than did principals in non-high performing districts.
 - b. High performing district administrators believed their evaluation of a teacher was a more accurate description of the teacher's performance than did non-high performing district administrators believed.
 - c. High performing administrators were more likely to evaluate teacher's total performance each year, rather than on a three year cycle which was a process more favored by non-high performing district administrators.
 - d. High performing district administrators also favored a rating scale rather than a narrative style evaluation. However, there was clear evidence the administrators preferred clear, descriptive rating scales with concrete examples.
 - e. High performing districts were more likely to view the teacher supervision and evaluation system as a means to improve teachers rather than dismiss them.

- f. High performing district administrators were more likely to have relationships with teachers characterized by collaboration, trust, honesty, openness, and shared commitment to the teacher's professional growth.
5. Principals reported a difference in the level professional development they had received (present system) and the level of professional development they felt they should have received (ideal). This data suggested more professional development on teacher supervision and evaluation was desired by both groups of principals.

Limitations of the Study

The accuracy of the findings in representing the perception of Missouri supervisors of instruction is dependent upon the candidness of the participants in the responses and the degree of match between the characteristics of the study sample and the target population. By guaranteeing confidential protection of responses and ensuring that only aggregated data would be reported, the researcher tried to assuage participant concerns about the disclosure of the responses or possible reprisal for their expressed opinions and to encourage participants to provide their most candid responses.

With only 52 usable surveys returned for high performing districts and 37 usable surveys returned for non-high performing school districts, the sample was not as large as anticipated. Nonetheless, the data analysis and resulting figures were convincing enough to pay conclusions and recommendations. The Chronbach's alphas produced were encouraging and provided a launching pad for a great deal of implications and recommendations. Only a small number of comparisons between factors were considered to scope of the study within the boundaries to aided by the research questions. Exploration of other factors such as rural versus suburban schools or an analysis of each question for significant differences could be done.

Interviews were conducted in a sample of convenience with five high performing district administrators. Interviews were conducted via phone due to budget and time constraints. This method led to less dynamic interviews and was a limitation of this study.

Significant Differences of the Study with Previous Research

While there were limitations to the generalizability of this investigation, results gleaned from this study permit comparisons to be made in the identification of practices of the teacher supervision and evaluation process. Work done by Colby, Bradshaw, and Joyner (2002) established foundational criteria for effective teacher evaluation. Through analysis of Likert-type survey questions, open ended survey questions, and interview data, this study expanded on these foundational criteria by supporting it was found that the principals in this study believed their present supervision system reflected these foundational criteria, including: teacher evaluation was tightly connected to district priorities and school functions such as school improvement, professional development, and student learning, educational leaders played a strong, positive role in evaluation, and the evaluation environment was supportive of ongoing, professional learning.

Blasé and Blasé (2004) defined supervision as a combination of supervisory beliefs and philosophies with the purpose of building trust, empowering teachers, and fostering reflection. Principals in this study reported their present evaluation system had encouraged teachers to self-reflect about their teaching expanding on the research of Blasé and Blasé. This study found principals from high performing schools were more likely to report a relationship with the teacher characterized by collaboration, honesty, trust, openness and a shared commitment to the teacher's professional development. This supported Blasé and Blasé's connection between supervisory actions and professional growth of teachers, teacher commitment, involvement, and

increased student learning (2004). Zepeda's (2007) research studied supervisors who connect the dots between supervision, teacher evaluation, and professional learning, assisted teachers in identification of professional development targets, provided resources, empowered teachers to build their own learning plans, and monitored results by presence in the classroom. This study expanded upon this research, noting high performing school principals were more likely to support the professional learning of teachers and view supervision as a means to improve teaching.

In a case study on instructional supervision, Glanz, Shulman, and Sullivan (2007) cited difficulty in distinguishing instructional leadership from supervision of instruction. Lezotte (2011) found that a school principal must be a well-informed student of teaching and learning in order to be an effective instructional leader, yet found few programs that were effective in developing school leaders. When examining the professional development practices of principals in this study for teacher supervision and evaluation, many of the resources reported by principals were more generalized instructional leadership methods rather than programs specifically designed for teacher supervision and evaluation. Lezotte (2011) found that developing collaborative groups was one effective strategy to develop principals into effective instructional leaders. Responses by participants of this study expanded on this idea, citing evaluation tool calibration exercises and discussion with other administrators as a viable means of professional development.

Recommendations for Future Research

Conclusions of this study offer several implications for future research. Building on this study, the next phase of research could look at the following:

1. A more in depth look at types of effective professional development practices for administrators to ensure high quality supervision and evaluation of teachers could be done.
2. A study of teacher perceptions on supervision and evaluation could be done, and a comparison made of those perceptions with administrative perceptions.
3. A case study could be conducted of supervision and evaluation practices for of non-tenured or unsatisfactory teachers observed from their first pre-observation conference to the final evaluation at the end of the year. The professional growth of the teachers should be observed in order to shed light on the effectiveness of the supervision and evaluation process on these two categories of teachers.
4. This study looked at school districts in Missouri that were high performing or not high performing. Further research could be done comparing high performing and not high performing districts disaggregated by urban or rural, average people expenditure, size, or location in the state.
5. Further research examining characteristics of the principal, such as level of education, number of years of supervisory experience, or number of years of supervisory experience in the district could also be examined for differences.
6. Finally, a longitudinal study reporting the educators' perceptions of supervision and evaluation compared to students measurable achievement outcomes and/or student ratings of the classroom instructors would be a valuable perspective to add to the current body of knowledge.

Implications for Practice

This study focused on exploring the links between teacher evaluation practices and improving teacher practices leading to higher student achievement. The findings of this study demonstrate even though supervisors in high performing and non-high performing districts in Missouri differ in current practices of teacher supervision and evaluation, they share common perceptions of an ideal supervision system. The ideal supervision system reflected recurring themes including: effective supervisors should demonstrate values of trust and cooperation and possess a belief the supervisory relationship will facilitate empowerment and reflection for teachers and supervisors; supervisors desire a supervision system with clearly articulated expectations; in an ideal system, supervisors see the process as differentiated based on individual teacher needs with emphasis on improving teachers rather than gathering evidence to dismiss teachers; and high performing school administrators emphasized a relationship with teachers of trust and communication.

Analysis of both the qualitative and quantitative data supports the need for quality professional development for the administrators in the area of supervising and evaluating teachers which was found to be lacking in the state of Missouri. Administrators often reported no training, and much variation was noted in professional development experiences of those who did report activities. Some principals reported their “life experience” or collaboration with other administrators as professional development, while others reported professional development in other leadership areas, such as Professional Learning Communities or cognitive coaching, rather than in the specific area of how to evaluate a teacher. This study indicated principals who participated in the study felt as if they had inadequate professional development to evaluate teachers. Interviews conducted with high performing school principals emphasized they valued

the training on Missouri's new teacher evaluation model and this experience provided promise there will be more importance placed on providing professional development for principals specifically in the area of supervising and evaluating teacher.

Information from this study provided valuable insight for principals, superintendents, boards of education, university administrator preparation programs and professional organization about the current practices and perceptions of principals in Missouri related to teacher evaluation systems. It also provides relevant information to State Department of Education and principal preparation programs for future training needs of principals related to effective teacher evaluation processes.

Reference List

- Adams, A. A. (2009). *Educator attitudes regarding evaluation: A study of the attitudes and opinions of southwest Missouri educators regarding the valued and outcome of the performance based teacher evaluation process*. (Doctoral dissertation). Retrieved from Proquest. (3390229)
- Adcock, E. P. (1995). Value-added effective schools study for elementary schools: 1994 Maryland school performance assessment program results. Research Report No. 36-9-95. Maryland: Prince George's Country Public Schools, Research, Evaluation and Accountability.
- Aspen Institute. (2012, January). *Means to an end: A guide to developing teacher evaluation systems that support growth and development*. Washington D. C.: Curtis, R., and Weiner, R.
- Ballard, K and Bates, A. (2008, December). Making a connection between student achievement, teacher accountability, and quality classroom instruction. Retrieved from <http://www.nova.edu/ssss/QR/QR13-4/ballard.pdf>
- Bill & Melinda Gates Foundation. (2011). Learning about teaching: Initial findings from the measures of effective teaching project. Seattle, WA.
- Bill & Melinda Gates Foundation. (2012). Gathering feedback about teaching: Combining high-quality observations student surveys and achievement gains. Seattle, WA.
- Blasé, J. & Blasé, J. (2004). *Handbook of instructional leadership: How successful principals promotes teaching and learning*. (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Bock, J. (2012, December 26). Missouri teacher preparation faces heightened standards for education. The St. Louis Post-Dispatch. Retrieved from <http://www.stltoday.com>

- Bogdan, R. C. & Biklen, S.K. (2007). *Qualitative research for education: An introduction to theories and methods*. Boston, MA: Pearson.
- Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago: Consortium on Chicago School Research: University of Chicago Urban Education Institute. Retrieved from <http://ccsr.uchicago.edu/publications/Teacher%20Eval%20Report%20FINAL.pdf>
- Bruce, C. D. & Ross, J. A. (2008). A model for increasing reform implementation and teacher efficacy: teacher peer coaching in grades 3 and 6 mathematics. *Canadian Journal of Education*, 31(2), 346-370.
- Bruggink, P. B. (2001). *Principal succession and school effectiveness: The relationship between the frequency of principal turnover Florida public schools from 1990--1991 to 1998--1999 and school performance indicators in 1998—1999*. (Doctoral dissertation). Retrieved from ProQuest. (304692467)
- Clark, P. L. (1998). *Perceptions of New Hampshire teachers and supervisors regarding teacher supervision*. (Doctoral dissertation). Retrieved from ProQuest. (9831940)
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). "How and why teacher credentials matter for student achievement." National Bureau of Economic Research Working Paper. 12828. Also available on the CALDER Center.org web site.
- Coggshall, J. G., Rasmussen, C., Colton, A., Milton, J., & Jacques, C. (2012). *Generating teaching effectiveness: The role of job-embedded professional learning in teacher evaluation*. Washington, DC: National Comprehensive Center for Teacher Quality.

- Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002(a), April) *Perceptions of teacher evaluation systems and their impact on school improvement, professional development and student achievement*. New Orleans, LA: Presented at Annual Meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. 464 916)
- Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002(b), April). *Teacher evaluation: A review of literature*. New Orleans, LA: Presented at Annual Meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. 464 915)
- Coulter, M. P. (2013). *A qualitative study of teacher and principal perceptions of Washington state teacher evaluation instruments: Danielson, marzano, and cel 5D+*. (Doctoral dissertation). Retrieved from ProQuest. (3587064)
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Fourth Ed. Los Angeles: Sage.
- Danielson, C. (2002). *Enhancing student achievement: a framework for school improvement*. Alexandria, VA: ASCD.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.
- Danielson, C. (2008). *The handbook for enhancing professional practice*. Alexandria, VA: ASCD.
- Danielson, C. (2009). *Talk about teaching! Leading professional conversations*. Thousand Oaks, CA: Corwin Press.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.

- Darling-Hammond, L. (1999). *Teaching as the learning profession: A handbook of policy and practice*. G. Sykes, (Ed.). San Francisco, CA: Jossey-Bass.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy and evidence. *Education Policy Analysis Archives*. 8(1). 171.
- Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teachers College Press.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012, March). Evaluating teacher evaluation. *Phi Delta Kappan*. 93(6). 8-15.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. Oxford, OH: Teachers College Press.
- Davis Washington, A. (2011). *Formal evaluation of teachers: An examination of the relationship between teacher performance and student achievement*. (Doctoral dissertation). Retrieved from PQDT Open. (3454838)
- DuFour, R. & Marzano, R. J. (2011). *Leaders of learning: How district, school, and classroom leaders improve student achievement*. Bloomington, IN: Solution Tree Press.
- Dunn, T. G., & Shriener, C. (1999). Deliberate practice in teaching: What do teachers do for self-improvement? *Teaching and Teacher Education*, 15(6), 631–651.
- Ebmeier, H. (2003). How supervision influences teacher efficacy and commitment: An investigation of a path model. *Journal of Curriculum and Supervision*, 18, 110-141.
- Edmonds, C. A., Waddle, J. L, Murphy, C. H., Ozturgut, O., & Caruthers, L E. (2007). Leading the learning: What Missouri principals say about their preparation programs [Electronic version]. *AASA Journal of Scholarship and Practice*, Winter 2007, 3(4), 14-21.

- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.). *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge, UK: Cambridge University Press.
- Gay, L. R., Mills, G. E., and Airisian, P. (2012). *Educational research: Competencies for analysis and applications*. 10th Edition. Boston: Pearson.
- Goe, L., Bell, C., and Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Glatthorn, A. A. (1997). *Differentiated supervision, 2nd edition*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (1998). *Supervision of instruction: A developmental approach, 4th edition*. Needham Heights, MA: Allyn & Bacon.
- Goldhammer, R. (1969). *Clinical supervision*. New York: Holt, Rinehart and Winston.
- Greene, T. (2004, July). Literature review for school-based staff developers and coaches. National Staff Development Council.
- Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205.
- Harris, B. M. (1998). Paradigms and parameters of supervision in education. In F. Gerald & E. Pajak (Eds.), *Handbook of research on school supervision* (pp. 1-34). NY: Simon and Schuster Macmillan.
- Hart, A. (1993). *Principal succession: Establishing leadership in schools*. Albany, New York: State University of New York Press.

- Hattie, J. (2012). *Visible learning for teachers maximizing impact on learning*. New York, NY: Routledge.
- Henry Barton, S. N. (2010). *Principals' perceptions of teacher evaluation practices in an urban school district*. (Doctoral dissertation). Retrieved from ProQuest. (3420551)
- Himmelein, M. E. (2009). *An investigation of principals' attitudes toward teacher evaluation processes*. (Doctoral dissertation). Retrieved from ProQuest. (3399151)
- Hopkins, G. (2005). Walk-throughs are on the move. *Education World*. Retrieved from http://www.education-world.com/a_admin/admin/admin405.shtml
- Hughes, V. M. (2006). *Teacher evaluation practices and job satisfaction*. (Doctoral dissertation). Retrieved from ProQuest. (3284776)
- Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics* 1(4): 85–108.
- Krajewski, R. J. & Anderson, R. H. (February 1980). Goldhammer's clinical supervision a decade later. *Educational Leadership*. 420-423.
- Larsen, M. A. (2009, October 26). Stressful, hectic, daunting: A critical policy study of the Ontario teacher performance appraisal system. *Canadian Journal of Educational Administration and Policy*, 8. Retrieved from http://www.umanitoba.ca/publications/cjeap/pdf_files/larsen.pdf
- Lezotte, L.W. & Snyder, K. M. (2011). *What effective schools do: Re-envisioning the correlates*. Bloomington, IN: Solution Tree Press.

- Link, T. (2011). *The impact of teacher evaluation models on teacher leadership, teacher collegiality, and instructional risk-taking*. (Doctoral dissertation). Retrieved from ProQuest. (3488091)
- Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). *Investigating the links to improved student learning: Final report of research findings*. New York: The Wallace Foundation.
- Lyon, L. K. (2009). *Teacher evaluation practices and student achievement*. (Doctoral dissertation). Retrieved from ProQuest. (3404049)
- McDonald, D. D. (2014). *The relationship between principal longevity and student achievement and middle schools in South Carolina*. (Doctoral dissertation). Retrieved from ProQuest. (3609207)
- McGreal, T. L. (1982, January). Effective teacher evaluation systems. *Educational Leadership*, 305.
- McGreal, T. L. (1983). *Successful teacher evaluation*. Alexandria, VA: Association for Supervision & Curriculum Development.
- McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.). New York: Addison Wesley Longman, Inc.
- Marshall, K., (2009). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap*. San Francisco, CA: Wiley, John & Sons.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.
- Marzano Research Laboratory. (2011). *The Marzano teacher evaluation scales*. Bloomington, IN: The Marzano Research Laboratory.

- Marzano, R. J. (2009). Setting the record straight on “high-yield” strategies. *Phi Delta Kappan*, 91(1), 30-37.
- Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Marzano, R. J. & Pickering, D. J. (2011). *The highly engaged classroom*. Bloomington, IN: Marzano Research Laboratory.
- Marzano, R. J. & Toth, M. D. (2013). *Teacher evaluation that makes a difference*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Marzano, R. J., & Waters, T. (2009). *District leadership that works: Striking the right balance*. Bloomington, IN: Solution Tree Press.
- Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation*. San Francisco: Jossey Bass.
- Mathers, C., Oliva, M., & Laine S. W. M. (2008). *Improving instruction through effective teacher evaluation: Options for states and districts*. Research and Policy Brief. Washington D. C.: National Comprehensive Center for Teacher Quality.
- Mathes, S. M., Mixon, S. A., Trutzel-Betts, T. T. (2009). *Enhancing administrators; capacity for leadership by improving evaluation practices and processes project*. (Doctoral dissertation). Retrieved from PQDT Open. (3383302)
- Missouri Department of Elementary and Secondary Education (DESE) (1999).

- Guidelines for Performance-Based Teacher Evaluation. Jefferson City, MO.\
- Missouri Department of Elementary and Secondary Education. (2011). Missouri model teacher and leader standards a resource for state dialogue, Vol. 45, No. 45. Retrieved from: <http://dese.mo.gov/eq/documents/StandardsInformationDocument.pdf>.
- Missouri Department of Elementary and Secondary Education. (2011). State Board Approves Teacher and School Leader Standards, Vol. 45, No. 45. Retrieved from: <http://dese.mo.gov/news/2011/teacherleaderstandards.htm>
- Missouri Department of Elementary and Secondary Education. (2011, April). *Educator standards and performance assessment systems powerpoint* [PDF document]. Retrieved from <http://dese.mo.gov/eq/eses/documents/ESESPowerpoint4-18.pdf>.
- Missouri Department of Elementary and Secondary Education. (July, 2013). Comprehensive guidelines for educator evaluation. Retrieved from: <http://dese.mo.gov/eq/documents/EssentialPrinciplesOverview-June2013.pdf>.
- Missouri Department of Elementary and Secondary Education. (June, 2012). ESEA waiver educator evaluation system FAQ. Retrieved from: <http://dese.mo.gov/qs/documents/qs-esea-flexibility-waiver-educator-faq.pdf>.
- National Commission on Excellence in Education (NCEE). (1983). A nation at risk: The imperative of educational reform. Washington, DC: U.S. Government Printing Office.
- National Institute for Excellence in Teaching. (2010, August). A teacher evaluation system that works. Daley, G. & Kim, L.
- The New Teacher Project. (2011). Rating a teacher observation tool: Five ways to ensure classroom observations are focused and rigorous. Retrieved from www.tntp.org/eval2.0

- Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010, September). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives* 18(23). Retrieved from <http://epaa.asu.edu/ojs/article/view/810>
- Nolan, J. F. & Hoover, L. A. (2008). *Teacher supervision & evaluation: Theory into practice*. (2nd edition). Hoboken, NJ: John Wiley & Son.
- Nye, B., Hedges, L., & Konstantopoulos, S. (2004, Fall). How large are teacher effects? *Educational Evaluation and Policy Analysis*, p. 237-257.
- Ohio Department of Education. (2013). Student growth measures for teachers. Retrieved from: <http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Student-Growth-Measures>.
- Owings, W. A., Kaplan, L. S., & Nunnery, J. (2005). Principal quality, ISLLC standard, and student achievement: A Virginia study. *Journal of School Leadership*, 15, 99-119.
- Pallas, A. M. (2010, December-2011, January). Measuring what matters. *Phi Delta Kappan*. 92(4). 68-71.
- Phillips, K. J. R. (2010, June). What does “Highly qualified” mean for student achievement? Evaluating the relationships between teacher quality indicators and at-risk students’ mathematics and reading achievement gains in first grade. *The Elementary School Journal*. 110(4). 464-493.
- Quinn, D. M. (2002). The impact of principal leadership behaviors on instructional practice and student engagement. *Journal of Educational Administration*. 40(5). 447-467.
- Reavis, C.A. (1978). Research in review/Clinical supervision: A review of the research. *Educational Leadership*. 580-584.

- Reeves, D. B. (2010). *Standards, assessment & accountability: Real questions from educators with real answers from Douglas B. Reeves*. Englewood, CO: The Leadership and Learning Center.
- Regional Educational Laboratory Northeast & Islands. (2012). An examination of performance-based teacher evaluation systems in five states. (No. 129). Shakeman, K.
- Sahlberg, P. (2010, September). The secret to Finland's success: educating teachers. Stanford Center for Opportunity Policy in Education.
- Sanders, W. L. (2000). Value-added assessment from student achievement data: opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329-339.
- Schmoker, M. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Sergiovanni, T. J. & Starratt, R. J. (2002). *Supervision: A redefinition* (7th ed.). Boston: McGraw-Hill.
- State Collaborative on Reforming Education. (2012). TVAAS: An introduction to value-added in Tennessee. Retrieved from: http://www.tnscore.org/wp-content/uploads/2010/12/SCORE_TVAAS4.pdf.
- Stronge, J. H. (2010). *Effective teachers = student achievement: What the research says*. Larchmont, NY: Eye on Education
- Sutton, S. R. (2008). Teachers' and administrators' perceptions of teacher evaluation. (Doctoral dissertation). Retrieved from ProQuest. (3340518)
- Thomas, D., Holdaway, E., & Ward, K. (2000). Policies and practices involved in the evaluation of school principals. *Journal of Personnel Evaluation in Education*, 14(2), 215-240.

- Toch, T., & Rothman, R. (2008, January). Rush to judgment: Teacher evaluation in public education. Washington, DC: Education Sector.
- Tucker, P. D. & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: Association for Supervision & Curriculum Development.
- United States Department of Education. (2009). Race to the top program executive summary. Retrieved from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>
- United States Department of Education. (2014). White house report: Race to the top setting the pace for gains across the education system. Retrieved from http://www.whitehouse.gov/sites/default/files/docs/settingthepacerttreport_3-2414_b.pdf
- Valentine, J. W. & Harting, R. D. (1988). Performance-based teacher evaluation in Missouri: A three-year report. Columbia: University of Missouri.
- Valentine, J. W. (2005, August). The instructional practices inventory: A process for profiling student engaged learning for school improvement. Middle Level Leadership Center. Retrieved from http://education.missouri.edu/orgs/mlc/4A_ipi_overview.php
- Wang, M. C., Haertel, G., & Walberg, H. J. (1993). Toward a knowledge base of school learning. *Review of Educational Research*, 73(3), 249-294.
- Watkins, S. E. (2011). Successful title 1 schools: Teacher and supervisor perceptions of the supervision and evaluation process. (Doctoral dissertation). Retrieved from ProQuest. (3487239)
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. (2nd ed.). The New Teacher Project.

- Wise, A. T., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practices*. National Institute of Education R-3139-NIE, Retrieved from <http://www.rand.org/pubs/reports/2006/R3139.pdf>
- Wolfe, P. (1987). What the seven-step lesson plan isn't! *Educational Leadership*, 44(5), 70.
- Xu, S. (2001, January 1). *Teacher evaluation and student learning: Perceptions of elementary school principals and teachers*. (Doctoral dissertation). Retrieved from Electronic Doctoral Dissertations for UMass Amherst. (AAI3012197)
- Zepeda, S. J. (2007). *Instructional supervision: Applying tools and concepts* (2nd ed.). NY: Eye on Education.
- Zepeda, S. J., & Mayers, R. S. (2000). *Supervision and staff development in the block*. NY: Eye On Education.

Appendix A: Survey Instrument

SECTION 1: DEMOGRAPHIC INFORMATION

Please respond to the following questions:

1. How many years have you been a supervisor?
2. How many years have you been a supervisor in the present district?
3. How many teachers do you presently supervise/evaluate?
4. At what grade level(s) do you supervise/evaluate?
5. Gender: Male Female
6. How many years of post secondary education have you completed?
7. What degrees do you hold?
8. Title of your position:
9. If you have received formal training in teacher evaluation/supervision, please indicate where you received your training (please check all that apply)
 - Through workshops or seminars offered within your school district
 - Through workshops or seminars offered outside your district
 - Through university/college courses
 - Other, please specify

SECTION 2: TEACHER SUPERVISION/EVALUATION PRACTICES, BELIEFS, VALUES

Please respond twice to each of the items in this section

In the left hand column circle the response which most accurately reflects your perceptions of your present supervision/evaluation system.

In the right hand column circle the response which most accurately represents your view of the ideal supervision/evaluation system.

Strongly Disagree 1	Disagree 2	Agree 3	Strongly Agree 4
<i>PRESENT</i>			<i>IDEAL</i>
1 2 3 4	1.	I assess the teacher's performance against a clearly articulated and written set of district teacher performance standards.	1 2 3 4
1 2 3 4	2.	The teacher and I meet once each year to establish goals for the teacher's professional growth and performance improvement.	1 2 3 4
1 2 3 4	3.	Prior to the classroom observation, the teacher and I meet to plan the observation	1 2 3 4
1 2 3 4	4.	I formally observe the teacher in the classroom two or more times a year.	1 2 3 4

1 2 3 4	5.	I use a variety of observation methods to gather data about the teacher's classroom performance.	1 2 3 4
1 2 3 4 1 2 3 4	6.	During the post observation conference, the teacher and I: a. analyze and interpret the data I collected during the observation. b. identify performance strengths and areas for improvement.	1 2 3 4 1 2 3 4
1 2 3 4	7.	I evaluate the teacher's total performance (planning & preparation, classroom environment, instruction, professional responsibilities) once each year.	1 2 3 4
1 2 3 4	8.	I evaluate the teacher's total performance (planning & preparation, classroom environment, instruction, professional responsibilities) at least once in 3 years.	1 2 3 4
1 2 3 4	9.	I examine a variety of information (e.g. lesson plans, teaching materials, student performance, etc.) to evaluate the teacher's total performance.	1 2 3 4
1 2 3 4	10.	The teacher's input helps to formulate the total performance evaluation.	1 2 3 4
1 2 3 4	11.	The teacher's total performance evaluation accurately reflects his/her performance.	1 2 3 4
1 2 3 4 1 2 3 4	12.	The total performance evaluation includes: a. a rating scale b. a narrative description	1 2 3 4 1 2 3 4
1 2 3 4	13.	I received training in how to supervise/evaluate effectively in my school district's teacher supervision/evaluation system.	1 2 3 4
1 2 3 4	14.	Teachers whose performance is judged to be unsatisfactory, receive assistance in the form of a mentor, training, or other support.	1 2 3 4
1 2 3 4	15.	First year and non-tenured teachers receive more intense supervision/evaluation than tenured teachers.	1 2 3 4
1 2 3 4	16	Teacher were actively involved in the development of the teacher supervision/evaluation system.	1 2 3 4
1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4	17.	The supervision evaluation system is intended to: a. facilitate school improvement efforts b. improve teacher classroom performance c. result in improved learning for students d. insure students receive competent instruction e. provide direction for the district's professional development program f. encourage teacher self-reflection about teaching g. remove incompetent teachers from the district	1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
1 2 3 4	18.	The supervision/evaluation system reflects the following district beliefs: a. teachers are adult learners with varied experiences, needs, and levels of development	1 2 3 4

1 2 3 4		b. teaching requires the ability to make judgments and decisions on complex matters	1 2 3 4
1 2 3 4		c. teachers learn from reflecting on their own teaching experiences	1 2 3 4
1 2 3 4		d. teachers and supervisors should collaborate in the supervision/evaluation process	1 2 3 4
1 2 3 4		e. teacher professional growth is a priority	1 2 3 4
1 2 3 4	19.	Throughout the supervision/evaluation process, my relationship with the teacher is characterized by collaboration, honesty, trust, openness, and a shared commitment to the teacher's professional growth.	1 2 3 4

SECTION 3: EFFECTIVENESS OF THE TEACHER SUPERVISION/EVALUATION SYSTEM

Please rate the effectiveness of your present supervision/evaluation system in achieving the results listed. Please circle the most appropriate response.

Strongly Disagree 1	Disagree 2	Agree 3	Strongly Agree 4
20.	During the past two years, the supervision/evaluation process has:		
	a. facilitated school improvement efforts		1 2 3 4
	b. improved teacher classroom performance		1 2 3 4
	c. resulted in improved learning for students		1 2 3 4
	d. insured that students receive competent instruction		1 2 3 4
	e. provided direction for teacher professional development activities		1 2 3 4
	f. encouraged teachers to self-reflect about their teaching		1 2 3 4
	g. resulted in the removal of incompetent teacher from the district		1 2 3 4

SECTION 4: GENERAL PERCEPTIONS

Please provide a response to each of the following questions:

21. What do you feel are the greatest strengths of your school district's present teacher supervision/evaluation system?
22. What do you feel are the greatest weaknesses of your school district's present teacher supervision/evaluation system?
23. What recommendations do you have for improving the district's present teacher supervision/evaluation system?
24. Describe any professional development that has helped you become a better teacher supervisor.

Appendix B: Interview Questions

1. What is your current position and explain how you evaluate teachers?
2. Do you view the teacher evaluation process more as a means of helping the teacher or rating the teacher?
3. How effective do you feel your present teacher evaluation system is?
4. What are some strengths and weaknesses of the teacher evaluation system used by your district?
5. What components of teaching should be included in the evaluation of teachers?
6. Describe what an ideal teacher evaluation model would include. How many observations and conferences? Are there other criteria that would help administrators evaluate teachers?
7. What type of professional development have you received for the evaluation of teachers? How did the professional development strengthen your evaluation procedures?
8. Do you have any final comments as we conclude this interview?

Appendix C: Consent Email

Dear Colleague,

My name is Shannon Snow and I am a doctoral student at Southwest Baptist University, and I am conducting a research study to gather information about principal and assistant principal perceptions regarding teacher supervision and evaluation practices. I am surveying principals in 100 randomly selected school districts in Missouri. Since your district was selected, I would like to ask for your participation. I realize that you are very busy; the survey should take no more than 15-20 minutes of your time to complete. The survey is completely anonymous. It will ask you for demographic information and your practices, beliefs, and values regarding your present teacher supervision/evaluation and a teacher supervision/evaluation system that you view as ideal. The survey will also ask questions regarding your view of the effectiveness of the present teacher supervision/evaluation system used at your school.

Your privacy is important; your answers will be combined with other participants and reported in aggregate form. Information reported will not indicate individual participants or school districts. There is no penalty should you choose not to participate or answer all of the questions. Your completion and submission of the survey will indicate your consent to participate and permission to use the information that you have provided in my study.

Before you make a final decision about participation, please read the following statements about how your responses will be used and how your rights as a participant will be protected:

- Participation in the study is completely voluntary. You may stop participating at any point without penalty.
- You need not answer all of the questions.
- Your answers will be kept confidential. Results will be presented to others in summary form only, without names or other identifying information.
- Your participation will take approximately 15-20 minutes. During this time you will answer questions about how you perceive teacher evaluation practices in your district.

This project has been reviewed and approved by the RRB Committee at Southwest Baptist University (328-1992). Dr. Terry Cox is the Chair of the RRB Committee. The committee believes that the research procedures adequately safeguard the subject's privacy, welfare, civil liberties and rights.

You may contact me at 417-276-9982 if you have questions or concerns about your participation. If you would like a copy of the results of this study, you may contact me via email at ssnow@stockton.k12.mo.us. Thank you for your time and consideration.

Sincerely,
Shannon Snow

THIS PROJECT HAS BEEN REVIEWED BY THE SOUTHWEST BAPTIST
UNIVERSITY RESEARCH REVIEW BOARD FOR RESEARCH AND
RESEARCH-RELATED ACTIVITIES INVOLVING HUMAN SUBJECTS
(417) 326-1659.