

**Kansas Superintendents' Perceptions of Teacher Evaluation Meeting the ESEA
Flexibility Waiver**

Lachelle L. Sigg
B.S., University of Nebraska at Kearney, 1997
Ed.S., University of Nebraska at Kearney, 2001

Submitted to the Graduate Department and Faculty of the School of Education of
Baker University in partial fulfillment of the requirements for the degree of
Doctor of Education in Educational Leadership

Susan K. Rogers, Ph.D.
Major Advisor

Russ Kokoruda, Ed.D.

Doug Sumner, Ed.D.

Date Defended: May 2, 2016

Copyright 2016 by Lachelle L. Sigg

Abstract

The purpose of this study was to survey superintendents to determine their perceptions regarding the effectiveness of their districts' evaluation system in meeting the components of Principle Three of the Kansas ESEA waiver: supporting continual improvement of instruction; meaningfully differentiating teacher performance using at least three performance levels; using multiple measures in determining teacher performance levels; including data on student growth for all students as a significant factor in determining performance levels; including other measures of professional practice; evaluating educators on a regular basis; providing clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and informing personnel decisions. The sample group included 128 superintendents serving during the 2015-2016 school year in public Pre-K through twelfth grade school districts throughout Kansas.

The results of the study indicated that superintendents perceive their districts' evaluation systems meet all of the guidelines of Principle Three of the Kansas ESEA waiver. Superintendents perceived their districts' evaluation systems effectively support continual improvement of instruction; meaningfully differentiates teacher performance using at least three performance levels; uses multiple measures in determining teacher performance levels; includes data on student growth for all students as a significant factor in determining performance levels; includes other measures of professional practice; evaluates educators on a regular basis; provides clear, timely, and useful feedback, includes feedback that identifies needs and guides professional development; and informs

personnel decisions. It was further noted that the size of the school district did not affect these perceptions.

The results of this research provides specific feedback for lawmakers and leaders in the state of Kansas in reference to the perceived level of effectiveness of teacher evaluation tools in meeting the ESEA waiver guidelines. Kansas districts' leaders can now focus their attention toward the expanded implementation and inter-rater reliability of district evaluation systems. Likewise, future research can now focus on specific perceptions held by other stakeholder groups within the school community.

Dedication

I dedicate this work to the many family members, friends, and educators who have supported me throughout my life and career. I share my most sincere thanks to my husband, Heath, and daughter, McLaine. Whether it was the many weekends and evenings that I was locked in a room writing or the nights I was attending class, I know this work took valuable time away from both of you. Heath, thank you for acting as Mr. Mom during this time. Your constant love and support for me throughout our relationship is at the heart of any success I have experienced and I cannot imagine my life without you. Not only are you an incredible husband and father, but one of the best educators I know. I am so very fortunate to have you by my side and love you dearly. McLaine, I hope that your love of books and curiosity in exploring will be a foundation for a love of learning. My hope for you is that you will have as many wonderful teachers and mentors in your life as I have had in mine. I love you to the moon and back!

This work would not have been possible without a small army of family supporting me. I dedicate this study, and all of my accomplishments, to my mom and dad, Mike and Linda O'Connor. I have seen firsthand the powerful role of parents and family and am blessed with the very best. You have modeled for my brothers and me the importance of being honest, working hard and always putting family first. The only thing greater than your expectations for us was your support and love, I thank you for all of the sacrifices you made and love you so much. I owe great thanks to my brothers and sisters-in-law, Jason, Kalee, Shawn and Kristi (and Marcia) who have all supported me in so many ways. It goes without saying that your babysitting, words of encouragement and

belief in me were an important part of making this all possible. Thank you and I love you all.

Lastly, I dedicate this work to the many educators who have so positively influenced me. Entering Hoover Middle School in Sioux City, Iowa as a “new kid,” I was fortunate to be placed in Mr. Brian Linafelter’s sixth grade classroom. Mr. Linafelter demonstrated all that anyone would hope to have in a teacher. This experience molded me in many ways and was the catalyst for my love of middle school. Likewise, Dr. Max McFarland, Director of the UNK School Psychology Department, set a foundation for me and my educational philosophy that I continue to appreciate to this day. I was fortunate to have these educators influence me during my own schooling; however, I have been fortunate and blessed to work alongside and under some of the very best educational leaders. Dr. Bill Gilhaus, Mrs. Lana Gerber, and Dr. Jessica Dain are incredible role models for me, and I thank them for their continued support, guidance, and friendship. Jessica and Lana, you exemplify all that it means to be incredible working moms and the truest of friends. Thank you for the laughter, honesty, and shoulder to lean on when I did not think I could do anymore. I have had the sincere honor of working under the leadership of Dr. Doug Sumner for much of my educational career. There are no words to express all that you have taught me and even scarcer are the words to impart my thanks. Not only have you taught and modeled what it means to be a great educational leader, but also your dedication to serving others is evident in all that you do. I thank you for believing in me, even when I did not deserve it, and most importantly for your friendship. I look forward to this continued learning, as well as the daily laughs.

Acknowledgements

There are many individuals worthy of my thanks and gratitude for their guidance and support throughout this study. My advisor, Dr. Susan Rogers, set high expectations for me and has been a constant source of guidance throughout my research and writing. I owe incredible thanks for her dedication of time in providing feedback and suggestions. Likewise, I must acknowledge Ms. Peg Waterman, who also supported my research through her expertise and feedback. Whether it was a phone call, Zoom or personal meeting, you were both so giving of your time and expertise. A great deal of thanks to Dr. Russ Kokoruda for his time in serving on my dissertation committee as I sincerely appreciated the time dedicated to providing me with valuable feedback. Lastly, my sincerest thanks goes to Dr. Doug Sumner for his willingness to serve on my dissertation committee. Your support of me extends well beyond the scope of this endeavor and I cannot begin to articulate the respect and appreciation I have for your leadership and friendship.

Table of Contents

Abstract	ii
Dedication	iv
Acknowledgements	vi
Table of Contents	vii
List of Tables	x
Chapter One: Introduction	1
Background	2
Statement of the Problem	6
Purpose of the Study	8
Significance of the Study	8
Delimitations	10
Assumptions	11
Research Questions	12
Definition of Terms	15
Overview of the Methodology	16
Organization of the Study	16
Chapter Two: Review of the Literature	18
History of Teacher Evaluation	19
Best Practices in Teacher Evaluation	23
Perceptions of Teacher Evaluation Effectiveness	32
Summary	42
Chapter Three: Methods	44

Research Design.....	44
Population and Sample	45
Sampling Procedures	45
Instrumentation	45
Measurement.....	47
Validity and Reliability.....	50
Data Collection Procedures.....	51
Data Analysis and Hypothesis Testing	52
Limitations	61
Summary	62
Chapter Four: Results	63
Descriptive Statistics.....	63
Hypothesis Testing.....	64
Summary.....	82
Chapter Five: Interpretation and Recommendations	84
Study Summary.....	84
Overview of the Problem	84
Purpose Statement and Research Questions	85
Review of the Methodology.....	86
Major Findings.....	86
Findings Related to the Literature.....	90
Conclusions.....	93
Implications for Action.....	94

Recommendations for Future Research	95
Concluding Remarks.....	96
References.....	98
Appendices.....	104
Appendix A. Teacher Evaluation Survey	105
Appendix B. Institutional Review Board (IRB) Request.....	110
Appendix C. Teacher Institutional Review Board (IRB) Approval	115
Appendix D. Teacher Evaluation Electronic Mail Message to Superintendents.	117

List of Tables

Table 1. Evaluation Systems Utilized in the State of Kansas in April of 2014	5
Table 2. District Enrollment Demographics	64
Table 3. Descriptive Statistics for the Results of the Test for H2.....	66
Table 4. Descriptive Statistics for the Results of the Test for H4.....	68
Table 5. Descriptive Statistics for the Results of the Test for H6.....	70
Table 6. Descriptive Statistics for the Results of the Test for H8.....	71
Table 7. Descriptive Statistics for the Results of the Test for H10.....	73
Table 8. Descriptive Statistics for the Results of the Test for H12.....	74
Table 9. Descriptive Statistics for the Results of the Test for H14.....	76
Table 10. Descriptive Statistics for the Results of the Test for H16.....	77
Table 11. Descriptive Statistics for the Results of the Test for H18.....	79
Table 12. Descriptive Statistics for the Results of the Test for H20.....	80
Table 13. Descriptive Statistics for the Results of the Test for H22.....	82

Chapter One

Introduction

Education in the United States is ever changing; however, some of the most impacting transformation sweeping the nation has evolved over the last decade following the enactment of the Elementary and Secondary Education Act (ESEA), most commonly known as the No Child Left Behind Act (NCLB, 2001). This groundbreaking legislation laid the foundation for some significant changes in schools, including the assessment of students in both mathematics and reading with minimum progress requirements, the provision of teacher qualification criteria or highly qualified classification, as well as school attendance options for parents of children attending underperforming schools. As noted in the 2014-2015 Kansas State Department of Education ESEA Waiver Fact Sheet, the U.S. Department of Education provided states the chance to move away from the ESEA in 2011. The state of Kansas submitted a waiver requesting flexibility from provisions of the NCLB Act and received approval for their flexibility waiver in July of 2012. Upon receiving approval of the waiver, Kansas began work in piloting a new teacher/leader evaluation system. The Kansas Educator Evaluation Protocol (KEEP), met all criterion included as part of the waiver, except a tie to student achievement. As a result, this reference was added to the KEEP system and piloted during the 2013-2014 school year. All districts were then expected to develop an evaluation tool that included all six components outlined in the waiver or adopt the KEEP as their teacher and leader evaluation system during the 2014-2015 school year.

Background

The public education system within the state of Kansas is comprised of 286 unified public school districts. During the 2014-2015 school year, approximately 520,000 students, pre-kindergarten through 12th grade were served within these districts. Public school districts range in enrollment size from approximately 70 to 51,000 students.

As NCLB was expiring, it was evident that Kansas needed to move away from the narrowly defined and fixed growth mindset of this legislation. In its place, educators agreed that a new focus was needed to ensure students were college and career ready (KSDE, 2015). As part of this expectation, four key components were addressed to support this transition and included as part of the ESEA Flexibility Waiver Request.

The first component addressed in the Kansas ESEA waiver request was the expectation that the state adopt college and career-ready standards in reading and mathematics. Likewise, it stipulated that Kansas would have in place, high-quality assessments aligned with the state's college and career ready standards (KSDE, 2014). While this component has been implemented, it has not been without its difficulties. Curriculum creation, adoption, and professional development have continued in districts throughout the state. However, the Kansas State Board of Education voted to leave the Smarter Balanced Assessment Consortium (SBAC) before the 2014 state assessments. Thus, the state worked closely with the Center of Education, Testing and Evaluation (CETE) through the University of Kansas to develop appropriate math and English language arts assessments.

The second principle of the ESEA Waiver Request was a state-developed differentiated recognition, accountability, and support system. The ESEA waiver

stipulates that the state set new annual measurable objectives (AMOs) for schools in English language arts and mathematics that are ambitious but achievable (KSDE, 2014). Schools have three options to meet the AMO, which include increasing student achievement growth, reducing the achievement gap, or decreasing the percent of non-proficient students. Information regarding performance in each of these areas is reported; however, a school must meet just one of the AMO targets to be considered making progress (KSDE, 2014). In addition to these targets, the state has committed to identifying Title I buildings that meet these expectations, as well as those that are not meeting these expectations. Appropriate supports and recognitions are implemented based on the needs and successes of these higher at-risk buildings.

The third principle addressed in the ESEA Waiver is effective instruction and leadership. More specifically, this principle requires changes to the teacher and administrator evaluation systems in all districts to meet specified standards. This waiver criterion stipulated the development and implementation of a teacher and principal evaluation and support system that

- Will be used for continual improvement of instruction;
- Meaningfully differentiate performance using at least three performance levels;
- Use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students, and other measures of professional practice;
- Evaluate teachers and principals on a regular basis;

- Provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and
- Will be used to inform personnel decisions (KSDE, 2015, p. 6)

This component is the primary focus of this study, as much change has occurred in the teacher evaluation systems in districts throughout Kansas. The Kansas State Department of Education (KSDE) immediately began its work in the creation of a statewide evaluation tool for districts to access and utilize in meeting this requirement. While the Kansas Educator Evaluation Protocol (KEEP) was offered free of charge to any district interested in using it, there were other evaluation systems available for districts to consider. As shown in Table 1, the responses provided to KSDE in April of 2014 indicate the evaluation system utilized at that time.

Table 1

Evaluation Systems Utilized in the State of Kansas in April of 2014

Evaluation System	Number of Districts
Danielson	10
Educators 4 Excellence	44
Greenbush	7
Kansas Educators Evaluation Protocol	72
Locally-Developed	45
Marzano	2
McREL's Teacher Evaluation System	87
National Association of Secondary School Principals	1
Observation and Appraisal Management System	1
Teacher Collaboration Assessment Survey	1

Note: Adapted from ESEA Flexibility Request, Kansas State Department of Education, 2015, p. 255.

“The Kansas State Department of Education (2015) believes that a high-quality educator evaluation system is critical for informing educators about performance and key to continual improvement of instruction leading to increased student learning and achievement” (p. 250).

The fourth and final principle to be addressed by the waiver is that of reducing duplication in reporting and unnecessary burden on district staff in preparing data and reports for state agencies. While this principle will continue to receive attention, the KSDE had already addressed many of the concerns related to the data gathered and burden placed on school districts in providing data for state reporting. Due to the work

that had already been done, the state was not required to do anything further to meet this principle's waiver criteria (KSDE, 2014).

While each principle of the Kansas ESEA Waiver request directly addressed the curriculum, instruction, and assessment of Kansas schools, the third principle of the waiver is the focus of this study. Although each principle is critical to the forward momentum of Kansas' public schools, none can be achieved without effective, high-quality teachers leading classrooms. As Danielson and McGreal (2000) shared, "the research on teaching has made clear, schools are paying increased attention to the importance of the teacher to student learning and the issue of teacher quality" (p. 15). Danielson (2002) further contends, "One of the most significant influences on a school's culture is its system for teacher evaluation, which must convey expectations for performance while simultaneously promoting professional learning" (p. 35). Districts have undergone changes in the last two years to ensure their teacher evaluation systems are meeting the ESEA Waiver requirements while ensuring the highest quality instruction is occurring in Kansas classrooms.

Statement of the Problem

Prior to the approval of the ESEA Flexibility Waiver, Article 90 – Evaluation of Licensed Personnel Evaluation, as found in Kansas Statute (1970), stipulated that all districts across Kansas implement a teacher evaluation system. However, not all evaluation tools met the specific requirements of Principle Three of the Kansas ESEA Flexibility Waiver. Thus, in preparation for the shift to the ESEA Flexibility Waiver, the 2013-2014 school year was considered a pilot year for implementation of an evaluation system that would be used for continual improvement of instruction, meaningfully

differentiate performance using at least three performance levels, use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students, and other measures of professional practice, evaluate teachers and principals on a regular basis, provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development, and will be used to inform personnel decisions (KSDE, 2015).

At the beginning of the 2014-2015 school year, all districts were mandated to implement appropriate evaluation tools that met all expectations outlined as part of Principle Three (KSDE, 2015). Districts went to work in reviewing their evaluation instruments to identify areas already meeting criteria outlined in the waiver, as well as areas of need. In response to this review, districts adopted new evaluation tools; some purchased these systems while others revised their current evaluation tool to enhance its effectiveness and ability to meet the guidelines of Principle Three. Another option to meet this requirement was the adoption of the Kansas Educators' Evaluation Protocol (KEEP) (KSDE, 2015). The KEEP was developed by the Kansas State Department of Education and offered an on-line teacher evaluation tool that met Principle Three requirements. In addition, the KEEP directly connected with reports and data provided to the state. While districts were not mandated to use the KEEP, it was an instrument available to all districts that provided assurance that all components of Principle Three were met. Districts implemented updated teacher evaluation systems throughout the 2014-2015 school year (KSDE, 2014); however, no research was conducted to identify the effectiveness of the various district tools in meeting the six components included as part of Principle Three of the Kansas ESEA Flexibility Waiver.

Purpose Statement

The purpose of this study was to determine superintendents' perceptions regarding the effectiveness of their districts' evaluation systems in meeting the six components of Principle Three of the ESEA waiver; supporting continual improvement of instruction; meaningfully differentiating teacher performance using at least three performance levels; using multiple measures in determining teacher performance levels; including data on student growth for all students as a significant factor in determining performance levels; including other measures of professional practice; evaluating educators on a regular basis; providing clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and informing personnel decisions. While there are six components listed as part of Principle Three, these components included eleven specific expectations, which were delineated in the survey created for the study. An additional purpose of this study was to determine whether the demographic factor of district size affected superintendents' perceptions regarding the effectiveness of their districts' evaluation systems in meeting the six components of Principle three of the ESEA waiver.

Significance of the Study

The work of Weisberg, Sexton, Mulhern, and Keeling (2009) provided further support for changes to teacher evaluation models as their research of 15,000 teachers, 1,300 administrators, and more than 80 local and state officials found current teacher evaluations to be insufficient and ineffective. With the expiration of the No Child Left Behind Act, many states throughout the nation have reviewed and revised teacher evaluation in an effort to improve their effectiveness. While much research has been

conducted into teacher and administrator perceptions of these new systems, no specific research has focused on the perceptions of superintendents throughout the US or the state of Kansas.

Much research has been conducted into the perceptions of these ever-changing systems. The work of Sheppard (2013) provided a baseline of teacher perceptions regarding their evaluation system in the state of Georgia, prior to any new regulations or standards. Conversely, many studies have looked at the perceptions held by teachers and administrators (Clark, 2014; Coulter, 2013; Derrington, 2014; Doherty, 2009; Heyde, 2013; Killian, 2010; Ruffini, Makkonen, Tejwani, & Diaz, 2014; Winslow, 2015) concerning newly adopted state and district evaluation protocols. The focuses of these studies varied but were similar in that they looked at the effectiveness of the newly adopted evaluation tool. Although Bridich (2013) researched the perceptions of Colorado teachers, administrators and policy makers, the specific focus was on their perceived role in the implementation of a new evaluation system.

While much research has been conducted into the perceptions of teachers and administrators about the implementation of a new evaluation tool, no research has yet been conducted in the state of Kansas. In addition to this hole in the research, Derrington's (2014) research targeting the perceptions of superintendents in a southeastern state appears to be an isolated body of study. Thus, the focus on superintendents' perceptions of new evaluation criteria throughout the state of Kansas is uncharted territory. Because of the important role the superintendent plays in selecting and implementing the teacher evaluation system within the school district, it is important to garner their feedback regarding their perceptions of effectiveness. The results of this

work may be helpful to lawmakers, as well as the Kansas State Department of Education (KSDE), in making future policy and process decisions regarding the ESEA Waiver requirements as they pertain to Principle Three, effective instruction and leadership. The perception held by superintendents that their evaluation tool is not meeting one of the six ESEA principle three guidelines may prompt additional discourse with policymakers regarding changes to these expectations. The KSDE might also use information of this nature to devise greater communication structures and strengthen supports and professional development provided to school districts throughout the state. Furthermore, looking for any connection that the size of the district may have in relation to the evaluations effectiveness in meeting each of the six guidelines of the ESEA waiver might also aid in future planning and policy updates. More specifically, should a link be found between the enrollment of a district and superintendent's perceptions, targeted support could be provided to these districts. In addition to these potential benefits, results of this research may assist district leaders in identifying the level of discrepancy in their district's evaluation tools effectiveness in meeting a specific ESEA guideline, as compared to other Kansas school districts, therefore allowing further discourse in finding methods to improve areas of deficit.

Delimitations

As noted by Lunenburg and Irby (2008), "delimitations are self-imposed boundaries set by the researcher on the purpose and scope of the study" (p. 134). The delimitations of this study were:

1. This study was delimited to Kansas public schools serving pre-kindergarten through 12th grade students.

2. An online survey tool was created addressing superintendents' perceptions of their teacher evaluation system in meeting each of the six components of the Kansas ESEA waiver. Superintendents were given the opportunity not to respond to any or all of the questions included in the survey. Thus, individual research question responses may not be representative of all respondents.
3. The collection of survey data was delimited to the timeframe of the collection window.

Assumptions

Lunenburg and Irby (2008) have clarified that “Assumptions are positions, premises, and propositions that are accepted as operational for purposes of the research” (p. 135). The following assumptions were made concerning this research study:

1. Kansas superintendent email addresses acquired through the Kansas State Department of Education School Directory were correct
2. Superintendents had sufficient knowledge of their district's evaluation system.
3. Superintendents were aware of and understood the provisions included as part of component three of the ESEA waiver.
4. Superintendents responding to the survey understood the items included as part of the survey tool.
5. Superintendents responded honestly and accurately to the items on the survey.
6. Superintendents provided accurate enrollment figures as verified by the official September 2015 enrollment count.
7. The interpretation of the survey results was accurate and complete.

Research Questions

This research proposes to answer the question of superintendents' perceptions of their teacher evaluation instruments' effectiveness in meeting the parameters outlined as part of Principle Three: Supporting Effective Instruction and Leadership of the ESEA Waiver. Bearing in mind the six components comprising Principle Three of the ESEA Waiver include 11 specific expectations, the specific research questions addressed in this study have been tailored to measure each of the expectations delineated in the six components. Specifically, there are two research questions to address component three and five research questions to measure component five of the ESEA Waiver. In addition to the 11 research questions focused on these six components, the extent to which superintendents' perceptions were impacted by the size of the district was also measured. Thus, the 22 research questions are as follows:

RQ1. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction?

RQ2. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction affected by the size of the district?

RQ3. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels?

RQ4. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels affected by the size of the district?

RQ5. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on student growth for all students in determining performance levels?

RQ6. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on student growth for all students in determining performance levels affected by the size of the district?

RQ7. What are superintendent's perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice?

RQ8. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice affected by the size of the district?

RQ9. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively supporting the evaluation of educators on a regular basis?

RQ10. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in evaluating educators on a regular basis affected by the size of the district?

RQ11. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback?

RQ12. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing clear feedback affected by the size of the district?

RQ13. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing timely feedback?

RQ14. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing timely feedback affected by the size of the district?

RQ15. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing useful feedback?

RQ16. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in providing useful feedback affected by the size of the district?

RQ17. What are superintendents' perceptions of the effectiveness of their districts' evaluation system in including feedback that identifies needs?

RQ18. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in including feedback that identifies needs affected by the size of the school district?

RQ19. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing feedback that guides professional development?

RQ20. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in guiding professional development affected by the size of the school district?

RQ21. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions?

RQ22. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in informing personnel decisions affected by the size of the district?

Definition of Terms

Roberts (2010) noted that this section of the dissertation “provides the definition of terms that do not have a commonly known meaning or that have the possibility of being misunderstood” (p. 139). To that end, further definition is being provided for key terms used in this study.

Elementary and Secondary Education Act. ESEA offered various grants to support K-12 and post-secondary education. “Additionally, the law provided federal grants to state educational agencies to improve the quality of elementary and secondary education” (U.S. Dept. of Ed., 2015, para. 7).

Kansas Educators Evaluation Protocol. The KSDE (2015) noted the “Kansas Educator Evaluation Protocol (KEEP), with the addition of Student Growth Measures and a final summative evaluation rating for 2014-2015, is the default State of Kansas educator evaluation system” (p. 250).

Kansas ESEA Flexibility Waiver Request. In 2011, the U.S. Department of Education provided states the chance to request relief from some of the requirements of the federal Elementary and Secondary Education Act (ESEA), also known as No Child Left Behind (NCLB). This opportunity was provided as the result of the expiration of the reauthorization of the ESEA five years earlier and was offered as a means to increase

academic achievement and quality instruction through state and local reforms (KSDE, 2014).

Overview of the Methodology

A quantitative inquiry approach was utilized in this study. More specifically a non-experimental research design using an online survey was employed. The population surveyed as part of this study included all public school superintendents throughout the state of Kansas in the fall of 2015. The survey examined superintendents' perceptions of their teacher evaluation system's effectiveness in meeting the six components of Principle Three of the ESEA flexibility waiver. A quantitative Likert-type survey scale was utilized in gathering these perceptions. The survey was distributed during the 2015-2016 school year to current superintendents via electronic mail that included a link to the survey. The data was then compiled online using SurveyMonkey. The data collected from SurveyMonkey was downloaded and imported to IBM® SPSS® Statistics Faculty Pack 23 for Windows for analysis. Statistical analysis testing used for this study included one-sample *t* tests and one-factor analyses of variance (ANOVAs) to analyze differences in multiple variables.

Organization of the Study

This study is organized into five chapters. The first chapter was comprised of an introduction, background, statement of the problem, the purpose of the study, the significance of the study, delimitations, assumptions, research questions, the definition of terms, an overview of the methodology, and organization of the study. Chapter two includes a review of the literature, which comprises the history of teacher evaluation, best practices in teacher evaluation, and a review of the literature regarding teacher

evaluation effectiveness. Provided in chapter three are the methodology, research design, population and sample, sampling procedures, instrumentation, measurement, validity and reliability, data collection procedures, data analysis and hypothesis testing and limitations. Chapter four includes descriptive statistics, hypothesis testing, and a summary. Provided in chapter five are the interpretation and recommendations, study summary, an overview of the problem, purpose statement and research questions, review of the methodology, major findings, findings related to the literature, implications for action, recommendations for future research, and concluding remarks.

Chapter Two

Review of the Literature

The purpose, process, and impact of teacher evaluation systems in school districts throughout the nation have been under great scrutiny in the last two decades (Danielson, 2000; Hull, 2009; McGuinn, 2012; Reform Support Network, 2011). Never was this focus more evident than in November of 2009 when President Barack Obama signed into law the American Recovery and Reinvestment Act of 2009 (ARRA). As noted in this historic legislation, “the ARRA provides \$4.35 billion for the Race to the Top Fund, a competitive grant program designed to encourage States that are implementing ambitious plans in developing, rewarding, and retaining effective teachers and principals” (Race to the Top Executive Summary, 2009, para. 2). As a result of this investment, states were given the opportunity to seek flexibility from the soon-to-expire Elementary and Secondary Education Act (ESEA), also known as the No Child Left Behind Act. The state of Kansas received approval of the ESEA waiver in 2011 and since that time, various updates have been made regarding the timeline and rollout of the expectations. One of the four principle areas addressed through the waiver was teacher evaluation (ESEA, 2014).

The purpose of this study was to determine superintendents’ perceptions regarding the effectiveness of their districts’ evaluation system in meeting the six components of Principle Three of the ESEA waiver. Chapter two is organized into three sections providing insight and a review of literature in each. The first section shares the history of teacher evaluation and its evolution in the United States. Section two provides a review of best practices in teacher evaluation. The third and final section delineates

previous research that was conducted into perceptions of the effectiveness of teacher evaluation systems.

History of Teacher Evaluation

Throughout U.S. history, teacher evaluation has been revised to reflect the evolving purpose assigned to public school education. Tracy (1995) found that this evolution could be organized into seven phases; the Community Accountability Phase, the Professionalization Phase, the Scientific Phase, the Human Relations Phase, the Second-Wave Scientific Phase, the Second-Wave Human Relations Phase and the Human Development Phase (p. 1). Each of these phases was marked by notable changes in the manner in which schools were organized and led, as well as the philosophy concerning the evaluation of teachers.

“From the colonial period through the early 1800s, the responsibility for supervision rested with various members of the community because of the strong American belief in local, lay control of education” (Tracy, 1995, p. 2). Community leaders were seen as the individuals responsible for the evaluation of teachers, and because of the heavy focus on the teaching of religion in schools, clergy often filled this role. A community visiting team would provide minimal oversight through on-site visits to the school building. Bolin and Panaritis (1992) shared, “Respected but poorly paid, the teacher was trusted to inculcate the values of family and church in much the way that the master craftsman of days gone by instructed apprentices in the skills and values of practicing a trade” (p. 30).

Tracy (1995) asserts that the second phase of teacher evaluation evolved into the professionalization phase as a greater focus was made on strong instruction and the

impact of a good teacher in the classroom. In addition to a greater focus on the instruction provided in the classroom, the organization of schools also evolved. It was during this time in the late 1800s that schools began to organize themselves with neighboring schools, thus leading to the creation of local school districts. “This type of school organization quickly became the prevailing model, leading to the creation of a hierarchical system for overseeing instruction” (Tracy, 1995, p. 4). Tracy further found that as school systems became more highly organized, a need for a good instructor was acknowledged. The acknowledgement of this need resulted in a shift in evaluators from individuals within the community to professionals within the hierarchical school system.

The twentieth century also provided a time of great change and growth in the area of school organization and supervision. Tracy (1995) coined this third period as the scientific phase as there was a shift to a more scientific approach to the supervision and evaluation of instruction provided in the classroom, and to the teachers responsible for providing it. The first use of formalized checklists by administrators in the observation of teachers was one example of this more structured and scientific approach. Tracy (1995) further stated that the methodology for applying the scientific process to teacher evaluation was not done to micro-manage teachers; rather, its intent was to develop a more cohesive and structured foundation on which to judge instruction. A natural assumption derived from this philosophy was that supervisors were the experts who could provide the greatest support and insight in supporting a teacher’s growth.

The Great Depression of the 1930s and 1940s also influenced education, more specifically the manner in which teachers were evaluated and supervised. While the early 1900s was marked by a scientific approach, Tracy (1995) found that the next phase of

teacher supervision could best be described as the human relations phase. Due to the impact the Depression had on the financial and emotional needs of all, educators soon noted the differing levels of needs among their students. Likewise, Tracy (1995) found that an even greater focus was made on the well-being of the teacher and the teachers' feelings of support in their role. During this phase of teacher evaluation and supervision, Tracy (1995) notes that teachers were thought to be more productive and able to do their jobs successfully if they had a voice in the management of the school.

While Tracy (1995) had identified distinct date ranges for phases one through four, the fifth phase, a second-wave of the scientific phase, was less bound by a time frame. This phase continued and expanded the focus and application of the scientific process on instruction and supervision. Throughout the late 1940s and through the early 1960s, teacher supervision took on a greater scientific focus. As Tracy (1995) notes, "Although the primary classroom observers remained the supervisors and principals, in a few instances in this second wave, teachers analyzed their classroom data through the use of taped recordings" (p. 4). As Danielson and McGreal (2000) noted, at this time it was believed that certain variables, specifically teacher's tone of voice, appearance, and warmth were key indicators of effectiveness. Thus, some evaluation systems included these items as part of their feedback and ratings. This era has continued to influence our practices today, as much of the scientific model is acknowledged and referenced in many evaluation and supervision processes.

"The mixture of supervisory principles in the first five decades of the twentieth century created a climate for direct teacher involvement in the 1960s" (Tracy, 1995, p. 4). A more cooperative spirit among evaluators and teachers marked the 1960s and 1970s;

however, the primary focus was on supporting new teachers and providing on-going training for existing educators. Danielson and McGreal (2000) also noted, “Not until the 1960s, however, did any coherent focus on teacher appraisal begin to emerge” (p. 15). During this resurgence of the human relations phase as noted by Tracy (1995), a spotlight was again shown on the importance of the relationship between the administrator and teacher; however, there was also a higher expectation held for the administrator’s knowledge and use of data to support good instruction. Danielson and McGreal (2000) shared that throughout this time in history, there was a strong push for further research on teaching and specific methods for recording what was happening in the classroom. Likewise, Danielson and McGreal (2000) identified that “It was important work in that it began to form the basis for a set of fundamental teaching skills that are a part of the current framework for teaching” (p. 13).

Glickman (1990) noted that, as much as student learning should be considered in the support and supervision of teachers, so should adult learning. Thus, the seventh and final phase of change as noted by Tracy (1995), builds upon the aforementioned human relations focus. As Danielson and McGreal (2000) identified, “Many evaluation systems in use today were developed in the early to mid-1970s and reflect what educators believed about teaching at that time” (p. 3). Their work further noted the impact of Madeline Hunter, who developed a theory-based method for evaluating teaching (Danielson and McGreal, 2000). The seven steps of teaching derived from Madeline Hunter’s work quickly became the basis for many school districts’ evaluation systems; however, there was no clear evidence or linkage of these methods positively impacting student achievement (Danielson and McGreal, 2000).

The 1980s and 1990s were marked by concerns about the global economy and the United States' place in it (Danielson and McGreal, 2000). The narrow focus of previous evaluation systems was expanded to include new research on cognitive learning theory. As Danielson and McGreal (2000) noted, "Adding also to the changing focus of teaching was a new understanding of content knowledge and how content is taught" (p. 14). This increased understanding and research set the stage for much of the evaluation systems in place today. Likewise, with the recent revision of the ESEA, states have been given the flexibility and authority to develop evaluation tools that include a more holistic approach to evaluating the effectiveness of teachers in the classroom.

Best Practices in Teacher Evaluation

Danielson and McGreal (2000) contended "our goals for student achievement have evolved - we are now interested in more complex learning, in problem-solving, the application of knowledge to unfamiliar situations" (p. 3). To that end, teacher expectations and the resulting evaluation and support of educators must also change. Danielson and McGreal (2000) identified four components that must be addressed in the adoption of a new teacher evaluation system; the new evaluation system should be linked to the mission of the district as a continuing process emphasizing student outcomes, and there must be a commitment to allocating adequate resources to all the new systems to be successful (pp. 18-19). Danielson and McGreal (2000) also noted an effective teacher evaluation system is more complex than the document used as the evaluation tool and must contain a succinct definition of the domain of teaching, techniques, and procedures for assessing all aspects of teaching, and well-qualified evaluators who can provide reliable feedback to classroom leaders. To that end, Chetty, Friedman and Rockoff

(2012) linked the importance of a good teacher in the classroom to many lasting effects, most importantly the earning power of the students in the future.

As a key leader in the field of instruction and teacher evaluation, Danielson (2007) worked with the Educational Testing Service (ETS) in 1987 to assist in the development of The Praxis Series: Professional Assessment for Beginning Teachers®. This series consisted of three assessments, the first two, Praxis I: Pre-Professional Skills Assessments and Praxis II: Subject Assessments, are used in many states to grant an initial teaching license (Danielson, 2007). Danielson's (2007) greatest impact was on The Praxis III: Classroom Performance Assessments, which was developed as a tool for administrators to assess actual skills and classroom performance. While many states utilized the Praxis III as part of its licensure expectations, Danielson (2007) received feedback from the Praxis III assessors indicating an interest in using the framework with all novice and experienced educators. As a result of this interest and the effectiveness of the Praxis III, Danielson (2007) developed the *Enhancing Professional Practice: A Framework for Teaching*. Danielson (2007) shared "the framework is based on the Praxis III criteria, augmented to apply to experienced as well as to novice teachers and used for purposes beyond the licensing of beginning teachers" (p. viii). Danielson's (2007) framework has been the foundation for many district and state evaluation systems and consists of four domains, with five to six components then identified as part of each domain. The Kansas ESEA Flexibility Waiver (2015), notes, "careful alignment of the teacher rubrics were made with the Interstate New Teacher Assessment and Support Consortium (INTASC) Performance Standards" (p.253). Likewise, Danielson's (2007) work is also reflective of the INTASC standards with correlations of the framework made

to each of the INTASC standards. Knowing the important impact this framework had in the development of the Kansas Educators Evaluation Protocol, further explanation of each of these domains is necessary to understand best practices in the evaluation of educators.

The first domain noted by Danielson (2007) is that of planning and preparation. While at first glance this domain may appear to focus primarily on the lesson plan, it has a broader application. More specifically, this domain has six components that Danielson (2007) identified as demonstrating knowledge of content, pedagogy and students, setting instructional outcomes, demonstrating knowledge of resources, as well as designing coherent instruction and student assessments.

Danielson's (2007) second domain is the classroom environment and includes five components as indicators. While the physical environment may be the first item to come to mind when the environment is mentioned, Danielson (2007) finds that "the components of the second domain establish a comfortable and respectful classroom environment that cultivates a culture for learning and creates a safe place for risk taking" (p. 28). As further noted in the five components, Danielson (2007) identifies the need for a teacher to create an environment of respect, rapport, and learning, which includes management of procedures and student behavior, as well as the organization of the physical space.

Danielson's third domain is that of instruction and includes five components to consider as indicators for the domain. The five components identified by Danielson (2007) include communicating with students, using questioning and discussion techniques, engaging students in learning, using assessment in instruction and

demonstrating flexibility and responsiveness. Danielson (2007) specifically noted that teachers who master the third domain have “their students engaged in meaningful work, which carries significance beyond the next test and which can provide skills and knowledge necessary for answering important questions or contributing to important projects” (p. 29).

Danielson (2007) shared that “the components in Domain 4 are associated with being a true professional educator; they encompass the roles assumed outside of and in addition to those in the classroom with students” (p. 30). To that end, Danielson (2007) identified six components as essential indicators for this final domain. These include the teacher’s ability to “reflect on teaching, maintaining accurate records, communicating with families, participating in a professional community, growing and developing professionally and showing professionalism” (Danielson, 2007, p. 30).

Danielson (2007) shared, “In the framework for teaching, levels of performance are provided for the four domains and for each of the elements that make up the 22 components of the domains” (p. 39). There are four rating levels referenced in Danielson’s framework (2007), including unsatisfactory, basic, proficient, and distinguished. Criterion is provided for each of the four ratings on each component within each of the four domains. Danielson (2007) found, this feedback regarding the level of performance was particularly useful when “used to support mentoring, coaching, or professional growth” (p. 39).

While Danielson is a recognized leader in the field of teacher effectiveness and evaluation, the research and findings of Marzano have also provided significant insight on maximizing teacher effectiveness through the evaluation process. Marzano et al.

(2011) concluded, “A well-articulated knowledge base for teaching is supported by the success of the Hunter model and the utility of the Danielson model” (p. 27). To that end, Marzano et al. (2011) identified and developed four domains as a foundation of knowledge for all teachers. Within these four domains are 60 elements, which provide more specified traits to support the overarching domain. While similarities can be found between the Danielson model and the model developed by Marzano, the emphasis placed on specific traits/areas differs (Marzano et al., 2011).

Marzano et al. (2011) identified the first domain as classroom strategies and behaviors, which was arranged into the three subcategories of Routine Segments, Content Segments, and Segments Enacted on the Spot. Within these three subcategories are the 41 elements identified in supporting the first domain. Marzano et al. (2011) further state, “This domain directly addresses what teacher do in classrooms” (p 33). Upon further inspection, this domain is directly linked to previous works by Marzano and provides a framework for instructional planning.

Planning and preparing is the second domain noted by Marzano et al. (2011) and includes subcategories for the planning of lessons and units, the use of materials and technology, and preparation and planning for special needs of students. The planning and preparing domain then includes eight elements that are organized within each of the three aforementioned subcategories. While the first domain focused on the activity within the classroom, the second domain’s focus is on the preparation and strategic planning a teacher must consider in their instruction. Marzano et al. (2011) share, “The better a teacher plans and prepares, the more effective are his or her choices of classroom strategies and behaviors” (p 46).

Marzano et al. (2011) find “A teacher who is able to evaluate his or her performance has taken a giant step on the road to expertise” (p 46). To that end, the third domain noted in their work is reflecting on teaching. The two subcategories within domain three highlight the need for teachers to reflect on their teaching and effectiveness. In addition, classroom leaders are pushed to develop and enact plans for their professional growth based on this reflection. To aid in this process, Marzano et al. (2011) have identified five activities to meet this domain: identification of pedagogical strength and weakness within Domain 1, evaluation of the effectiveness of lessons and units, evaluation of the effectiveness of strategies across differing student demographics, development of written growth plans and monitoring subsequent progress.

The final domain identified by Marzano et al. (2011) is collegiality and professionalism. The three subcategories within this domain note the need for teachers to promote a positive culture, share ideas and strategies while promoting the growth and development of the district and school. Found within these three subcategories are six elements for teachers to reference and demonstrate in support of the overarching domain of collegiality and professionalism.

These four domains work together to provide teachers and administrators with a tool for evaluating teacher performance and effectiveness. However, just as important to identifying this knowledge base, is the need for focused feedback. Marzano et al. (2011), found “Focused feedback required clear descriptions of levels of performance” (p 51). A five-point scale is used in Marzano et al.’s (2011) framework and includes ratings from 0-4 with not using (0), beginning (1), developing (2), applying (3), and innovating (4). As shared previously, the four domains within the Marzano et al. (2011) framework are

organized into subcategories and ultimately specified elements. Every element is provided a scale for rating with specified indicators to describe the rating level. Marzano et al. (2011) find this specified feedback to be in sharp contrast with many evaluation and feedback systems, which utilize a checklist format. Furthermore, they note the important role this focused response provides teachers in improving their practice rather than simply reporting what they are or are not doing currently.

In addition to the Danielson and Marzano frameworks, Mid-continent Research for Education and Learning (McREL) has also developed an evaluation system to meet the ever-increasing complexities of evaluating teachers. While many comparisons can be made and similarities noted between the Marzano and Danielson frameworks, the foundation for, and organization of, the McREL evaluation cannot be as easily compared. While many themes can be identified among all three evaluation frameworks, it takes greater analysis to find these commonalities in the McREL framework. As noted in the McREL's Teacher Evaluation System (2009) "This evaluation instrument and accompanying process is based on elements of a 21st century education and a set of rigorous research-based standards" (p. 1). McREL has identified five standards with components provided under each to detail further the necessary skills and attributes.

The first standard identified in the McREL (2009) structure is teacher leadership. This standard is delineated into five components, which focus on the teacher's leadership in the classroom, school, and profession. Additionally, there is a focus on the teacher's attributes in advocating for the school and students while holding high moral and ethical standards for themselves (McREL, 2009). McREL (2009) further noted, "Leadership

among the staff and with the administration is shared in order to bring consensus and common, shared ownership of the vision and purpose of the school's work" (p. 3).

The second standard addressed through the McREL (2009) evaluation framework is the teacher's ability to create a respectful environment for all students. This standard includes five components to be addressed, which more specifically detail the need for a classroom instructor not only to provide a supportive classroom environment and treat every child as an individual, but also to embrace differences and diversity while addressing the needs of special populations. Also referenced in this standard is the need for teachers to develop and maintain strong collaborative partnerships with families and the support network of their students outside of the school walls. McREL (2009) shared, "Teachers facilitate instruction, encouraging all students to use 21st century skills so they discover how to learn, innovate, collaborate, and communicate their ideas" (p. 3).

McREL (2009) noted, "Teachers no longer cover material; they, along with their students, uncover solutions" (p. 3). In support of this 21st century skill, the third standard included in the McREL (2009) framework addresses and measures teachers' mastery of the content they are teaching. More specifically, McREL identifies four components necessary for meeting this standard. The first component or skill is the teacher's ability to align his/her instruction to district and state standards, the second addresses the need for teachers to possess a depth of knowledge in their content area, while the third looks for a teacher's ability to make cross-curricular connections. The fourth and final component measures the teacher's ability to make learning relevant to students and their personal lives. McREL again connects this standard with 21st century skills by sharing,

“Teachers make the content they teach engaging, relevant, and meaningful to students’ lives” (p. 3).

McREL’s (2009) fourth standard incorporates a focus on eight components, all of which support the expectation that, “Teachers facilitate learning for their students” (p. 19). To that end, the eight components or skills addressed within this standard focus on the teacher’s awareness of, and planning for, students’ instructional, emotional, and social levels. Likewise, emphasis is placed on the teacher’s ability to differentiate instruction based on the needs of students while finding innovative ways to incorporate technology and support critical thinking. Collaboration and teaming among students is also addressed, as well as the teacher’s ability to clearly communicate and assess student’s progress.

The fifth and final standard included in the McREL (2009) model, is teacher reflection. McREL (2009), linked this fifth standard to the 21st century skill which notes, “Teachers are reflective about their practice and include assessments that are authentic, structured, and that demonstrate student understanding” (p. 3). To provide more specificity regarding this expectation, this standard is organized into three components, which focus on a teacher’s ability to ask critical questions regarding student learning and develop professional goals based on this data. Lastly, this standard incorporates a measure of a teacher’s ability to adapt to and navigate an ever-changing environment within their classroom, school, and district.

These five standards and their supporting components/skills comprise the McREL evaluation system. Similar to the Danielson and Marzano frameworks, the McREL system also includes a five-point rating system for each of the 25 components identified

within the five standards. Accompanying these ratings is a descriptor as to what must be observed and/or evidenced to support that specified rating. McREL (2009) delineated these ratings into “developing, proficient, accomplished, distinguished, and not demonstrated” (p. 11).

Hanushek (2012) shared that “we have recently seen a number of brave states step out and legislate better evaluations of teachers including, when possible, the use of value-added measures” (para. 8). The comprehensive work of Darling-Hammond (1999), noted that the impact of a well-prepared teacher on student achievement can outweigh the effects of student demographic factors, including poverty, minority status, and language of origin. Hutchison-Lupardus and Snyder’s (2012) investigation found that many state systems included value-added measurements as part of teacher evaluations that were not in alignment with current research and best practice. However, they did find a “need for continual strides to be made toward the development of national teaching standards and a national teacher evaluation model” (p. 129).

Perceptions of Teacher Evaluation Effectiveness

The recent implementation of new teacher evaluation systems in many states has prompted much research into the perceived effectiveness of these new systems. Doherty (2009) shared, “Since 2003, the Massachusetts suburban school district being studied has involved teachers and administrators in the development and implementation of a standards-based teacher evaluation system, called the Teacher Assessment Process (TAP)” (p. 3). To measure teacher and administrator’s perceptions of the effectiveness of this newly adopted evaluation system, Doherty (2009) conducted a quantitative study, which included 170 teachers and 14 administrators in a suburban Massachusetts school

district. The research questions focused on various measures of teacher's perceptions regarding the impact on instruction because of the standards-based evaluation model. However, of particular interest was the difference in perception held by administrators and teachers concerning a shift in the focus of the teacher evaluation system from employment decisions to teacher growth. Doherty (2009) disaggregated the quantitative and qualitative data into elementary and secondary teachers' perceptions. Results indicated that "elementary teachers perceived the teacher evaluation system as having a stronger impact on improved teacher instruction, sustained school improvement, increased student learning and elevated professional growth than their secondary counterparts" (p. iii). Conversely, Doherty (2009) found that high school teachers and administrators found these four areas to be least impacted by the teacher evaluation system. Results of Doherty's (2009) work indicated that administrators found the standards-based evaluation had positively influenced new teachers, as well as improved instruction with technology and educational objectives. In addition, Doherty (2009) found that administrators viewed the evaluation tool as a means of improving instruction and less as a tool for making employment decisions.

Killian (2010) initiated research into the perceptions of secondary teachers and secondary administrators throughout Missouri. Killian's (2010) work included 859 secondary school administrators and 4,825 teachers from public school districts throughout Missouri as the population referenced in the study. Furthermore, Killian (2010) shared "A random sampling of 1,000 secondary teachers and administrators from this pool were invited to respond by an invitational letter to an online survey" (p. 78). While the majority of the survey was quantitative in nature and included responses to

twenty-three statements with a corresponding six-point Likert-type scale, two constructed-response items were also included. This mixed-method study was specifically inquiring as to teachers' and administrators' perceptions of the Missouri Performance Based Evaluation Model (Killian, 2010). One notable finding of Killian's (2010) work was teacher's preference for professional growth to be impacted by professional development and mentoring versus the evaluation tool as the primary source in devising professional learning plans.

The work of Sheppard (2013) provides a source of insight regarding teacher and administrator's perceptions of teacher evaluation prior to any change. More specifically, Sheppard (2013) noted, "In many states, including Georgia, school systems are looking at ways to evaluate teachers that offer a somewhat more structured and more systematic approach to teacher evaluations" (p. 3). Although at the time of the study, Georgia had not required any systematic change in teacher evaluations, Sheppard's work provides insight into teacher and administrator's perceptions of their current teacher evaluation system. In this endeavor, Sheppard (2013) also utilized the Teacher Evaluation Profile (TEP) survey instrument, which consists of 46 questions with a five-point Likert-type scale for responses. In addition to these responses, an added open-construct question was also included as part of the survey. A purposive sampling of 12 administrators and 277 teachers were included in the analysis of this research. Sheppard (2013) found, "A number of teachers (43.73%) believed that the evaluation process in their system was average and that these evaluations had a strong impact on professional practices (20.15%)" (p. 57). Administrators and teachers differed in their perception of the strongest attribute of the evaluation system, with teachers identifying feedback based on

standards as the greatest strength, while administrators found the timing of the evaluation to be the most beneficial (Sheppard, 2013). Similarly, administrators were found to believe that student learning was impacted the most by the teacher evaluation process (Sheppard, 2013).

Bridich (2013) also contributed to research into teacher and administrative perceptions of their roles in the implementation of a new evaluation system as mandated by Colorado Senate Bill 10-191. The sample group used in the study were teachers and administrators in the Rockies School District, located in the greater Denver, Colorado area, and policymakers who were members of Colorado's 2013 General Assembly (Bridich, 2013). This mixed method research found differences in the perceptions of teachers and administrators regarding their role in the implementation of the new evaluation system. Bridich (2013) noted the disparity between administrators and teachers' views of their role in the process. More specifically, administrators held the perception they had a much greater voice in the adoption and implementation of the evaluation system than their teacher counterparts. A common theme of concern noted by Bridich (2013), held by the administrator and teacher sample groups, was the inclusion of student growth measures as part of the evaluation system. The three sample groups of policymakers, teachers, and administrators did not report a strong belief that the new evaluation system would result in improved student learning (Bridich, 2013).

Coulter (2013) provided qualitative research into the perceptions of the level of support six teachers and six principals held concerning the newly adopted teacher evaluation systems in the state of Washington. Similar to many other states, Washington also had updated its teacher evaluation processes and expectations through Senate Bill

5895 in 2012 (Coulter, 2013). As part of the passing of this new legislation, school districts were given the option to select from three evaluation systems, the Danielson Model, Marzano Model, or Center for Educational Leadership's Five Dimensions of Teaching and Learning (CEL 5D+). Coulter's (2013) work included interviews with two teachers and two administrators working in districts that had adopted one of the three models. Thus, two teachers and two principals, working in districts that had adopted the Danielson Model, as well as two teachers and two principals working in districts that had adopted the Marzano Model, and two teachers and two principals working in districts that were using the CEL 5D+ evaluation system, were interviewed. The focus of Coulter's (2013) work was on the items used in identifying effective teachers, as well as the newly adopted evaluation system's effectiveness. One finding through Coulter's (2013) research was the common agreement among all teachers and administrators in their opposition to the use of student growth as a component of teacher evaluation. Although Senate Bill 5895 mandated the inclusion of student growth in teacher evaluation, all participants in Coulter's (2013) study disagreed with this expectation.

Similar to the legislation changes referenced in Colorado and Washington, Heyde (2013) evaluated the effects of the Illinois Performance Evaluation Reform Act of 2010 (Public Act 96-861), as modified by Senate Bill 7. This bill was signed into law in June of 2011 and required districts in the state of Illinois to begin shifting to a performance-based system of teacher evaluation for the purpose of teacher retention (Heyde, 2013). Through a mixed method research design, 120 teachers and 16 administrators in three suburban K-8 Chicago school districts were administered surveys and interviewed as part of Heyde's research. Heyde (2013) found a general level of contentment with the current

evaluation systems, “But the current systems do not provide any real differentiation among teachers, and both administrators and teachers know this” (p. 55). Another weakness in the current system, as cited by Heyde’s (2013) work, is the limited use of walk-through data in current evaluation systems. Although Heyde’s (2013) research was focused on perceptions of the current evaluation system, prior to any required change from Senate Bill 7, it is helpful baseline data. Knowing this study can then be replicated in the future to measure and compare perceptions of teachers and administrators after implementing reformed evaluation systems, this insight will be most beneficial. Based on results of Heyde’s (2013) work, there is also an implication of the need to get teachers involved in the evaluation selection and adoption process. As Heyde (2013) found, most teachers were unaware of any impending change in the evaluation systems. Thus, providing teachers with a voice may assist in setting a foundation for a more successful and effective evaluation system, as teachers could then speak to the value of the system rather than seeing it as a political movement.

As another important insight and point of comparison, the work of Clark (2014) focused on principals’ perceptions of teacher evaluations in the state of Nebraska. Of particular interest in this study was the evaluation system’s ability to identify effective teachers. While Nebraska had instituted a teacher and principal performance framework, a specified evaluation system was not mandated. Likewise, there was no incorporation of student growth measures as part of the evaluation system’s means for determining teacher effectiveness (Clark, 2014). While this void in the system was further discussed by Nebraska legislators, it did not result in a mandate for the addition of any value added measures in the teacher evaluation systems. The lack of any reference to student

achievement garnered much scrutiny from various professional organizations, as well as various research bodies. With this in mind, Clark (2014) used a mixed methods approach to garner responses from all Nebraska principals as to their perceptions of their current evaluation systems ability to appropriately identify and measure teacher effectiveness. Results indicated that Nebraska principals utilized their respective evaluation tools as a means to promote professional growth; however, they also viewed it as a vehicle for identifying and dismissing ineffective educators (Clark, 2014). He further noted that Nebraska is currently piloting an evaluation system that includes multiple measures, including student growth; however, it had not yet been implemented at the time of this study.

As a result of this national movement, Arizona has also implemented sweeping changes and expectations for the evaluation of teachers and principals. Senate Bill 1040 was enacted in May of 2010 and requires school districts to evaluate teachers using an evaluation instrument that meets the requirements of the state board–approved Arizona Framework for Measuring Educator Effectiveness (Ruffini et al., 2014). Partnering with the Regional Educational Laboratory West and the Arizona Department of Education, Ruffini et al. (2014) conducted a descriptive study tapping into the perceptions and feedback of 10 school districts across Arizona. While five of the districts included in the study had implemented the Arizona State Department of Education tool, the other five were utilizing district-developed tools that met the criteria outlined in the Arizona Framework for Educator Effectiveness. Ruffini et al. (2014) focused on the challenges and unintended consequences found after the first year of implementation of these new or revised evaluation tools. More specifically, teacher’s perceptions of the precision and

value of the new evaluation, as well as any perceived changes in teaching or collaboration between teachers and/or administrators were measured. Fifty-four principals and 92 teachers participated in the focus group interviews as part of the study, and 157 teachers responded to a survey. Although principals found the training sufficient for implementation of the new evaluation system, teachers reported feeling they were not fully informed of all necessary evidence needed to support the nonobservational domains. Furthermore, principals shared concerns with the time commitment necessary to complete the new evaluation system. Ruffini et al. (2014) reported support from principals and teachers regarding the effectiveness of the classroom observations and resulting feedback. Concerns were raised by teachers regarding inter-rater reliability, as well as the methodology applied in factoring student achievement as part of the pilot-year evaluation system. However, Ruffini et al. (2014) found greater support by teachers regarding the use of student standardized assessments in measuring student progress and ultimately teacher effectiveness. Another concern noted by teachers in the sample group was the manner in which the summative rating was calculated. More specifically, Ruffini et al. (2014) noted concerns with the fairness of weightings used, as well as supplemental assessments that were considered and tabulations made for those instructing in other content and grade levels.

A similar body of research was conducted by Paufler in 2014 and included elementary teachers and principals in a large Arizona school district. Paufler's (2014) work sought to answer questions regarding teacher and administrator's perceptions of the goals and purpose of the new evaluation tool. Likewise, Paufler (2014) included measures of the actual and intended implementation procedures for the evaluation system

and its effectiveness. Through a mixed-methods research design, both surveys and interviews were utilized with meta inference on a sample teacher group. The results of Paufler's (2014) multi-phased research identified five common assertions. The first identified administrators and teachers shared belief the evaluation system should be used to improve professional practice. While both sample groups held agreement on this belief, teachers disagreed that this stated purpose was the reason for the implementation. The second assertion made by Paufler (2014), was that administrators had an increased understanding of the evaluation system components and process. Conversely, teachers did not share this increased understanding to the same degree as their administrative counterparts. Paufler (2014) asserted that differences in administrator and teacher's perception of validity, reliability, and fairness were impacted by a perceived misalignment, idealistically and in reality, of the evaluation purposes. The fourth assertion of Paufler's (2014) work was a difference in administrator and teacher perception of the evaluation system's impact on professional practice, as well as a teacher's ability to influence the outcome of their evaluation. More specifically, administrators held the perception that teachers could affect their evaluation and that the evaluation positively influenced teacher's professional practice. However, teachers shared their belief that they had little control over the outcome of their evaluation, and the system itself had little impact on their professional practice. The final finding of Paufler (2014) was a major difference in administrator and teacher's perceptions of the evaluation's popularity and adaptiveness. Both sample groups shared concerns with unintended consequences of the new evaluation system, which could affect the system's long-term effectiveness and viability. Ultimately, Paufler (2014) concludes, "As subjects

of, and actors within, teacher evaluation policies in context, school administrators and teachers should largely determine and assess the utility of the standards used to measure system effectiveness” (p. 184).

A similar method of inquiry was conducted by Winslow (2015) whose research was specifically aimed at the perceptions of elementary and middle school teachers and administrators in the Triad Community School District #2 in Troy, IL. Research questions were focused on the perceptions held by these two groups about the implementation of the Danielson Framework as their evaluation tool. The mixed-methods study utilized both surveys (for teachers) and personal interviews of both teachers and principals to garner a better understanding of the qualitative data received. Due to these survey results and interviews, Winslow (2015) identified seven themes; observation feedback is provided in a different manner which teachers favored, both teachers and principals perceived the Danielson process to be formative but had process concerns, teachers and principals note increased student-centered activities which include more higher-order thinking, teachers and principals referenced face-to-face conversations which included feedback, teachers and administrators were aware of the impact of evidence following observations and the impact on summative ratings, administrators identified the time commitment in implementing the new evaluation system was great, and both principals and teachers had concerns with the training needed for implementation. While Winslow’s (2015) work specifically measured teacher and principal perceptions, there were implications referenced to the important role of the superintendent in the adoption of a new evaluator process. More specifically, Winslow (2015) noted, “The superintendent must also support the immense amount of time

necessary for a quality evaluation system by creating opportunities where teachers and principals may have dialogue throughout the school day and during other non-instructional time” (p 78).

While much research has been conducted on teachers’ and principals’ perceptions of teacher evaluation systems, the work of Derrington (2014) provided insight from the perspective of superintendents. Through this qualitative study, which focused on principals and superintendents from a southeastern state, Derrington (2014) found unanimous agreement between both sample populations regarding the benefits, concerns, and unintended consequences of the new evaluation system. More specifically, both principals and superintendents identified instructional leadership, increased principal instructional knowledge, and increased principal evaluation competency as benefits (Derrington, 2014). Concerns cited by both principals and superintendents were time and training tensions and the unintended consequence of increased principal burnout (Derrington, 2014).

Summary

The review of the literature provided insight into the history and evolution of teacher evaluations in the United States. In addition, three strategic models for effectively evaluating teachers were also discussed, with details regarding similarities and differences among the frameworks noted. A summary of research that has been conducted in the area of teacher and administrator perceptions of teacher evaluation was also included as part of the discourse and foundation for the current body of research. Chapter three includes the research design; population and sample; sampling procedures; instrumentation including measurement, validity and reliability; data collection

procedures; data analysis and hypothesis testing; and limitations related to this research study.

Chapter Three

Methodology

The purpose of this study was to determine superintendents' perceptions regarding the effectiveness of their districts' evaluation system in meeting the six components of Principle Three of the ESEA waiver. An additional purpose of this study was to determine whether the demographic factor of district size affected superintendents' perceptions regarding the effectiveness of their districts' evaluation system in meeting the six components of Principle Three of the ESEA waiver. This chapter includes the specific methodology utilized in this research. Provided in this chapter is information regarding the research design; population and sample; sampling procedures; instrumentation including measurement, validity and reliability; data collection procedures; data analysis and hypothesis testing; and limitations.

Research Design

Creswell (2009) stated, "A survey design provides a quantitative or numeric description of trends, attitudes, or opinions of a population by studying a sample of that population" (p. 145). A quantitative descriptive research design was used in this study, to assess superintendents' perceptions of the effectiveness of their districts' evaluation system in meeting the six guidelines for component three of the ESEA waiver in Kansas. More specifically, the dependent variables were the superintendents' perceptions regarding the effectiveness of their districts' evaluation system in supporting continual improvement of instruction; differentiating teacher performance using at least three performance levels; using multiple measures in determining teacher performance levels; including data on student growth for all students as a significant factor in determining

performance levels; including other measures of professional practice; evaluating educators on a regular basis; providing clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and informing personnel decisions. The independent variable was the district demographic factor of size.

Population and Sample

The population utilized in this study included all K-12 public school superintendents in the state of Kansas. The sample included all superintendents in Kansas with current active email accounts included in the Kansas State Department of Education 2015-2016 Superintendent Directory. Only responses received from the superintendents were included as part of the sample.

Sampling Procedures

A nonrandom purposive sampling method was used in this study. This sampling method allowed for the purposeful selection of “participants and sites who...best help in understanding the research problem or questions” (Creswell, p. 231). The criterion for participation was set to include public school superintendents with current active email accounts throughout the state of Kansas.

Instrumentation

An online survey instrument, the Teacher Evaluation Survey, was developed specifically for this study. This instrument was developed after conducting research on the Kansas State Department of Education ESEA Flexibility request (KSDE, 2015) document to represent succinctly and accurately the components included as Principle Three of the Kansas ESEA Waiver Request. Background for the study and a summary of

the six components of Principle Three of the waiver were provided to all respondents before administering the survey. The first 11 items on the survey were taken directly from the six components of Principle 3 of the ESEA waiver. Bearing in mind the six components comprising Principle Three of the ESEA Waiver provide 11 specific expectations, the specific research questions addressed in this study have been tailored to measure each of the expectations delineated in the six components. Specifically, there are two research questions to address component three and five research questions to address component five of the ESEA Waiver. A five-point Likert-type scale was utilized to report quantitatively superintendents' perceptions of the effectiveness of their districts' evaluation system in meeting the six components of Principle Three of the ESEA waiver. Superintendents were asked to provide their level of agreement on the following scale: *1- Strongly Disagree, 2- Disagree, 3- Neither Agree nor Disagree, 4- Agree, 5- Strongly Agree*. A copy of the Teacher Evaluation Survey can be found in Appendix A. The final item included as part of the survey was an open- response format inquiring as to the enrollment of the superintendent's school district per the September 20, 2015 enrollment count. Items 1-11 addressed superintendents' perceptions of their evaluation system by indicating their level of agreement with the following:

- *My district's evaluation system supports continual improvement of instruction.*
- *My district's evaluation system meaningfully differentiates performance using at least three performance levels.*
- *My district's evaluation system includes, as a significant factor, data on student growth for all students in determining performance levels.*

- *My district's evaluation system effectively includes other measures of professional practice.*
- *My district's evaluation tool effectively supports the evaluation of educators on a regular basis.*
- *My district's evaluation system effectively provides clear feedback.*
- *My district's evaluation system effectively provides timely feedback*
- *My district's evaluation system effectively provides useful feedback.*
- *My district's evaluation system includes feedback that identifies needs.*
- *My district's evaluation system includes feedback that guides professional development.*
- *My district's evaluation system is effectively used to inform personnel decisions.*

Participants responded on the five-point Likert-type scale to each of these eleven statements.

Measurement. The dependent variable in RQ1 and RQ2 was superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction. The variable was measured using the calculated mean response to item 1 on the survey. The independent variable for RQ2 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ3 and RQ4 was superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels. The variable was measured using

the calculated mean response to item 2 on the survey. The independent variable for RQ4 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ5 and RQ6 was superintendents' perceptions of their districts' evaluation systems effectiveness in including, as a significant factor, data on student growth for all students in determining performance levels. The variable was measured using the calculated mean response to item 3 on the survey. The independent variable for RQ6 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ7 and RQ8 was superintendents' perceptions of their districts' evaluation systems in effectively including other measures of professional practice. The variable was measured using the calculated mean response to item 4 on the survey. The independent variable for RQ8 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ9 and RQ10 was superintendents' perceptions of their evaluation systems effectiveness in supporting the evaluation of educators on a regular basis. The variable was measured using the calculated mean response to item 5 on the survey. The independent variable for RQ10 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ11 and RQ12 was superintendents' perceptions of their districts' evaluation systems in effectively providing clear feedback. The variable was measured using the calculated mean response to item 6 on the survey. The

independent variable for RQ12 was district size, which was measured utilizing the response to item 12 (district enrollment as provided in the survey).

The dependent variable in RQ13 and RQ14 was superintendents' perceptions of their districts' evaluation systems in effectively providing timely feedback. The variable was measured using the calculated mean response to item 7 on the survey. The independent variable for RQ14 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ15 and RQ16 was superintendents' perceptions of their districts' evaluation systems in effectively providing useful feedback. The variable was measured using the calculated mean response to item 8 on the survey. The independent variable for RQ16 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ17 and RQ18 was superintendents' perceptions of their districts' evaluation systems in effectively including feedback that identifies needs. The variable was measured using the calculated mean response to item 9 of the survey. The independent variable for RQ18 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ19 and RQ20 was superintendents' perceptions of the effectiveness of their evaluation systems in providing feedback that guides professional development. The variable was measured using the calculated mean response to item 10 of the survey. The independent variable for RQ20 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

The dependent variable in RQ21 and RQ22 was superintendents' perceptions of the effectiveness of their districts' evaluation systems in informing personnel decisions. The variable was measured using the calculated mean response to item 11 of the survey. The independent variable for RQ22 was district size, which was measured utilizing the response to item 12 (district enrollment as provided on the survey).

Validity and reliability. To ensure validity in the survey design, doctoral program advisors collaborated in the development of the tool. Two university dissertation supervisors reviewed and provided comments regarding the evaluation tool and procedures. Although initial drafts of the survey included demographic questions regarding the gender of the respondent, it was agreed that this factor was not central to the purpose of the study; therefore, this question was removed. In addition, because of this feedback, changes were made to the ordering and wording of the questions included in the survey. A second draft of the survey was then submitted to the university committee. Feedback from this version was again gathered, and the survey was updated with the new ordering of the survey items, as well altering the items from a question format to statements with the five-point Likert-type scale utilized to gather responses. Upon receiving approval from the university committee, further testing of the survey took place with four assistant superintendents. Each of these individuals had recently served as a superintendent in the state of Kansas. Feedback from each of the previous superintendents indicated a clear understanding of the questions, as well as alignment with the purpose of the study. Therefore, no further revisions were necessary.

Sackett and Larson (1990) found “most commonly used single-item measures can be divided into two categories: (a) those measuring self-reported facts and (b) those

measuring psychological constructs. If the construct being measured is sufficiently narrow or is unambiguous to the respondent, a single item may suffice” (p. 631). The individual items used in this research were self-reported facts that were sufficiently narrow and unambiguous. Therefore, reliability was not an issue for the measurement using this survey instrument.

Data Collection Procedures

Approval to collect data and conduct the research was initially sought from the Baker University Institutional Review Board on November 2, 2015 (see Appendix B). Formal approval was received from the IRB on November 4, 2015 (see Appendix C). All Kansas superintendents’ email addresses were gathered via the Kansas Department of Education 2015-2016 directory. These email contacts were then utilized to send each superintendent an invitation to participate in this study. The email (see Appendix D) provided an explanation for the purpose of the study, as well as clarification that participation was voluntary. Likewise, the email communication indicated that by responding to the survey, the respondent was providing consent to participate. Lastly, the email also confirmed that the reporting of data would only be in a summary format. A link to the survey platform SurveyMonkey was also included in the body of the email communication. The survey link was initially sent to all Kansas public school superintendents on December 2, 2015. A reminder email was then sent to those, who had not yet responded, on December 19, 2015. A third and final reminder was then provided on January 27, 2016.

Data Analysis and Hypothesis Testing

Data were downloaded from SurveyMonkey and imported into IBM® SPSS® Statistics Faculty Pack 23 for Windows addressing the twenty-two research questions. The research question, corresponding hypothesis, and method for statistical analysis are provided below for each item.

RQ1. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction?

H1. Superintendents perceive that their districts' evaluation systems are effective in supporting continual improvement of instruction.

A one-sample *t* test was conducted to test H1. The average perception of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction was tested against a null value of 3. The level of significance was set at .05.

RQ2. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction affected by the size of the district?

H2. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H2. The categorical variable used to group the mean perception of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction was the size of the district. The districts were organized into six categories, the first category

included districts with a Pre-K through 12th grade enrollment of 68-299 students, the second category included districts with an enrollment of 300-599 students, the third category included districts with an enrollment of 600-899 students, the fourth category included districts with an enrollment of 900-1,999 students, the fifth category included districts with an enrollment of 2,000-4,999 students, and the sixth and last category included districts with an enrollment of 5,000-22,000 students. The level of significance was set at .05.

RQ3. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels?

H3. Superintendents perceive that their districts' evaluation systems meaningfully differentiates performance using at least three performance levels.

A one-sample *t* test was conducted to test H3. The average perception of the effectiveness of their districts' evaluation systems meaningfully differentiating performance using at least three performance levels was tested against a null value of 3. The level of significance was set at .05.

RQ4. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels affected by the size of the district?

H4. Superintendents' perceptions of the effectiveness of their districts' evaluation systems meaningfully differentiating at least three performance levels are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H4. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems meaningfully differentiating at least three performance levels was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ5. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on student growth for all students in determining performance levels?

H5. Superintendents perceive that their districts' evaluation systems effectively includes, as a significant factor, data on student growth for all students in determining performance levels.

A one-sample *t* test was conducted to test H5. The average perception of the effectiveness of their districts' evaluation systems effectively including, as a significant factor, data on student growth for all students in determining performance levels was tested against a null value of 3. The level of significance was set at .05.

RQ6. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in including student growth data for all students as a significant factor in determining performance levels affected by the size of the district?

H6. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including student growth for all students as a significant factor in determining performance levels are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H6. The categorical variable used to group the average perception of the effectiveness of their

districts' evaluation systems meaningfully differentiating at least three performance levels was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ7. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice?

H7. Superintendents perceive that their districts' evaluation systems effectively includes other measures of professional practice.

A one-sample *t* test was conducted to test H7. The average perception of the effectiveness of their districts' evaluation systems effectively including other measures of professional practice was tested against a null value of 3. The level of significance was set at .05.

RQ8. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice affected by the size of the district?

H8. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including other measures of professional practice are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H8. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems meaningfully differentiating at least three performance levels was the size of the district. The districts were organized into six categories, As described following H2. The level of significance was set at .05.

RQ9. What are superintendents' perceptions of their evaluation systems effectiveness in effectively supporting the evaluation of educators on a regular basis?

H9. Superintendents perceive that their districts' evaluation systems effectively supports the evaluation of educators on a regular basis.

A one-sample *t* test was conducted to test H9. The average perception of the effectiveness of their districts' evaluation systems effectively supporting the evaluation of educators on a regular basis was tested against a null value of 3. The level of significance was set at .05.

RQ10. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively evaluating educators on a regular basis affected by the size of the district?

H10. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively evaluating educators on a regular basis are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H10. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems evaluating educators on a regular basis was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ11. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback?

H11. Superintendents perceive that their districts' evaluation systems effectively provides clear feedback.

A one-sample *t* test was conducted to test H11. The average perception of the effectiveness of their districts' evaluation systems effectively providing clear feedback was tested against a null value of 3. The level of significance was set at .05.

RQ12. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback affected by the size of the district?

H12. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H12. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in effectively providing clear feedback was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ13. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing timely feedback?

H13. Superintendents perceive that their districts' evaluation systems effectively provides timely feedback.

A one-sample *t* test was conducted to test H13. The average perception of the effectiveness of their districts' evaluation systems effectively providing timely feedback was tested against a null value of 3. The level of significance was set at .05.

RQ14. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing timely feedback affected by the size of the district?

H14. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing timely feedback are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H14. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in effectively providing timely feedback was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ15. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing useful feedback?

H15. Superintendents perceive that their districts' evaluation systems effectively provides useful feedback.

A one-sample *t* test was conducted to test H15. The average perception of the effectiveness of their districts' evaluation systems effectively providing useful feedback was tested against a null value of 3. The level of significance was set at .05.

RQ16. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in providing useful feedback affected by the size of the district?

H16. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing useful feedback are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H16. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in effectively providing useful feedback was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ17. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in including feedback that identifies needs?

H17. Superintendents perceive that their districts' evaluation systems effectively includes feedback that identified needs.

A one-sample *t* test was conducted to test H17. The average perception of the effectiveness of their districts' evaluation systems effectively including feedback that identified needs was tested against a null value of 3. The level of significance was set at .05.

RQ18. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in including feedback that identifies needs affected by the size of the school district?

H18. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including feedback that identified needs are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H18. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in effectively including feedback that identifies needs was

the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ19. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing feedback that guides professional development?

H19. Superintendents perceive that their districts' evaluation systems effectively provides feedback that guides professional development.

A one-sample *t* test was conducted to test H19. The average perception of the effectiveness of their districts' evaluation systems is effectively providing feedback that guides professional development was tested against a null value of 3. The level of significance was set at .05.

RQ20. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in guiding professional development affected by the size of the school district?

H20. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively guiding professional development are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H20. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in effectively guiding professional development was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

RQ21. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions?

H21. Superintendents perceive that their districts' evaluation systems effectively informs personnel decisions.

A one-sample *t* test was conducted to test H21. The average perception of the effectiveness of their districts' evaluation systems effectively informing personnel decisions was tested against a null value of 3. The level of significance was set at .05.

RQ22. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation system in effectively informing personnel decisions affected by the size of the district?

H22. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions are affected by the size of the district.

A one-factor analysis of variance (ANOVA) was conducted to test H22. The categorical variable used to group the average perception of the effectiveness of their districts' evaluation systems in informing personnel decisions was the size of the district. The districts were organized into six categories, as described following H2. The level of significance was set at .05.

Limitations

As with any research study, there are factors that may influence the outcome but cannot be controlled by the researcher. While they may not be within the control of the researcher, these limitations must be declared as they "may have an effect on the interpretations of the findings or on the generalizability of the results" (Lunenburg & Irby, 2008, p. 133). Knowing superintendents were providing feedback based on the teacher evaluation tool in their districts, they may not be critical of the processes.

Likewise, some superintendents answered the questions without having the knowledge to respond accurately to the items. Although participants were provided insight as to how the collected data would be reported, some individuals who prefer not to respond to electronic surveys may not have participated. Likewise, every effort was taken to develop a tool that included twelve items and would take five to ten minutes to complete; some participants may not have taken the time to complete and submit the survey. Therefore, data analysis was limited to those fully-completed and submitted response surveys.

Summary

Included in Chapter three were further explanations of the research design, population and sample, sampling procedures, instrumentation, measurement, and validity and reliability of the study. Similarly, the research questions and hypotheses for each of these questions were discussed in detail. The population was Kansas superintendents at the time the survey was administered. The sample included the Kansas superintendents who chose to complete the survey. Lastly, the data collection and analysis procedures for each of the hypotheses were also described, as well as the limitations of the study.

Chapter four includes the results of the one-sample t tests and ANOVAs conducted in the hypothesis testing.

Chapter Four

Results

The purpose of this study was to determine superintendents' perceptions regarding the effectiveness of their districts' evaluation systems in meeting the six components of Principle Three of the ESEA waiver: supporting continual improvement of instruction; meaningfully differentiating teacher performance using at least three performance levels; using multiple measures in determining teacher performance levels; including data on student growth for all students as a significant factor in determining performance levels; including other measures of professional practice; evaluating educators on a regular basis; providing clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and informing personnel decisions. While there are six components listed as part of Principle Three, these components included eleven specific expectations. Therefore, there were eleven questions included as part of the survey to address each of these expectations. An additional purpose of this study was to determine whether the demographic factor of district size affected superintendents' perceptions regarding the effectiveness of their districts' evaluation systems in meeting the six components of Principle Three of the ESEA waiver. This chapter contains the descriptive statistics and the results of the one-sample *t* tests and one-factor ANOVAs.

Descriptive Statistics

The population for this research was superintendents throughout the state of Kansas. The survey was sent to the 286 superintendents throughout the state, and 128 responses were received. As part of the survey, demographic information was requested

regarding the enrollment of their respective school district. Of the 128 respondents, 126 provided their district's enrollment. The enrollments for the 126 districts were organized into six categories. These demographics are shared in Table 1 and provide the six categories as well as the number of districts included in each category.

Table 2

District Enrollment Demographics

Category	Enrollment	<i>N</i>
1	68-299	25
2	300-599	38
3	600-899	21
4	900-1,999	18
5	2,000-4,999	17
6	5,000-22,000	7

The descriptive statistics calculated as part of this research provided a frame of reference regarding the sample group. More specifically, this enrollment data indicates the great disparity in the size and number of districts represented within each category. The subsequent section provides the results of the hypothesis testing to support the conclusions drawn regarding superintendents' perceptions of the effectiveness of their evaluation tool in meeting the six guidelines of the ESEA Waiver.

Hypothesis Testing

The focus of the study surrounded 22 specific research questions. Each question is presented below with corresponding hypotheses and results of the statistical analysis for each. The level of significance was set at .05.

RQ1. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction?

H1. Superintendents perceive that their districts' evaluation systems are effective in supporting continual improvement of instruction.

A one-sample t test was conducted to test H1. The average perception of the effectiveness of superintendents' districts' evaluation systems in supporting continual improvement of instruction was tested against a null value of 3. The results of the one-sample t test indicated a statistically significant difference between the two values, $t = 19.306$, $df = 127$, $p = .000$. The sample mean ($M = 4.188$, $SD = .696$) was higher than the null value (3). Superintendents perceived their evaluation systems were effective in supporting continual improvement of instruction.

RQ2. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction affected by the size of the district?

H2. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction are affected by the size of the district.

A one-factor ANOVA was conducted to test H2. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in supporting continual improvement of instruction was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.466$, $df = 5, 120$, $p = .206$. See Table 2 for the means and standard deviations for this analysis. A follow up post hoc was not warranted.

Superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction were not affected by the size of the district.

Table 3

Descriptive Statistics for the Results of the Test for H2

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	4.160	0.473
2 (300-599)	38	4.132	0.704
3 (600-899)	21	4.381	0.590
4 (900-1,999)	18	4.389	0.502
5 (2,000-4,999)	17	4.059	0.659
6 (5,000-22,000)	7	3.714	1.604

RQ3. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels?

H3. Superintendents perceive that their districts' evaluation systems meaningfully differentiates performance using at least three performance levels.

A one-sample *t* test was conducted to test H3. The average perception of the effectiveness of superintendents' districts' evaluation systems in meaningfully differentiating performance using at least three performance levels was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 17.560$, $df = 127$, $p = .000$. The sample mean ($M = 4.211$, $SD = .780$) was higher than the null value (3). Superintendents perceived their

evaluation systems were effective in meaningfully differentiating performance using at least three performance levels.

RQ4. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels affected by the size of the district?

H4. Superintendents' perceptions of the effectiveness of their districts' evaluation systems meaningfully differentiating at least three performance levels are affected by the size of the district.

A one-factor ANOVA was conducted to test H4. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in meaningfully differentiating performance using at least three performance levels was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = .597$, $df = 5$, 120 , $p = .703$. See Table 3 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels were not affected by the size of the district.

Table 4

Descriptive Statistics for the Results of the Test for H4

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	4.160	0.624
2 (300-599)	38	4.053	0.868
3 (600-899)	21	4.286	0.956
4 (900-1,999)	18	4.389	0.502
5 (2,000-4,999)	17	4.294	0.588
6 (5,000-22,000)	7	4.286	1.254

RQ5. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on student growth for all students in determining performance levels?

H5. Superintendents perceive that their districts' evaluation systems effectively includes, as a significant factor, data on student growth for all students in determining performance levels.

A one-sample *t* test was conducted to test H5. The average perception of the effectiveness of superintendents' districts' evaluation systems in including, as a significant factor, data on student growth for all students in determining performance levels was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 6.012$, $df = 127$, $p = .000$. The sample mean ($M = 3.508$, $SD = .956$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively included, as a significant factor, data on student growth for all students in determining performance levels.

RQ6. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in including student growth data for all students as a significant factor in determining performance levels affected by the size of the district?

H6. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including student growth for all students as a significant factor in determining performance levels are affected by the size of the district.

A one-factor ANOVA was conducted to test H6. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in including student growth for all students as a significant factor in determining performance levels was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = .521$, $df = 5, 120$, $p = .760$. See Table 4 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including student growth for all students as a significant factor in determining performance levels were not affected by the size of the district.

Table 5

Descriptive Statistics for the Results of the Test for H6

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.400	0.764
2 (300-599)	38	3.500	0.980
3 (600-899)	21	3.524	0.928
4 (900-1,999)	18	3.333	1.029
5 (2,000-4,999)	17	3.706	1.047
6 (5,000-22,000)	7	3.857	1.069

RQ7. What are superintendent's perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice?

H7. Superintendents perceive that their districts' evaluation systems effectively includes other measures of professional practice.

A one-sample *t* test was conducted to test H7. The average perception of the effectiveness of superintendents' districts' evaluation systems effectively includes other measures of professional practice was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 15.838$, $df = 126$, $p = .000$. The sample mean ($M = 4.031$, $SD = .734$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively include other measures of professional practice.

RQ8. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice affected by the size of the district?

H8. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including other measures of professional practice are affected by the size of the district.

A one-factor ANOVA was conducted to test H8. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in including other measures of professional practice was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.389$, $df = 5, 119$, $p = .233$. See Table 5 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in including other measures of professional practice were not affected by the size of the district.

Table 6

Descriptive Statistics for the Results of the Test for H8

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.960	0.611
2 (300-599)	38	3.816	0.955
3 (600-899)	21	4.190	0.602
4 (900-1,999)	17	4.294	0.470
5 (2,000-4,999)	17	4.059	0.429
6 (5,000-22,000)	7	4.143	1.069

RQ9. What are superintendents' perceptions of their districts' evaluation systems effectiveness in supporting the effective evaluation of educators on a regular basis?

H9. Superintendents perceive that their districts' evaluation systems effectively support the evaluation of educators on a regular basis.

A one-sample t test was conducted to test H9. The average perception that superintendents' districts' evaluation systems effectively support the evaluation of educators on a regular basis was tested against a null value of 3. The results of the one-sample t test indicated a statistically significant difference between the two values, $t = 19.033$, $df = 127$, $p = .000$. The sample mean ($M = 4.195$, $SD = .711$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively support the evaluation of educators on a regular basis.

RQ10. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively evaluating educators on a regular basis affected by the size of the district?

H10. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively evaluating educators on a regular basis are affected by the size of the district.

A one-factor ANOVA was conducted to test H10. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in evaluating educators on a regular basis was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.448$, $df = 5, 120$, $p = .212$. See Table 6 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in evaluating educators on a regular basis were not affected by the size of the district.

Table 7

Descriptive Statistics for the Results of the Test for H10

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.960	0.790
2 (300-599)	38	4.237	0.751
3 (600-899)	21	4.333	0.577
4 (900-1,999)	18	4.278	0.575
5 (2,000-4,999)	17	4.000	0.791
6 (5,000-22,000)	7	4.571	0.535

RQ11. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback?

H11. Superintendents perceive that their districts' evaluation systems effectively provides clear feedback.

A one-sample *t* test was conducted to test H11. The average perception that superintendents' districts' evaluation systems effectively provides clear feedback was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 14.336$, $df = 127$, $p = .000$. The sample mean ($M = 3.930$, $SD = .734$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively provides clear feedback.

RQ12. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback affected by the size of the district?

H12. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback are affected by the size of the district.

A one-factor ANOVA was conducted to test H12. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in effectively providing clear feedback was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.584$, $df = 5, 120$, $p = .170$. See Table 7 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in effectively providing clear feedback were not affected by the size of the district.

Table 8

Descriptive Statistics for the Results of the Test for H12

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.680	0.748
2 (300-599)	38	3.842	0.789
3 (600-899)	21	4.238	0.539
4 (900-1,999)	18	3.944	0.639
5 (2,000-4,999)	17	3.941	0.659
6 (5,000-22,000)	7	4.143	1.069

RQ13. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing timely feedback?

H13. Superintendents perceive that their districts' evaluation systems effectively provides timely feedback.

A one-sample t test was conducted to test H13. The average perception that superintendents' districts' evaluation systems effectively provides timely feedback was tested against a null value of 3. The results of the one-sample t test indicated a statistically significant difference between the two values, $t = 15.452$, $df = 126$, $p = .000$. The sample mean ($M = 4.055$, $SD = .770$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively provides timely feedback.

RQ14. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing timely feedback affected by the size of the district?

H14. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing timely feedback are affected by the size of the district.

A one-factor ANOVA was conducted to test H14. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in effectively providing timely feedback was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.623$, $df = 5, 119$, $p = .159$. See Table 8 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in effectively providing timely feedback were not affected by the size of the district.

Table 9

Descriptive Statistics for the Results of the Test for H14

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.960	0.611
2 (300-599)	38	3.842	0.945
3 (600-899)	20	4.350	0.587
4 (900-1,999)	18	4.278	0.575
5 (2,000-4,999)	17	4.059	0.659
6 (5,000-22,000)	7	4.143	1.069

RQ15. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing useful feedback?

H15. Superintendents perceive that their districts' evaluation systems effectively provides useful feedback.

A one-sample *t* test was conducted to test H15. The average perception that superintendents' districts' evaluation systems effectively provides useful feedback was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 16.257$, $df = 125$, $p = .000$. The sample mean ($M = 3.968$, $SD = .669$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively provides useful feedback.

RQ16. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in effectively providing useful feedback affected by the size of the district?

H16. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing useful feedback are affected by the size of the district.

A one-factor ANOVA was conducted to test H16. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in effectively providing useful feedback was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = .850$, $df = 5, 118$, $p = .517$. See Table 9 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in effectively providing useful feedback were not affected by the size of the district.

Table 10

Descriptive Statistics for the Results of the Test for H16

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.960	0.539
2 (300-599)	37	3.838	0.800
3 (600-899)	20	4.100	0.447
4 (900-1,999)	18	4.111	0.471
5 (2,000-4,999)	17	3.824	0.728
6 (5,000-22,000)	7	4.143	1.069

RQ17. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in including feedback that identifies needs?

H17. Superintendents perceive that their districts' evaluation systems effectively includes feedback that identifies needs.

A one-sample t test was conducted to test H17. The average perception that superintendents' districts' evaluation systems effectively includes feedback that identifies needs was tested against a null value of 3. The results of the one-sample t test indicated a statistically significant difference between the two values, $t = 19.048$, $df = 127$, $p = .000$. The sample mean ($M = 4.023$, $SD = .608$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively includes feedback that identifies needs.

RQ18. To what extent are superintendents' perceptions of the effectiveness of the districts' evaluation systems in including feedback that identifies needs affected by the size of the school district?

H18. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively including feedback that identified needs are affected by the size of the district.

A one-factor ANOVA was conducted to test H18. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in effectively including feedback that identifies needs was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = .939$, $df = 5, 120$, $p = .458$. See Table 10 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in

effectively including feedback that identifies needs were not affected by the size of the district.

Table 11

Descriptive Statistics for the Results of the Test for H18

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.880	0.666
2 (300-599)	38	3.974	0.677
3 (600-899)	21	4.143	0.478
4 (900-1,999)	18	4.222	0.428
5 (2,000-4,999)	17	3.941	0.429
6 (5,000-22,000)	7	4.143	1.069

RQ19. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing feedback that guides professional development?

H19. Superintendents perceive that their districts' evaluation systems effectively provides feedback that guides professional development.

A one-sample *t* test was conducted to test H19. The average perception that superintendents' districts' evaluation systems effectively includes feedback that guides professional development was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 11.805$, $df = 126$, $p = .000$. The sample mean ($M = 3.772$, $SD = .737$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively includes feedback that guides professional development.

RQ 20. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in guiding professional development affected by the size of the school district?

H20. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively guiding professional development are affected by the size of the district.

A one-factor ANOVA was conducted to test H20. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in guiding professional development was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.199$, $df = 5, 119$, $p = .314$. See Table 11 for the means and standard deviations for this analysis. A follow up post hoc was not warranted. Superintendents' perceptions of their districts' evaluation systems in guiding professional development were not affected by the size of the district.

Table 12

Descriptive Statistics for the Results of the Test for H20

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.760	0.663
2 (300-599)	38	3.632	0.751
3 (600-899)	21	4.048	0.590
4 (900-1,999)	17	3.882	0.697
5 (2,000-4,999)	17	3.588	0.795
6 (5,000-22,000)	7	3.857	1.069

RQ21. What are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions?

H21. Superintendents perceive that their districts' evaluation systems effectively informs personnel decisions.

A one-sample *t* test was conducted to test H21. The average perception that superintendents' districts' evaluation systems effectively informs personnel decisions was tested against a null value of 3. The results of the one-sample *t* test indicated a statistically significant difference between the two values, $t = 12.803$, $df = 127$, $p = .000$. The sample mean ($M = 3.813$, $SD = .761$) was higher than the null value (3). Superintendents perceived their evaluation systems effectively inform personnel decisions.

RQ22. To what extent are superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions affected by the size of the district?

H22. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions are affected by the size of the district.

A one-factor ANOVA was conducted to test H22. The categorical variable used to group the average perception of the effectiveness of superintendents' districts' evaluation systems in informing personnel decisions was the size of the district. The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.370$, $df = 5, 120$, $p = .240$. See Table 12 for the means and standard deviations for this analysis. A follow up post hoc was not warranted.

Superintendents' perceptions of their districts' evaluation systems in informing personnel decisions were not affected by the size of the district.

Table 13

Descriptive Statistics for the Results of the Test for H22

Category	<i>N</i>	<i>M</i>	<i>SD</i>
1 (68-299)	25	3.880	0.526
2 (300-599)	38	3.711	0.768
3 (600-899)	21	3.952	0.740
4 (900-1,999)	18	3.833	0.618
5 (2,000-4,999)	17	3.529	1.068
6 (5,000-22,000)	7	4.286	0.756

In summary, 11 one-sample *t* tests were conducted to determine Kansas superintendents' perceptions of their teacher evaluation systems in meeting the six guidelines of the ESEA waiver. Additionally, 11 one-factor ANOVAs were conducted to determine any relationship between superintendents' perceptions and the size of their respective school district. The findings indicate that superintendents perceive their evaluation systems meet the guidelines as outlined in the ESEA waiver. There were no relationships found between the perceptions held and the size of the school district.

Summary

Chapter four included the descriptive statistics and hypothesis testing related to Kansas superintendents' perceptions of the effectiveness of their teacher evaluation systems in meeting the six guidelines of the ESEA waiver. Likewise, results were shared regarding any relationships between those perceptions and the school district's

enrollment (size). The results of conducting the one- sample t tests and ANOVAs were provided. Chapter five includes a summary of the study, major findings, findings related to the literature, implications for action and recommendations for future research.

Chapter Five

Interpretation and Recommendations

Chapter five contains a summary of the study, including an overview of the problem, purpose statement, and research questions. In addition, a review of the methodology and major findings are also provided. These findings and their relation to the literature are then expanded upon as well. The chapter concludes with implications for action, as well as recommendations for future research and concluding remarks.

Study Summary

The following section includes a synopsis of key components to the study. An overview of the problem, which focused on the perceptions of Kansas superintendents' perceptions of the effectiveness of their districts' teacher evaluation systems in meeting the ESEA waiver guidelines, is shared. In addition, the purpose statement and research questions are also reviewed. The section concludes with a review of the methodology applied in answering the study's twenty-two research questions and the major findings.

Overview of the problem. Ensuring quality teachers are leading every classroom has been a long sought goal in American classrooms; however, research now supports the theory that teachers are one of the single most important factors in the academic and future success of our students (Chetty et al., 2012). With this in mind, further research has been conducted to identify the characteristics of these effective teachers and the best means by which to identify, support and retain them (Weisberg et al., 2009). As a result, the importance and role of evaluating teachers and the tools used in doing so have undergone much scrutiny. This response has been further supported by the adoption of the Race to the Top program instituted by President Obama's administration in 2011

(Hull, 2013; McGuinn, 2011; Pennington, 2014). Hull (2013) noted that in response to this program, many states took advantage of waivers to redesign their evaluation systems. The state of Kansas submitted a waiver request, which included specific provisions and expectations for the evaluation of educators (KSDE, 2015). As a result, many districts throughout the state revised and adopted a new teacher evaluation system. Much research has been conducted into teacher and building administrators' perceptions of the effectiveness of evaluation systems in other states (Bridich, 2013; Clark, 2014; Coulter, 2013; Doherty, 2009; Heyde, 2013; Killian, 2010; Paufler, 2014; Ruffini et al., 2014; Sheppard, 2013; Winslow, 2015); however, there is no published research regarding the evaluation systems utilized throughout the state of Kansas. Similarly, little research (Derrington, 2014) has been conducted on the perceptions and input of district superintendents in the implementation and effectiveness of these teacher evaluation systems. Knowing the superintendent is the leader of the school district and as such is a key player in the selection and implementation of teacher evaluations, the perceptions of this leader regarding the effectiveness of the district's evaluation system is insightful. This study was conducted to gather superintendents' perceptions of the effectiveness of their districts' evaluation systems in meeting the ESEA waiver guidelines, and to fill the void in the research.

Purpose statement and research questions. The first purpose of the study was to determine superintendents' perceptions of their district's evaluation systems in meeting each of the six guidelines of Principle Three of the ESEA waiver. While there were six guidelines identified, many of the guidelines included multiple expectations. Thus, eleven research questions were designed to address these expectations. The second

purpose of the study was to determine any relationship that may have existed between the superintendents' perceptions of the effectiveness of their district's evaluation systems in meeting each of the eleven expectations included in the six components of Principle Three of the ESEA waiver and the size of the school district. Eleven research questions were designed to address the second purpose. Therefore, 22 research questions were posed in this study.

Review of the methodology. All K-12 Kansas public school superintendents were included as part of the sample in this study. A perceptive survey tool was developed based on the eleven expectations within the six components of Principle Three of the ESEA waiver. There were eleven questions surrounding each of these expectations included on the survey instrument. A quantitative Likert-type survey scale was utilized in gathering these perceptions. A twelfth, open response question, inquired as to the Pre-K-12 enrollment of the district. SurveyMonkey was then used to compile data, which was subsequently downloaded and imported to IBM® SPSS® Statistics Faculty Pack 23 for Windows for analysis. One-sample *t* tests and one-factor ANOVAs were used to test the hypotheses in this study.

Major findings. Upon receiving 128 responses from the 286 Kansas superintendents surveyed, the statistical analysis was conducted. The results of the analysis indicate Kansas superintendents' perceived their teacher evaluation systems effectively address the parameters of the six components of the ESEA waiver. Likewise, survey responses and analysis of these results indicated the size of the school district did not affect perceptions that the evaluation system was effective in meeting the components of the ESEA waiver.

RQ1 and RQ2 focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction and whether these perceptions were affected by the size of the school district.

Superintendents perceived their evaluation systems were effective in supporting continual improvement of instruction. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction were not affected by the size of the district.

RQ 3 and RQ 4 focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels and whether these responses were affected by the size of the district. Superintendents perceived their evaluation systems were effective in meaningfully differentiating performance using at least three performance levels. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in meaningfully differentiating performance using at least three performance levels were not affected by the size of the district.

RQ 5 and RQ 6 were focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on student growth for all students in determining performance levels and whether these perceptions were affected by the size of the school district. Superintendents perceived their evaluation systems effectively uses multiple measures including, as a significant factor, data on student growth for all students in determining performance levels. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in using multiple measures, including, as a significant factor, data on

student growth for all students in determining performance levels were not affected by the size of the district.

RQ7 and RQ8 focused on superintendent's perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice and whether these perceptions were affected by the size of the district.

Superintendents perceived their evaluation systems effectively include other measures of professional practice. Superintendent's perceptions of the effectiveness of their districts' evaluation systems in effectively including other measures of professional practice were not affected by the size of the district.

RQ9 and RQ10 looked at superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting the effective evaluation of educators on a regular basis and whether these perceptions were affected by the size of the district.

Superintendents perceived their evaluation systems effectively support the evaluation of educators on a regular basis. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in supporting the effective evaluation of educators on a regular basis were not affected by the size of the district.

RQ11 and RQ12 focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback and the effect of district size on these perceptions. Superintendents perceived their evaluation systems effectively provide clear feedback. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing clear feedback were not affected by the size of the district.

RQ13 and RQ14 looked at superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing timely feedback and whether these perceptions are affected by the size of the district. Superintendents perceived their evaluation systems effectively provides timely feedback. Superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing timely feedback were not affected by the size of the district.

RQ15 and RQ16 focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively providing useful feedback and whether the size of the district affected these perceptions. Superintendents perceived their districts' evaluation systems effectively provides useful feedback. Superintendents' perceptions of their districts' evaluation systems in effectively providing useful feedback were not affected by the size of the district.

RQ17 and RQ18 looked at superintendents' perceptions of the effectiveness of their districts' evaluation systems in including feedback that identifies needs and whether these perceptions were affected by the size of the district. Superintendents perceived their districts' evaluation systems effectively includes feedback that identifies needs. Superintendents' perceptions of their districts' evaluation systems in effectively including feedback that identifies needs were not affected by the size of the district.

RQ19 and RQ20 focused on superintendents' perceptions of the effectiveness of their districts' evaluation systems in providing feedback that guides professional development. Superintendents perceived their districts' evaluation systems effectively provides feedback that guides professional development. Superintendents' perceptions of

their districts' evaluation systems in guiding professional development were not affected by the size of the district.

RQ21 and RQ22 looked at superintendents' perceptions of the effectiveness of their districts' evaluation systems in effectively informing personnel decisions and whether these perceptions were affected by the size of the district. Superintendents perceived their districts' evaluation systems effectively informs personnel decisions. Superintendents' perceptions of their districts' evaluation systems in informing personnel decisions were not affected by the size of the district.

In summary, this study included 22 research questions addressing Kansas Superintendents perceptions of the effectiveness of their teacher evaluation systems. The study found that Kansas superintendents' perceive their teacher evaluation systems were effective in meeting each of the six components of Principle Three of the Kansas ESEA waiver. Likewise, these perceptions were not affected by the size of the district.

Findings Related to the Literature

Little research exists regarding the Kansas ESEA waiver requirements and subsequent changes made to teacher evaluations and their effectiveness throughout the state. There is a breadth of research and study into the perceptions held by teachers and building administrators surrounding teacher evaluation in various states, including Arizona, Colorado, Illinois, Massachusetts, Missouri, Nebraska, Pennsylvania and Washington (Bridich, 2013; Clark, 2014; Coulter, 2013; Doherty, 2009; Heyde, 2013; Killian, 2010; Paufler, 2014; Ruffini et al., 2014; Sheppard, 2013; Winslow, 2015). However, research into superintendents' perceptions of teacher evaluation systems and their effectiveness in meeting the new standards imposed through the Rise to the Top

program, was virtually non-existent. As Bridich (2013) specifically referenced, there had not been research on superintendents' or school boards' perceptions of teacher evaluations. In addition, Bridich (2013) also noted that this is an area needing to be studied further as these leaders play a key role in the selection and implementation of teacher evaluations. This hole in the body of research was the impetus for this study into Kansas superintendents' perceptions of the effectiveness of their districts' evaluation systems in meeting the ESEA waiver guidelines.

Doherty's (2009) work was focused on the perceptions of teachers and administrators in a suburban Massachusetts school district that had recently adopted a standards-based teacher evaluation system. The results of the study showed agreement amongst the principals regarding their perception of the standards-based tool effectively being used to support teacher growth and relegating its use as a tool to make personnel decisions as a last resort. Likewise, the work of Coulter (2013) in studying the implementation of the Washington state teacher evaluation system also found that building principals perceived the new evaluation system focused on the professional growth of teachers. Another source of research, which supported this perception, was that of Clark (2014), who found Nebraska principals also perceived their teacher evaluation system was effective in supporting teacher growth. These findings were supported in the current study as Kansas superintendents also found their districts' evaluation systems effectively supported the improvement of instruction while also informing personnel decisions.

Previous research into the use of student growth/achievement as a factor in teacher evaluation has revealed differing opinions by school administrators. Bridich's

(2013) work referenced the recent implementation of Colorado SB 191, which mandated changes to teacher evaluations throughout the state. Results of his study specifically noted that most building administrators (72%) supported the inclusion of student assessment results as a component of teacher evaluation. Conversely, the work of Coulter (2013) found that all administrators sampled in Washington, disagreed with the mandate that student growth data must be included in teacher evaluations. However, this was a mandate included in the passing of Washington SB 6696. Although the work of Bridich (2013) and Coulter (2013) utilized principals and assistant principals as part of the study's sample, there was a connection to the present research. While the current study did not ask whether superintendents agreed with the inclusion of student growth data in teacher evaluations, it did find that superintendents found their districts teacher evaluation systems effectively included student growth measures for all students.

Of particular interest and relativity to the current study is the work of Derrington (2014), which was the only study to sample school superintendents and their perceptions of the implementation of a new teacher evaluation system in a southeastern state. While the study focused on the principal and superintendent's role in the implementation process, as well as the benefits and concerns in this endeavor, connections can also be drawn between this research and the current study. Derrington (2014) found that all superintendents and principals identified an increase in principals' awareness and understanding of good instruction. Specifically, Derrington (2014) notes "Principals and superintendents unanimously agreed that more conversation about recognizing good instruction occurred both in the schools and at the district level as a result of the new rubric" (p 127). While the current study did not specifically inquire as to the

implementation process, Derrington's (2014) findings support the value and importance of studying superintendents' perceptions. A key factor in utilizing superintendents in the current study was the belief that they play an important role in supporting building leaders. Similarly, the importance of including superintendents' perceptions of their districts' teacher evaluation systems is critical due to the need for their leadership and support in the implementation process as further noted in the work of Derrington (2014). More specifically, Derrington (2014) noted unanimous support for the need for superintendent support in the implementation of a new teacher evaluation system. A final connection to Derrington's (2014) research and the current study can also be made as it relates to the implementation of state policy. The current study found that all superintendents perceived their districts' evaluation systems were effective in meeting the six components of Principle Three of the Kansas ESEA waiver. Conversely, Derrington (2014) noted, "Inconsistent implementation and different interpretations of state policy between districts were two other unintended consequences that all 18 interviewees reported" (p. 131). This finding was not supported in the current study, as there was a strong level of agreement with each of the eleven expectations included as part of Principle Three of the Kansas ESEA waiver.

Conclusions

This section includes conclusions drawn from the current study that can be utilized to improve teacher evaluation systems in meeting the eleven expectations included in Principle Three of the Kansas ESEA waiver. To that end, implications for further action are included, as well as recommendations for future research. Concluding remarks are also provided.

Implications for action. The results of this study provide many potential areas for future action, particularly for Kansas superintendents and policymakers. Although superintendents perceived their current evaluation systems are meeting the eleven expectations included in the six components of Principle Three of the Kansas ESEA waiver, they may now take this research and conduct further investigation into the implementation phase of this system. Given the positive perceptions by superintendents that the evaluation systems were meeting the ESEA waiver criteria, the focus can now surround ensuring the system is implemented properly and updated as changes to the waiver are made. Furthermore, superintendents can also begin to investigate and ensure the teacher evaluation system is supported by strong inter-rater reliability through continued support and professional development with all evaluators.

The study further noted no significant difference between the perceived effectiveness of the district's evaluation system and the size of the district. Thus, superintendents can now collaborate and share, across the state, their specific system with other leaders as a means to improve and refine all teacher evaluation systems. With the focus now shifted to perfecting the use of the evaluation tool, this cross-district collaboration could be useful to perfecting and ensuring the authentic evaluation of all teachers.

The finding that districts' teacher evaluation systems are perceived to meet the ESEA waiver criteria, and that this perception is not affected by the size of the district, should also provide policy makers peace of mind. While much discussion surrounds the perceived inequality inherent in the great disparity in district enrollments throughout the state of Kansas, this study further supports the notion that no matter the size of the

district, the teacher evaluation system is perceived to be meeting policy makers' expectations. While there is already discussion regarding the need for further consolidation of Kansas school districts, this study would support lawmakers continued focus on other areas of school reform and finance.

Recommendations for future research. This study was intended to fill a void in the research regarding teacher evaluation systems and the perception of these systems meeting expectations outlined through the Race to the Top program. While the current study found Kansas superintendents perceived their districts' evaluation systems met all eleven expectations included in the six components of Principle Three of the ESEA waiver, there is a need for further research.

Specifically, the current study only included quantitative research; thus, further qualitative study could be conducted. This qualitative data could focus on the basis for superintendent's perceptions. An assumption made in this study was that the superintendent had sufficient understanding of their district's current teacher evaluation system. Thus, further qualitative research would provide an opportunity to ensure this understanding and delve further into the specific reasons the systems meet the ESEA waiver criteria.

A comparison study could also be conducted, utilizing the same survey tool with the elimination the size of district factor. The population sample could then be expanded to include Kansas teachers, principals, and superintendents. By using the same survey tool, direct comparisons could be made among the responses from teachers, principals, and superintendents in an effort to better identify differences in perceptions.

Another area for continued research is in response to potential changes lawmakers are proposing regarding Kansas teacher evaluation systems. While the research for this study was compiled during the 2015-2016 school year and based on the ESEA waiver guidelines as of July of 2015, further changes have been proposed to the ESEA waiver. In response to these proposed changes, it may be useful to study district's response to this updated expectation.

A final suggestion for future research concerns the sample population. While the current study utilized only Kansas superintendents' perceptions, it may be useful to expand the study to include other district leaders. Knowing many districts have an identified Human Resources Directors and/or Instructional Leadership Directors, these individuals may have more "hands-on" experience with the teacher evaluation system. Likewise, expanding the sample group may also increase the number of respondents to the survey.

Concluding remarks. Teacher evaluation has been a hot topic in school reform since the institution of the Race to the Top program. As Weisberg et al. (2009) found "In the absence of policy systems based on instructional effectiveness, districts make decisions about teachers in other ways" (p. 24). The approval of the Kansas ESEA waiver application and its inclusion of expectations for teacher evaluation provided just such policy provisions. With the implementation of these expectations throughout Kansas, the current study provides insight currently lacking in the research.

Knowing the ESEA waiver provides six specific components regarding the expectation for teacher evaluations, the focus of this study provides insight into the current perceptions held by superintendents in meeting these expectations. Results

indicate perceived success in the development and implementation of a teacher evaluation system that meets these requirements. Furthermore, the study found shared support for meeting these expectations across all responding districts, no matter the district size. Now having this insight, district leaders and lawmakers can focus their efforts on ensuring the effectiveness of teacher evaluation system in truly ensuring quality teachers are leading every classroom.

References

- Bolin, F. S., & Panaritis, P. (1992). Searching for a common purpose: A perspective on the history of supervision. In C. Glickman (Eds.), *Supervision in transition: 1992 yearbook for the Association for Supervision and Curriculum Development* (pp. 30-44). Alexandria, VA: Association for Supervision and Curriculum Development.
- Bridich, S. M. (2013). *Perceptions surrounding the implementation of Colorado senate Bill 10-191's new teacher evaluations* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3588313)
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2012). *Great teaching: Measuring its effects on students' future earnings*. Retrieved from <http://educationnext.org/great-teaching/>
- Clark, T. K. (2014). *Principals' perceptions concerning teacher evaluation in Nebraska public schools* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3666954)
- 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). Available from ProQuest Dissertations and Theses database. (UMI No. 3587064)
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publication.

- Danielson, C. (2002). *Enhancing student achievement: A framework for school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Retrieved from http://depts.washington.edu/ctpmail/PDFs/LDH_1999.pdf
- Derrington, M. L. (2014). Teacher evaluation initial policy implementation: Principal and superintendent perceptions. *Planning and Changing*, 45(1/2), 120-137.
- Doherty, J. F. (2009). *Perceptions of teachers and administrators in a Massachusetts suburban school district regarding the implementation of a standards-based teacher evaluation system* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3431766)
- Hadfield, T. E., Hutchinson-Lupardus, T. R., & Snyder, J. E. (2012). *Assessing state models of value-added teacher evaluations: Alignment of policy, instruments, and literature-based concepts* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3516314)

- Hanushek, E. (2012). *Low performing teachers have high costs*. Retrieved from http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202012%20Education%20Next_Low-Performing%20Teachers%20Have%20High%20Costs.pdf
- Heyde, C. R. (2013). *Teacher and administrator perceptions of the effectiveness of current teacher evaluation practices and the impact of the new Illinois performance evaluation reform act of 2010* (Doctoral dissertation). Retrieved from <http://digitalcommons.ni.edu/cgi/viewcontent.cgi?article=1002&context=pe>
- Hull, J. (2011). *Building a better evaluation system: Full report*. Retrieved from <http://www.centerforpubliceducation.org/Main-Menu/Staffingstudents/Building-A-Better-Evaluation-System/Building-A-Better-Evaluation-System.html>
- Hull, J. (2013). *Trends in teacher evaluation: How states are measuring teacher performance*. Retrieved from <http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/Trends-in-Teacher-Evaluation-At-A-Glance/Trends-in-Teacher-Evaluation-Full-Report-PDF.pdf>
- Kansas State Department of Education (2014). *Kansas ESEA flexibility waiver fact sheet 2014-2015*. Retrieved from <http://www.ksde.org/Portals/0/ECSETS/FactSheets/FactSheet-Waiver-WaiverOverview.pdf>
- Kansas State Department of Education (2015). *ESEA flexibility renewal form Kansas*. Retrieved from <http://www.ksde.org/Portals/0/Title%20Programs%20and%20Services/ESEAWaiver/20150611-ESEA-FlexRequest.pdf>

- Killian, N. (2010). *Administrators' and teachers' perceptions of the efficacy of the Missouri performance-based teacher evaluation model*. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3419724)
- Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social behavioral sciences*. Thousand Oaks, CA: Corwin Press
- Marzano, R. J., Frontier, T., & Livingston, D., (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD
- McGuinn, P. (2012). *The state of teacher evaluation reform*. Retrieved from <https://www.americanprogress.org/issues/education/report/2012/11/13/44494/the-state-of-teacher-evaluation-reform/>
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002). Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>
- Paufler, N. (2014). *Investigating a teacher evaluation system: School administrator and teacher perceptions of the system's standards of effectiveness*. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3646906)
- Pennington, K. (2014). *ESEA waivers and teacher-evaluation plans: State oversight of district-designed teacher-evaluation systems*. Retrieved from <https://www.americanprogress.org/issues/education/report/2014/05/07/89121/esea-waivers-and-teacher-evaluation-plans/>
- Race to the Top Executive Summary. (2009). Retrieved from <https://www2.ed.gov/programs/racetothetop/executive-summary.pdf>

- Reform Support Network. (2011). *Great teachers and leaders: State considerations on building systems of educator effectiveness*. Retrieved from <https://www2.ed.gov/programs/racetothetop/great-teachers.doc+&cd=1&hl=en&ct=clnk&gl=us>
- Roberts, C. M. (2010). *Dissertation journey: A practical and comprehensive guide to planning, writing, and defending our dissertation*, (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Ruffini, S., Makkonen, R., Tejwani, J., & Diaz, M. (2014). *Principal and teacher Perceptions of implementation of multiple-measure teacher evaluation systems in Arizona*. Washington, DC: U.S. Department of Education.
- Sheppard, J. (2013). *Perceptions of teachers and administrators regarding the teacher evaluation process* (Doctoral dissertation). Retrieved from <http://digitalcommons.georgiasouthern.edu/cgi/viewcontent.cgi?article=1856&context=etd>. Paper 852
- U.S. Department of Education. (2012). *ESEA flexibility*. Retrieved from <https://www.ed.gov/sites/default/files/esea-flexibility.doc>
- U.S. Department of Education, Elementary and Secondary Education Act. (2015). *History of ESEA*. Retrieved from <http://www.ed.gov/esea>
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Retrieved from <http://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness>
- Williams, J. (2009). *McRELS teacher evaluation system*. Retrieved from <http://www.medford.k12.nj.us/cms/lib07/NJ01001377/Centricity/Domain/3/McRELTeacher%20Evaluation%20Users%20Guide.pdf>

Winslow, R.A. (2015). *Administrative feedback following classroom observations as part of a Danielson-Based teacher evaluation system: Teacher and administrator perceptions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No: 3705399)

Appendices

Appendix A: Teacher Evaluation Survey

Teacher Evaluation

As part of this study I am investigating Kansas Superintendents' perceptions that their current teacher evaluation system meets the six guidelines included as Principle 3 of the Elementary and Secondary Education Act Waiver (ESEA).

The entire survey should take no longer than 10 minutes for you to complete. Your answers will be kept anonymous as no district identifying information will be asked or gathered through the survey. All responses will be kept confidential and combined with responses of other participants in summary form. Therefore, the results and information reported will not include any individual or school district specific responses. The completion of the survey will indicate your consent to participate and permission to use the information provided by you in my research study. You have the option to not answer any question included on the survey that causes you concern. Likewise, you may discontinue participation at any point during the survey.

Thank you in advance for your time and participation in the study. I sincerely appreciate your willingness to support this work. Please do not hesitate to let me know if you have any questions or concerns regarding the survey or if you would like a copy of the results. I can be reached any time at lachellesigg@gmail.com or you are welcome to call me personally at (913)957-0423.

Please answer each of the questions to the best of your ability and thank you in advance for your time, it is sincerely appreciated!

Teacher Evaluation

With the approval of the Elementary and Secondary Education Act (ESEA) Waiver by the Federal Department of Education, Kansas has now replaced No Child Left Behind with the four principles included in this waiver. The goal of the third principle is to ensure that all school districts implement evaluation instruments that will support systems that:

Will be used for continual improvement of instruction;

Meaningfully differentiate performance using at least three performance levels;

Use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students, and other measures of professional practice;

Evaluate educators on a regular basis;

Provide clear, timely and useful feedback, including feedback that identifies needs and guides professional development;

Bearing in mind these six components to Principle 3 of the ESEA Waiver for teacher evaluation systems, please respond with your level of agreement to the following statements:

1. My district's evaluation system supports continual improvement of instruction

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. My district's evaluation system meaningfully differentiates performance using at least three performance levels.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. My district's evaluation system includes, as a significant factor, data on student growth for all students in determining performance levels.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. My district's evaluation system effectively includes other measures of professional practice.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. My district's evaluation tool effectively supports the evaluation of educators on a regular basis.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. My district's evaluation system effectively provides clear feedback.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. My district's evaluation system effectively provides timely feedback.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. My district's evaluation system effectively provides useful feedback.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. My district's evaluation system includes feedback that identifies needs.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. My district's evaluation system includes feedback that guides professional development.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. My district's evaluation system is effectively used to inform personnel decisions.

Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12. What was the 2015 Student Enrollment (Pre-K-12 FTE) for your school district?

Appendix B: Institutional Review Board (IRB) Request



SCHOOL OF EDUCATION
GRADUATE DEPARTMENT

IRB PROTOCOL NUMBER

Date:

(IRB USE ONLY)

IRB REQUEST

Proposal for Research

Submitted to the Baker University Institutional Review Board

I. Research Investigator(s) (Students must list faculty sponsor first)

Department(s) School of Education Graduate Department

Name	Signature	
1. Dr. Susan Rogers		Major Advisor
2. Margaret Waterman		Research Analyst
3. Dr. Russ Kokoruda		University Committee Member
4. Dr. Doug Sumner		External Committee Member

Principal Investigator: Lachelle Sigg
 Phone: (913)957-0423
 Email: lachellesigg@gmail.com
 Mailing address: 11261 S. Race St.
 Olathe, KS 66061



Faculty sponsor: Dr. Susan Rogers
 Phone: (913)491-4432
 Email: Susan.Rogers@bakeru.edu

Expected Category of Review: Exempt Expedited Full

II: Protocol: (Type the title of your study)

Kansas Superintendents' Perceptions of their Teacher Evaluation in Meeting the Requirements of the Elementary and Secondary Education Act Flexibility Waiver

Summary

In a sentence or two, please describe the background and purpose of the research.

In preparation for the impending expiration of the No Child Left Behind Act, states across the country were forced to evaluate and address four specific areas outlined within the Elementary and Secondary Education Act. This study is focusing on the third component of the ESEA Flexibility Waiver, effective instruction and leadership. More specifically, the purpose of this study is to determine Kansas superintendents' perceptions regarding the effectiveness of their districts' evaluation systems in supporting continual improvement of instruction; meaningfully differentiating teacher performance using at least three performance levels; using multiple measures in determining teacher performance levels; including data on student growth for all students as a significant factor in determining performance levels; including other measures of professional practice; evaluating educators on a regular basis; providing clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and informing personnel decisions. An additional purpose of this study is to determine whether the demographic factor of district size affected superintendents' perceptions.

Briefly describe each condition or manipulation to be included within the study.

There will be no conditions and/or manipulations as part of this study.

What measures or observations will be taken in the study? If any questionnaire or other instruments are used, provide a brief description and attach a copy.

Will the subjects encounter the risk of psychological, social, physical or legal risk? If so, please describe the nature of the risk and any measures designed to mitigate that risk.

The Teacher Evaluation Survey will be used to gather the necessary responses for this study. The tool was created by the researcher with support and input from university advisors as well as previous Kansas superintendent reviews and insights. A copy of the survey has been attached.

Subjects will not encounter the risk of any psychological, social, physical or legal risk in taking the survey.

Will any stress to subjects be involved? If so, please describe.

No stress to subjects will be experienced as a result of his/her participation.

Will the subjects be deceived or misled in any way? If so, include an outline or script of the debriefing.

Subjects will not be deceived or misled in any way in this study.

Will there be a request for information which subjects might consider to be personal or sensitive? If so, please include a description.

No personal or sensitive information will be asked for in the survey.

Will the subjects be presented with materials which might be considered to be offensive, threatening, or degrading? If so, please describe.

No, subjects will not encounter any offensive, threatening, or degrading material by participating in the study.

Approximately how much time will be demanded of each subject?

The Teacher Evaluation Survey will take approximately 10 minutes for each subject to complete.

Who will be the subjects in this study? How will they be solicited or contacted? Provide an outline or script of the information which will be provided to subjects prior to their volunteering to participate. Include a copy of any written solicitation as well as an outline of any oral solicitation.

Kansas superintendents will be the subjects in the study. Each superintendent will be contacted via their publicly listed email address as provided in the Kansas State Department of Education Directory. A solicitation letter explaining the study and providing detail regarding data collection will be provided to each superintendent. A copy of this letter is also attached.

What steps will be taken to insure that each subject's participation is voluntary? What if any inducements will be offered to the subjects for their participation?

Each participant will be informed of the voluntary nature of the survey via the solicitation letter, which is attached. In addition, all responses will be collected via an electronic survey link.

How will you insure that the subjects give their consent prior to participating? Will a written consent form be used? If so, include the form. If not, explain why not.

The solicitation letter included with the survey will provide explanation of their consent upon completing the survey. Written consent will not be needed as completion of the survey will provide this assurance.

Will any aspect of the data be made a part of any permanent record that can be identified with the subject? If so, please explain the necessity.

All responses to the survey will be provided through the electronic survey system, SurveyMonkey, thus no personal identification of individual subjects can be made.

Will the fact that a subject did or did not participate in a specific experiment or study be made part of any permanent record available to a supervisor, teacher or employer? If so, explain.

No information regarding a subject's participation will be kept as any permanent record made available to any supervisor or employer.

What steps will be taken to insure the confidentiality of the data?

An on-line, anonymous survey will be utilized, and no identifiable information will be reported in the study.

Where will it be stored?

All data will be stored within SurveyMonkey, which is password protected.

How long will it be stored?

All data will be stored for a period of five years.

What will be done with it after the study is completed?

After five years, all data will be deleted.

If there are any risks involved in the study, are there any offsetting benefits that might accrue to either the subjects or society?

There are no risks involved in the study.

Will any data from files or archival data be used? If so, please describe.

Archival data will be utilized. Data from SurveyMonkey will be downloaded and imported into IBM® SPSS® Statistics Faculty Pack for Windows.

Appendix C: Institutional Review Board (IRB) Approval



Baker University Institutional Review Board

November 4, 2015

Dear Lachelle Sigg and Dr. Rogers,

The Baker University IRB has reviewed your research project application and approved this project under Exempt Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.

Please inform this Committee or myself when this project is terminated or completed. As noted above, you must also provide IRB with an annual status report and receive approval for maintaining your status. If you have any questions, please contact me at CTodden@BakerU.edu or 785.594.8440.

Sincerely,

Chris Todden EdD
Chair, Baker University IRB

Baker University IRB Committee
Verneda Edwards EdD
Sara Crump PhD
Erin Morris PhD
Scott Crenshaw

Appendix D: Teacher Evaluation Survey electronic mail message to superintendents

December 3, 2015

Teacher Evaluation Survey electronic email message to superintendents

Dear Superintendent,

I am currently a doctoral student at Baker University, working to complete my dissertation. As part of this study I am investigating Kansas Superintendents' perceptions of their current teacher evaluation system in meeting the six guidelines included as Principle 3 of the ESEA (Elementary and Secondary Education Act) Waiver. Included in this study are Superintendents identified by the Kansas State Department of Education. I kindly ask for your participation in a survey, which can be found by clicking on the following link: www.surveymonkey.com

Bearing in mind the value of your time, the entire survey should take no longer than 10 minutes for you to complete. While the majority of the survey is in a multiple-choice format, it will also inquire as to the total Pre-Kindergarten through Twelfth grade enrollment (FTE) in your school district, as reflected in your September 20th, 2015 enrollment count.

Please rest assured that your answers will be kept anonymous as no district identifying information will be asked or gathered through the survey. Rather all responses will be kept confidential and combined with responses of other participants in summary form. Therefore, the results and information reported will not include any individual or school district specific responses. The completion of the survey will indicate your consent to participate and permission to use the information provided by you in my research study. Lastly, please know that you also have the option to not answer any question(s) included on the survey that causes you concern. Likewise, you may discontinue participation at any point during the survey.

Thank you in advance for your time and participation in the study. I sincerely appreciate your willingness to support this work. Please do not hesitate to let me know if you have any questions or concerns regarding the survey or if you would like a copy of the results. I can be reached any time at lachellesigg@gmail.com or you are welcome to call me personally at (913)957-0423.

Sincerely,
Lachelle Sigg

Director of Certified Human Resources
USD 512 - Shawnee Mission School District
Overland Park, KS 66204