

TEACHER PERCEPTIONS OF CHANGE:
AN EVALUATION OF THE NETWORK FOR EDUCATOR EFFECTIVENESS
SYSTEM

A Dissertation Presented to
The Faculty of the Graduate School
University of Missouri-Columbia

In Partial Fulfillment of the
Requirements for the Degree

Doctor of Education

By
JENNIFER L. KATZIN

Dr. Robert L. Watson, Dissertation Chair

MAY 2014

The undersigned, appointed by the Dean of the Graduate School, have examined
the dissertation entitled

TEACHER PERCEPTIONS OF CHANGE:
AN EVALUATION OF THE NETWORK FOR EDUCATOR EFFECTIVENESS
SYSTEM

presented by Jennifer L. Katzin

a candidate for the degree of Doctor of Education

and hereby certify that in their opinion it is worth of acceptance.

Dr. Robert Watson

Dr. Cynthia MacGregor

Dr. Kim Finch

Dr. Beth Hurst

DEDICATION

This dissertation is dedicated to my wonderful family. Without their endless support, it would have never been completed. To Greg, thank you for your love, encouragement, prodding, and for picking up the slack. To Carly, thank you for your constant encouragement. Your comments and back rubs made my work more tolerable. To Cleo, thank you for your hugs. Your patience was a big help. No more papers guys...let's go to the pool!

ACKNOWLEDGMENTS

I would like to thank Dr. Robert Watson and Dr. Cindy MacGregor for the instruction and wisdom that they both shared over the course of the program. Your instruction taught me a lot about research, writing, and leadership.

Thank you to Dr. Kim Finch and Dr. Beth Hurst who served on my dissertation committee along with Dr. Watson and Dr. MacGregor. Their feedback was invaluable and provided the guidance needed to complete this project.

Also, thank you to Donna Aldrich and the other members of Cohort 8. Their friendship and accountability made the completion of the coursework and dissertation enjoyable. I learned a lot from each of them.

Lastly, thank you to God for meeting all my needs according to the riches of your glory in Christ Jesus (Philippians 4:19).

TABLE OF CONTENTS

DEDICATION.....	i
ACKNOWLEDGMENTS	ii
LIST OF TABLES	ix
LIST OF FIGURES	x
ABSTRACT	xi
CHAPTER	
1. INTRODUCTION TO THE STUDY	1
History of Missouri Teacher Evaluation	2
Problem Statement	7
Research Purpose	8
Research Questions	9
Conceptual Framework	9
Design and Methods	12
Assumptions	13
Definition of Key Terms	14
Significance of the Research for Leadership Practice	16
Summary	17
2. REVIEW OF LITERATURE	18
History of Teacher Evaluation	18
Importance of Teacher Evaluation and Their Effect on Student Achievement ...	23
Teacher Effects on Student Achievement	24
Teacher Evaluation Effects on Student Achievement	25

Barriers and Obstacles to Conducting Teacher Evaluations	28
Qualification/Quantity of Evaluators	28
Types of Data Used to Evaluate	29
Infrequency of Evaluations	30
Stress and Nervousness	31
“Dog and Pony Show”	32
Best Practices for Conducting Teacher Evaluations	34
Marzano Framework	34
Danielson Framework	35
Tucker and Stronge Framework	37
Classroom Observations and Feedback	38
Preparing Units of Instruction	44
Teacher Professional Development	46
Evaluation by Students	50
Marshall Framework	52
3. RESEARCH DESIGN AND METHODOLOGY	57
Purpose of the Study	58
Research Questions	58
Design for the Study	59
Participants and Sampling Procedures	60
Data Collection	62
Data collection procedures	63
Human subjects protection and ethical considerations	64

Data Analysis	65
Role of the Researcher	65
Trustworthiness	66
Limitations and Assumptions	67
Limitations	67
Assumptions	68
Summary	68
4. ANALYSIS OF DATA	70
Demographic Information	71
Survey Demographics	71
Focus Group Demographics	75
Findings	78
Research Question 1a: Mini-observations	78
Increased use of technology	79
Increase in student engagement	80
Increased use of assessment	81
Improvement based on specific administrator/system recommendations ..	81
Explanations regarding no changes made	82
Research Question 1b: Unit of Instruction Evaluation	84
Increased rigor	85
Greater collaboration	85
Positive impact on daily instruction	85
Explanations regarding no changes made	86

Research Question 1c: Professional Development Plan	87
Sharper Instructional focus	88
Implementation of learned skills	89
Teachers as learners	90
Explanations regarding no changes made	90
Research Question 1d: Student Survey Data	91
Positive comments regarding student survey information	92
Misunderstood data-driven decision making	93
Concerns about evaluations completed by children	93
Explanations regarding no changes made	93
General Thoughts about the Effectiveness of NEE	95
Teaching changed for the better	95
Good system format	96
Value in increased administrator presence	96
Subjective and arbitrary	96
The unattainable 7	97
Teacher stress/anxiety	99
Administrator stress/anxiety	99
Length of evaluations too short	100
Frequency of evaluations too high	100
Excessive busy work	101
Summary	102
5. CONCLUSIONS	104

Conclusions	106
Research Question 1a: Mini-observations	106
Research Question 1b: Unit of Instruction Evaluation	107
Research Question 1c: Professional Development Plan	109
Research Question 1d: Student Survey Data	111
Discussion	113
Implications and Recommendations for Practice	114
Research Question 1a: Mini-observations	115
Research Question 1b: Unit of Instruction Evaluation	116
Research Question 1c: Professional Development Plan	117
Research Question 1d: Student Survey Data	118
Overcoming Barriers	118
Subjectivity	118
The unattainable 7	119
Stress and anxiety	120
Limitations	121
Recommendations for Further Study	121
Summary	123
REFERENCES	124
APPENDICES	134
A. FOCUS GROUP PROTOCOL	134
B. INFORMED CONSENT – FOCUS GROUP	135
C. SURVEY	136

D. INFORMED CONSENT – SURVEY	138
Vita	140

LIST OF TABLES

Table	
1. Survey Responses by District	72
2. Years of Experience of Survey Participants	73
3. Level Taught by Survey Participants	74
4. Gender of Survey Participants	74
5. Focus Group Participants by District	75
6. Years of Experience of Focus Group Participants	76
7. Level taught by Focus Group Participants	77
8. Gender of Focus Group Participants	77
9. Likert Responses to Amount of Change as a Result of Mini-observations	79
10. Summary of Themes for Research Question 1a: Mini-observations	83
11. Likert Responses to Amount of Change as a Result of Unit of Instruction Evaluation	84
12. Summary of Themes for Research Question 1b: Unit of Instruction Evaluation	87
13. Likert Responses to Amount of Change as a Result of the Professional Development Plan	88
14. Summary of Themes for Research Question 1c: Professional Development Plan	91
15. Likert Responses to Amount of Change as a Result of Student Survey Data	92
16. Summary of Themes for Research Question 1d: Student Survey Data	94
17. Effectiveness of NEE	102
18. Summary of Conclusions	112

LIST OF FIGURES

Figure

1. Percent of change indicated based on implementation of mini-observations..... 107
2. Percent of change indicated based on implementation of unit of instruction
evaluations..... 108
3. Percent of change indicated based on implementation of professional
development plans 110
4. Percent of change indicated based on implementation of student survey data ... 111

ABSTRACT

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. The study gathered teachers' perceptions of the impact NEE had on their teaching, specifically noting changes in teaching practice. This information served as an evaluation of the program and will be used to guide its further development.

The study was guided by the following research questions. What are teachers' perceptions of change impacted by the Network for Educator Effectiveness based on the framework presented by Kim Marshall? Specifically, what changes have been made as a result of the mini-observations, unit of instruction evaluation, professional development plan, and student survey data?

Data for the study was collected through focus group interviews and surveys. The population included two small southwest school districts in Missouri that had implemented the NEE system. An interpretive analysis of data was done to make the following conclusions. The mini-observation piece has affected the most change in classroom instruction. The unit of instruction evaluation, professional development plan and student survey data have affected little change due to incomplete or improper implementation. It is recommended that the NEE system continue with further implementation to see a greater effect on change.

CHAPTER ONE
INTRODUCTION TO THE STUDY

Under the requirements of the *No Child Left Behind (NCLB) Act*, all public schools must meet adequate yearly progress (AYP) every year in raising student achievement. The end result was supposed to be 100% of students reaching proficient or advanced by the year 2014. As 2014 quickly approached, *NCLB* came under much scrutiny due to recent reports on the high number of *failing* schools. In studying AYP reports from schools across the nation for the 2010-2011 school year, it was reported that only 48% of schools met AYP. This percentage of failing schools has slowly increased from 29% failing in 2006, to the most current percentage of 48% in 2012 (Usher, 2012). In 2009, President Obama and Education Secretary Arne Duncan announced a \$4.3 billion dollar education initiative called *Race to the Top*. This initiative would provide money to schools who agreed to complete certain reforms (White House, n.d.).

These initiatives prompted education leaders to look into other ways to hold schools accountable for student performance. One of the identified avenues needing improvement was that of teacher evaluation systems. The following portions of this chapter will identify Missouri's history of teacher evaluation systems and what led to their need for change. It will also identify the problem statement and purpose of this research project. The research questions presented will be used to guide the study, along with the conceptual framework. The chapter will then present the design and methods for the study. Assumptions will be addressed and definition of key terms will be presented.

Finally, the chapter will provide the significance for this particular study as it contributes to the educational practitioner.

History of Missouri Teacher Evaluation

In 1983, Missouri adopted statute 168.128 RSMo which directed the local board of education of each school district to adopt a “comprehensive performance-based evaluation for each teacher employed by the district” (Missouri Department of Elementary and Secondary Education, 2012C, p. 2). The following years, the Missouri Department of Elementary and Secondary Education (DESE) developed model evaluation instruments for districts to use. In 1999, DESE released the Performance-Based Teacher Evaluation (PBTE). This evaluation tool has been in place in districts since that time to document the evaluation of all teachers.

In January of 2002, President Bush instituted the *No Child Left Behind (NCLB) Act*. This act placed strict requirements on school districts based on their students’ performance on state regulated standardized assessments. Lofty goals were created for districts that would require all students to reach levels of proficient or advanced by the year 2014. If schools did not meet these specified goals, they were identified as schools in need of improvement. School districts were then held to stipulations that would demonstrate their efforts to make improvements.

The *Race to the Top* initiative was launched by President Obama and Education Secretary Arne Duncan in December 2009. This initiative, which was part of the *American Recovery and Reinvestment Act*, offered large financial resources to schools willing to implement favored reform efforts. These reforms included the development of rigorous standards and better assessments, adoption of better data systems to provide

information about student progress, support for teachers and school leaders to become more effective, and increased emphasis on resources for interventions to turn around the lowest performing schools (White House, n.d.).

In addition to stimulating the discussion about teacher evaluation, *Race to the Top* legislation generated substantial and solid changes. Marzano (2013) identified two major changes in teacher evaluation that are directly tied to *Race to the Top*: (1) use of student growth measures to indicate teacher performance, (2) greater rigor in measuring the pedagogical skills held by teachers. *Race to the Top* sparked significant school reform efforts and shows that policy change is possible (Marzano & Toth, 2013).

Following the *Race to the Top* initiative, the Missouri Department of Elementary and Secondary Education (DESE) began its work on updating the system's standards and indicators for effective teaching. This process involved numerous stakeholders including educators, partner organizations, school districts, and educator preparation programs. DESE believed the improvement of students' learning required the improvement of the practice of those teaching in classrooms and providing leadership in schools. The new system focused on the improvement of effective educational practices and the professional development of teachers, principals, and superintendents. The state created seven essential principles of effective evaluation.

- Principle 1: Research-Based Practices: Educators' performance is measured against research-based practices tied to the improvement of student learning.
- Principle 2: Differentiated Levels of Performance: Uses multiple ratings to differentiate levels of performance

- Principle 3: Probationary Period for New Educators: Provides a probationary period with adequate time for sufficient induction and socialization support for new teachers and leaders.
- Principle 4: Use of Measures of Student Growth in Learning: Uses measures of student growth as a significant part of the evaluation of professional practice and ensures that professional or distinguished ratings cannot be obtained if student growth is low.
- Principle 5: Ongoing, Deliberate, Meaningful, and Timely Feedback: Provides this type of feedback on performance relative to research-based targets.
- Principle 6: Standardized and Periodic Training for Evaluators: Requires standardized and periodic training for evaluators to ensure reliability and accuracy.
- Principle 7: Evaluation Results to Inform Personnel Employment Determinations, Decisions, and Policy: Utilizes the results and data to inform decisions in these areas regarding employment (Missouri Department of Elementary and Secondary Education, 2013).

Based on these seven principles, DESE created three professional frames for the educator. These frames outlined the data sources for the evaluation of educators. The first frame is professional commitment. Professional commitment refers to the professional agreement between a teacher and leader. The data sources included in this frame are personnel files and records and quality of planning (units of instruction). The second frame is professional practice. This related to specific effective actions or behaviors in which a teacher or leader engages. Data sources for this frame were

identified as observations and performance and applied professional learning. The third, and last, frame is that of professional impact. This identified the effect, consequence or result that occurred due to the behaviors and commitments of the teacher and leader. Student performance measures and student feedback on performance are the data pieces reviewed to support this frame (Missouri Department of Elementary and Secondary Education, 2012A).

As 2014 approached, many districts across the nation found themselves in failing categories with the possibility of reaching their goal looking grim. In June 2012, the United States Department of Education approved Missouri's Elementary and Secondary Education Act (ESEA) waiver. This waiver gave Missouri some reprieve from the stringent requirements put in place by *NCLB*. The waiver called for the state to create its own accountability system that would better identify schools in need of improvement as well as recognize districts with exemplary results.

One specific goal included in the ESEA waiver is the implementation of a new teacher evaluation system. The waiver calls for the "state education agency and local education agency to focus on developing and implementing more meaningful evaluation and support systems" (Missouri Department of Elementary and Secondary Education, 2012B, p. 6). This, in addition to the requirements set forth by *Race to the Top*, put the burden of creating and maintaining a new teacher evaluation system on all districts in the state of Missouri. While DESE had outlined seven principals of effective evaluation along with three frames for collecting data, districts would be required to create their own systems for collecting and analyzing this data. To alleviate the work districts would put

into creating and maintaining a new evaluation system, the Network for Educator Effectiveness (NEE) system was created.

The NEE system was born from two auxiliary units of the College of Education at the University of Missouri: the Heart of Missouri Regional Professional Development Center (RPDC) and the Assessment Resource Center (ARC). The NEE system is based on the Missouri educator standards and indicators. The comprehensive system is web-based, providing a place for storing and managing data on each teacher in the evaluation system. NEE includes five sources of data that can be accumulated for each teacher in the system: mini-observations, units of instruction, professional development plans, student survey data, and student assessment data.

However the last data source, student assessment data, has not been implemented. While this data source is part of the NEE system, data collected (type and amount) is left as a local district decision. The districts studied had not implemented this portion of the system. Therefore, this portion of NEE was not evaluated in this research project. The key goals of NEE are to aide in teacher growth and to help administrators in identifying areas of strength and weakness in teachers they supervise (University of Missouri, 2013).

The NEE trains evaluators to make reliable and consistent measurements of educator's effectiveness, including engagement with students, professional development, and achievement goals. It strives to create a common understanding among all educators and administrators of the standards, criteria, and rubrics used for evaluation. The training required for administrators included discussion, collaboration in evaluating using video clips, and individual practice. Administrators are required to complete qualifying rounds in which their accuracy of scoring is measured. If accuracy is not within 80%, additional

training is required. Administrators are provided periodic and ongoing training and assessment to ensure accuracy of scoring (University of Missouri, 2013).

The advantages of implementing the NEE system reach to administrators, teachers, and students. For administrators, training provided for reliable and valid evaluation of teachers that produces scores that have meanings that are clear and transparent to teachers. The emphasis of NEE is on *coaching* teachers which leads to growth; this also creates a sense of teamwork between administrators and teachers where both parties are working together to improve instruction. The advantages identified for teachers are frequent observations that include relevant feedback. While effective teaching is identified and acknowledged, teachers are also recognized for growth and improvement. Evaluation information is presently in a timely manner, aligned to Missouri's teacher standards and easy to interpret. Students benefit from having more effective teachers in the classroom, and optimal learning environments are created to improve student performance (University of Missouri, 2013).

Problem Statement

Leaders of school districts across the nation are working to create new teacher evaluation systems to meet the requirements of new legislation (Coggshall, Rasmussen, Colton, Milton, & Jacques, 2012). These new systems need to include multiple measures of student learning and teacher practice. Due to requirements in place by the ESEA waiver and expectations outlined by the *Race to the Top* initiative, the NEE system was created. The creators of NEE strove to provide a system that districts could utilize to meet these new legislative requirements.

The entire system was built to guide administrators in evaluating teachers with the purpose of helping teachers grow and better themselves as professionals. Evaluation for this purpose aligns with Bolman and Deal's (2008) human resource frame. The human resource frame recognizes that people "are capable of enormous amounts of learning and adaptation" (Bolman & Deal, 2008, p. 123). However, the creators of NEE do not know what impact the system has had on change, or growth, in teaching behaviors following their initial implementation. Minimal data collection has been done to evaluate the effectiveness of NEE. This is the origination of the problem of practice. Knowledge of teacher growth would guide the creators of NEE to make appropriate changes to maximize the amount of learning that occurs during each evaluation in order to promote change and improvement in teachers. Rossi, Lipsey, and Freeman (2004) stated that at various times policy makers and program managers need to distinguish worthwhile programs from ineffective ones and launch new ones or revise existing ones to achieve desirable results.

Research Purpose

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. The study gathered teachers' perceptions of the impact NEE had on their teaching, specifically noting changes in teaching practice. This information served as an evaluation of the program and will be used to guide its further development. The newness of the NEE system makes it a prime candidate for an evaluation of its effectiveness. Creators can make adjustments to improve its success in guiding educators in their professional growth efforts. An evaluation of the system's

effectiveness, from the standpoint of the teachers, would give the authors rich data to ensure the most highly effective program. This fills a gap in research related to the effectiveness of this new system. The research will also inform other states as they develop new teacher evaluation systems during this time of high accountability under *No Child Left Behind* as they try to meet the rigorous standards outline in the *Race to the Top* initiative. The research will benefit school leaders as they implement NEE with educators under their supervision.

Research Questions

The research questions which guided this study were as follows:

1. What are teachers' perceptions of change impacted by the Network for Educator Effectiveness based on the framework presented by Kim Marshall?
 - a. What changes have been made in teaching as a result of the mini-observations?
 - b. What changes have been made in teaching as a result of the unit of instruction evaluation?
 - c. What changes have been made in teaching as a result of the professional development plan?
 - d. What changes have been made in teaching as a result of analyzing the student survey data?

Conceptual Framework

The purpose of this study was to examine teachers' perceptions of change impacted by NEE based on the framework presented by Marshall. The University of Missouri created the online tool for evaluators to meet the need for new evaluation

system requirements put into place by Missouri's Department of Elementary and Secondary Education. The conceptual framework for this study was derived from the work of Marshall (2009). In Marshall's book, *Rethinking Teacher Supervision and Evaluation*, he presented a system that supports the improvement of teachers. Improving teacher quality will help close the achievement gap. Research has shown that quality teaching is the most important factor in student achievement (Clotfelter, Ladd, & Vigdor, 2007; Nye, Hedges, & Konstantopoulos, 2004; Rice, 2003; Rivkin, Hanushek, & Kain, 2005).

Marshall's (2009) concepts were based on personal experiences as well as prior research and studies in the area of teacher supervision and evaluation. His approach is a framework that includes four parts a leader should implement in an effective teacher evaluation system. The first piece is that of mini-observations. Marshall presented the idea of visiting teachers' classrooms frequently in small chunks of time. These frequent, unannounced visits give the principal a sense of whether or not the curriculum is being taught. An important aspect to this is the feedback teachers receive after each visit. The feedback should be timely and specific to be most effective. Marshall stated that giving feedback in person opens a two-way conversation about what was observed; "Ideally, teachers will leave feedback conversations with specific ideas for improving their practice. . ." (p. 81). These mini-observations are a piece of the data that accumulates in the NEE system for each enrolled teacher.

The second aspect to Marshall's (2009) plan is that of monitoring and evaluating curriculum unit planning. Teachers must submit a unit of instruction for evaluation in the NEE system. When the learning objectives are not clear, it is difficult to identify

problems in student learning along the way. Marshall believed a principal must know the bigger curriculum picture in order to know what to look for in a mini-observation.

Schmoker (1999) stated, “Every teacher deserves a clear, manageable, grade-by-grade set of standards and learning benchmarks that make sense and allow a reasonable measure of autonomy” (p. 21).

Similar is Marshall’s (2009) third aspect, monitoring and evaluating interim assessments. This piece requires the evaluator to consider what is being taught and how the success of that teaching is being measured. Marshall contended the widening gap in achievement stems from an embedded paradigm in teaching: teaching, testing, and moving on without fixing any learning problems. He stated most teachers do not have the training or tools to stop when students are below mastery and address the problems before continuing. This process widens the gap even further. This aspect leads to the analysis of student achievement data included in the NEE system.

The last piece to the system is that of end-of-year rubric evaluations. This rubric, created by the evaluator, is used as a *summative* type evaluation to give the teachers feedback on their overall performance. Marshall (2009) stated “they are best used to pull together all the impressions gathered in the course of a school year” (p. 165). These rubrics would include information such as student surveys and professional development plans.

The framework presented by Marshall (2009) was the measure to which the NEE was held. Teachers were asked to give their perspectives of NEE based on the recommendations of Marshall for an effective evaluation system. This information will be helpful in the further development of the system, as well as, guide leaders in their

presentation of the system in the future to make sure that it is being implemented correctly and communicated well to teachers.

Design and Methods

Rossi et al. (2004) defined program evaluation as “the use of social research procedures to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action to improve social conditions” (p. 16). A valid description of the performance of a program includes a correct representation of what the program truly accomplishes. According to these authors, program evaluations usually involve one or more of five program areas: need for the program, design of the program, program implementation and service delivery, program impact or outcomes, and program efficiency.

This research project was a program evaluation to measure teachers’ perceptions of the change that occurred in their instruction due to the implementation of the Network for Educator Effectiveness system. The study measured the program’s impact or outcomes based on the teachers’ views of how it changed their instruction and made them better teachers. According to Rossi et al. (2004), the purpose of this type of evaluation is to inform social action. This study will do so by informing those creating and implementing the program about how effective the system has been in creating change in teacher behavior.

This qualitative study collected data was analyzed to determine to what extent NEE created change in teacher behavior. This research included data collected from focus groups and surveys. By using these tools, the researcher hoped to glean a large

enough sample of qualitative data to make connections and identify recurring themes. Krueger and Casey (2009) stated focus groups are purposed to promote self-disclosure among those participating. One focus group was assembled from each of the nine buildings in the districts being studied. A large net was cast to collect data from a large group through an open-ended question survey distributed to all teachers in the districts.

Participants for this study came from two small school districts in southwest Missouri. The sample included 311 teachers from the pre-kindergarten through twelfth grade, including encore teachers such as art, music, and physical education, 84 of which chose to participate in the survey and 21 participated in a focus group. The two school districts were chosen due to their involvement in the NEE system, as well as, their availability to the researcher. This availability derived from their proximate location to the researcher's home district along with personal relationships with several administrators within those districts. Data were collected and analyzed using a coding system. Upon initial review, themes in the data were recorded and then a further study of the data revealed information that was coded into identified themes.

Assumptions

There were three major identifiable assumptions in this study. The first assumption was that the Marshall Framework is the most appropriate framework by which to measure the NEE. This framework was chosen because it so closely aligns with the components of NEE; however, the researcher acknowledges there are multiple frameworks for teacher evaluation that are research based and effective, as is the Marshall Framework.

Another major assumption was that the NEE has made an impact on the teachers of these two districts. Some bias may enter in this assumption in that the researcher has implemented this system in her building and has witnessed change first-hand. This does not mean this system has created change in other settings. This assumption comes from personal experience which may lead to bias.

There was also the assumption that the NEE system is implemented correctly in all of the schools studied and to its fullest extent. The creators of NEE provide principals and other school leaders with extensive training, even requiring qualification testing that must be completed at a mastery level before the administrator is able to participate in NEE. However, there is the human element involved in conducting the evaluations which opens the avenue for error. It is also possible that schools are not utilizing all four pieces of the system. This could cause skewed results based on a teacher's lack of experience with one or more of the pieces.

Definition of Key Terms

For the purposes of this study, the key terms listed were defined as follows:

Framework. A framework is the basic structure of something (Merriam-Webster, 2014). In this case it is a structure, or skeletal support, for conducting teacher evaluation. It provides all the elements needed to conduct meaningful, effective evaluations.

Mini-observation. The mini-observations are frequent, unannounced classroom visits that give prompt feedback to teachers. The high frequency and prompt feedback are important pieces to improving teaching and learning (Marshall, 2009).

Network for Educator Effectiveness (NEE). NEE is a comprehensive, web-based teacher evaluation system that includes five pieces of data. This data can be accumulated for the purpose of evaluating teacher effectiveness and promoting teacher growth (University of Missouri, 2013).

Professional development plan. The purpose of a professional development plan is to require teachers to set goals and encourage them to self-assess (Marshall, 2009). Job-embedded professional learning is grounded in day-to-day teaching practice, occurs regularly and frequently, considers student data to identify problem areas in need of improvement, and is aligned to school goals and curriculum (Coggshall et al., 2012).

Student assessment data. Data collected from student performance events constitute student assessment data. This data allow principals an ongoing look at evidence of student learning (Marshall, 2009).

Student survey data. Student survey data are data collected from the students about a teacher's performance. Carefully field-tested questions posed to students can provide more reliable data than any other form of teacher evaluation according to Ripley (2012).

Teacher evaluation. A process of data collection that will promote growth in teacher performance and aide in making decisions regarding personnel, employment decisions and policy making (Missouri Department of Elementary and Secondary Education, 2013).

Teacher perceptions. Perceptions are the ways a person thinks about or understands something (Merriam-Webster, 2014). Teacher perceptions are those ways of thinking and understanding held by teachers involved in the study.

Unit of instruction. A unit plan describes a teacher's plan for teaching skills and concepts over a three to five-week period (Marshall, 2009).

Significance of the Research for Leadership Practice

The most important factor affecting student learning is the teacher (Wright, Horn, & Sanders, 1997). Wright et al. also contended "more can be done to improve education by improving the effectiveness of teachers than by any other single factor" (p. 63). This study will aide in the future refinement of a system to promote teacher quality and teacher growth. Such a system will increase teacher effectiveness, therefore increasing student achievement.

Evaluations of programs whose mission is to "intervene in social conditions in ways that make them better" (Rossi et al., 2004, p. 25), or social programs, serve the purpose of improving the program itself. While the program improves the human condition, the evaluation improves the program (Rossi et al., 2004). The NEE system is a very new system. The creators of the system conducted a pilot year for schools to try it and provide feedback for improvement. This research will help to inform their future decision making when refining the system. The data collected will inform on the system as a whole, as well as, individual parts and components. Because there are four active pieces to the system, mini-observations, unit of instruction, professional development plans, and student survey information, the researcher's intent is to provide data to support the effectiveness of each piece along with areas of weakness that need improvement.

By collecting this data based on what teachers think about the system, administrators who implement the evaluation system will know how to better communicate with their teachers regarding the implementation of the system and then

throughout the evaluation process. Understanding teachers' thoughts and needs will aide them in understanding how to improve the systems effectiveness by their own implementation of the system.

Summary

This paper has outlined the importance of, and need for, effective teacher evaluation systems. NEE sought to meet the needs of school districts in the state of Missouri by providing a web-based teacher evaluation system. This system provided the guidelines for what should be included in the evaluations, along with an online storage of data collected. The research questions, based on the framework of Marshall (2009), guided the data collection for this study. The qualitative study will address what change has occurred following implementation of the NEE system.

CHAPTER TWO

REVIEW OF LITERATURE

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. This literature review will present information to outline the history of teacher evaluation systems. It will also outline the importance of teacher evaluation systems and their impact on student achievement. The literature review will identify areas of research regarding barriers and obstacles to conducting teacher evaluations. Research on best practices for teacher evaluation will be studied and presented in this chapter. Lastly, a review will be presented of Marshall's (2009) work on Rethinking Teacher Supervision and Evaluation.

History of Teacher Evaluations

“In the early 1700s, education was not considered a professional discipline or field of study” (Marzano, Frontier, & Livingston, 2011, p. 12). Clergy were assigned the duty of selecting teachers and making judgments about their teaching. The teacher was considered a community servant and was often supervised by a committee that was to monitor the quality of instruction. As communities grew with industry, school systems became more complex. As this occurred, one teacher in the building was often selected to perform certain administrative duties. This *principal* teacher role soon developed into the role of the building principal (Marzano et al., 2011).

During the late 19th century and early 20th century, education was dominated by two views of education, both different. One view was that of John Dewey. According to

Marzano et al. (2011), Dewey saw democracy, not scientific management, as the scaffold, or support, for learning. He felt schools should be structured so students could practice citizenship and develop ideas of democracy. This knowledge would help them move from passive learners to their role of active citizens (Marzano et al., 2011). On the opposing side of this were the views of Frederick Taylor. Taylor was a proponent of scientific management which espoused the measurement of specific behaviors to identify the one best way to complete a task. “In every case one method has been tried out, until the proper remedy has been found. That series of proper eliminations, that evolution, is what is called scientific management” (Taylor, 1916, p. 67). Taylor believed education could be conducted in this same way.

Teacher evaluation systems have changed dramatically over time. Literature can be found dating back to the 1940s documenting the types of qualities that were expected to be found in teachers (Hughes, 2006). Danielson and McGreal (2000) stated that during the 1940s and 1950s evaluators of teachers were looking for what was termed presage variables. These were variables that included personal traits that teachers naturally possessed, such as voice, appearance, demeanor, character, emotional stability, honesty, and enthusiasm. Konold et al. (2008) defined them as “a set of formative experiences (educational and environmental factors that influence teachers’ behaviors), demographic characteristics (race, ethnicity, gender, age, etc.), and personal properties (e.g., personality characteristics, attitudes, beliefs)” (p. 301). These traits were held with high esteem because educators during this time period thought them to be the best predictors of teacher success. Danielson and McGreal (2000) found no real evidence to link these variables to student learning. Kennedy (2010) stated, “In the past several decades, as

researchers have sought the qualities that make a good teacher, they have correlated numerous teacher characteristics with student learning outcomes; but none has demonstrated very high correlations” (p. 591). She explained that as people have looked to personal characteristics to define teacher quality, they have over-looked aspects of their work that is outside of their control. These aspects include resources, planning time, and other infrastructures of the school that may affect their performance.

Research has been underway for over a century on the relationship between teachers’ characteristics and teacher quality, yet little progress has been made in linking the two. Rockoff, Jacob, Kane, and Staiger (2011) conducted a study to identify if one can recognize an effective teacher at the time of hire based on factors that were observed at that time. The results of this study show that there is no single factor that can predict success. Thus, it is important that teacher evaluation systems have continued to change and evolve into the numerous systems that are available today.

Due to an increase in research during the 1960s and 1970s, the focus placed on teacher qualities shifted from presage variables to evaluations of what actually took place within classrooms. “At the beginning of the 1970s, researchers began to look at the effects of instruction on student learning” (Marzano, Pickering, & Pollock, 2001, p. 1). Danielson and McGreal (2000) noted that researchers began developing evaluation processes that were designed to enhance instruction. The indicators were more focused on creating an accurate depiction of what was happening in the classroom. Evaluation skills and classroom observations of teaching techniques became the real focus over personal attributes of the teacher. This idea of clinical supervision spread like wildfire

according to Marzano et al. (2011). They stated by 1980, about 90 percent of school administrators were employing some type of clinical supervision model.

In 1983, the National Commission on Excellence in Education released the report, *A Nation at Risk* (U.S. Department of Education, 1983) that detailed the rising level of mediocrity of educational performance. This report placed the public's attention on the need for a restructuring of the country's education system. The Carnegie Forum on Education and the Economy responded to one section of this report, teacher quality, by creating a task force of stakeholders. This task force argued that in order to have high performing schools, the nation's teachers would have to be better trained and supported. The task force made recommendations for strengthening standards in teaching in a report called *A Nation Prepared: Teachers for the 21st Century* (Carnegie Forum on Education and the Economy, 1986). This resulted in the creation of the National Board for Professional Teaching Standards (NBPTS). Its purpose was to implement the report's core recommendations (National Board of Professional Teaching Standards, 2014).

The NBPTS certification is an advanced teaching credential. Teachers are required to analyze their teaching context and students' needs, submit videos of their teaching, and provide student work samples that demonstrate growth and achievement.

The following criteria must be shown through this process:

1. A strong command of content.
2. The ability to design appropriate learning experiences that elevate student learning.
3. The use of assessments to inform instructional decision making.
4. Partnerships with colleagues, parents, and the community.

Through this structured process, teachers can obtain National Board Certification. These teachers show powerful teaching that improves student achievement (National Board for Professional Teaching Standards, 2014).

Since the beginning of the 21st century, the emphasis on supervision has shifted to an emphasis on evaluation. Several reports came out early in the century stating that teacher effectiveness should be measured based on classroom observations as well as student achievement gains (Marzano et al., 2011). There is much debate surrounding the inclusion of student performance data in teacher evaluation systems (Danielson & McGreal, 2000; Mathis, 2012). Tucker and Stronge (2005) stated that school reform efforts are occurring in a variety of forms, with two of the most prominent being a focus on higher teacher standards and improved student performance. They questioned, “If teachers do, in fact, make a difference in student learning, and if we are to have competent and caring teachers, shouldn’t we relate teacher work to student work? Shouldn’t student achievement be a fundamental measure of teacher effectiveness?” (p. 13)

Tucker and Stronge (2005) believed that student learning can, and should, be determined by a variety of assessments. The range of possibilities includes norm-referenced achievement tests, criterion-referenced achievement tests, and other types of student assessments. “The testing information that provides data on goal attainment at the school and district level can also be used at the classroom level for assessing teacher quality” (Tucker & Stronge, 2005, p. 91).

This concept of including student growth data in teacher evaluation, sometimes referred to as *value-added*, also has its deficiencies according to Danielson and McGreal

(2000). They warned that a student's home environment can affect a student's absolute performance as well as the rate of learning. Also, depending on how data are collected, a teacher could be misjudged if a good student were to move during the course of the year while a student more difficult to teach moved into the class. Some students possess cognitive or behavioral characteristics that make them more challenging to teach. If a system places too much emphasis on demonstrated student growth, it could provide a disincentive for teachers to teach this type of student. The argument is made that finding fair ways to evaluate students poses a dilemma for schools as well. These technical issues are significant and need to be addressed before high-stakes testing is implemented (Danielson & McGreal, 2000).

In light of these ongoing disagreements, teacher evaluation practices are currently under great debate. Across the United States, the school accountability theme is commonly discussed among state government officials and local community members. Public schools are being examined, and the people they serve are calling for improvement (Tucker & Stronge, 2005). One area that can be agreed upon is the importance of teacher quality in improvement efforts (Corcoran, 2007; McQuarrie & Wood, 1991; Owings, Kaplan, Nunnery, Marzano, Myran, & Blackburn, 2006; Range, Duncan, Scherz, & Haines, 2012).

Importance of Teacher Evaluation and Their Effect on Student Achievement

DuFour and Marzano (2009) stated, "principal evaluation of teachers is a low-leverage strategy for improving schools, particularly in terms of the time it requires of principals" (p. 64). Principals historically have spent an immense amount of time on a process that has not improved teaching and learning when research has shown that the

quality of instruction is the single most important factor in student achievement (Marshall, 2009). This section will introduce research to support the importance of teacher quality on the success of students. In the latter part, it will connect a quality teacher evaluation system to ensuring teacher quality.

Teacher Effects on Student Learning

Research has shown evidence that the quality of teachers has decreased steadily over time, while nearly all modern research claims teacher effectiveness to be among the most effective strategy schools have to improve student achievement (Corcoran, 2007; McQuarrie & Wood, 1991; Owings et al., 2006; Range et al., 2012). Goe (2007) also supported this stance by stating that teacher quality matters a great deal in regard to student learning. She stated that a large amount of research has been conducted measuring teacher quality based on student learning outcomes. Nye et al. (2004) conducted a four-year study to measure the effects of teacher effectiveness on student achievement. Their study found that teacher effects are larger than school effects when it comes to individual student achievement. They stated that the teacher a student gets matters more than which school a student happens to attend. Konold et al. (2008) stated, “Data tell a seemingly simple story: Teaching behaviors matter—they matter a great deal” (p. 309).

The research of Sanders and Rivers (1996) presented the additive or cumulative effect of teacher effectiveness on student achievement. Sanders and Rivers spent a multi-year period studying what happened to students who had teachers that produced high achievement results compared to those whose teachers produced low achievement results. They discovered that when students, beginning in third grade, had teachers with high

achievement results for a period of three years in a row, those students scored on average at the 96th percentile on their statewide mathematics assessment at the end of the third year. Students with a comparative achievement history that were placed with low-achieving teachers for three consecutive years, obtained a score on the same statewide mathematics assessment at the 44th percentile. This large difference of 52-percentile points led researchers a major conclusion that teachers make a difference.

Teacher Evaluation Effects on Student Learning

Understanding the importance of teacher effectiveness on student achievement, it is easy to make the assumption that evaluating teacher effectiveness also has an impact on student learning. However, there is research to support this assumption. “Faculty evaluation is widely understood to be the most effective tool to improve the quality of instruction in schools” (Koops & Winsor, 2006, p. 61). Heck (2009) emphasized the importance of increasing teacher effectiveness in order to improve student outcomes. His study included 9,196 students nested within 156 elementary schools.

Of Heck’s (2009) three major findings, the first was that teacher effectiveness of successive teachers was related to student achievement in reading and math. Second was collective teacher effectiveness, considered an organizational property of the school, was positively associated with achievement levels. Third, the quality of academic organization and teaching processes, along with the stability of the teaching staff, were positively related to achievement levels. His findings related to teacher effectiveness and student achievement led to his claim that a focus on hiring and retaining high-quality teachers and facilitating improved academic processes can bring about increases in

school effectiveness. This supports the assumption that teacher evaluation systems, used to improve the quality of teaching, have an effect on student achievement.

Range et al. (2012) stated, “Although teachers have the greatest impact on student achievement, school leaders also have a powerful, albeit indirect, impact on student success” (p. 303). They also contended instructional leadership is the primary method by which principals influence student achievement. Instructional leadership is defined as helping teachers improve their practice.

According to Khan, Khan, Shah, and Iqball (2009) instructional leadership is a complex task; it means that the principal becomes a leader of leaders. Instructional leadership involves working with others (students, teacher, and parents) to improve the quality of instruction and curriculum. Within this role, principals formatively supervise teachers by visiting their classrooms to collect data on their performance. Data are then used to summatively evaluate teachers. The primary purpose of evaluation is to provide evidence for retention or non-retention and to ensure teachers are held to common, rigorous standards (Range et al., 2012).

Focusing on student achievement is usually a result of school-wide improvement efforts. These efforts often, however, overlook the importance of teacher evaluations as part of the process. Normore (2005) stated if the evaluation of teachers is based on technical knowledge, shared language and evaluation skills, it can serve as a vital element for school improvement initiatives. Teacher evaluation can be a strong tool in school improvement initiatives because it addresses individual development of teachers and their sense of professionalism.

Normore (2005) summed these thoughts into one quote, “The benefits to rendering greater attention to personnel evaluation include a likely increase in the success of school improvement interventions” (p. 350). McQuarrie and Wood (1999) added to that argument by noting supervision, staff development, and teacher evaluation provide the framework for improving and maintaining effective instructional practices. This provides the avenue for administrators and teachers to work together to focus on meeting the needs of students and improving the teaching and learning in schools.

The most useful way to change an organization is to focus on people according to Khan et al. (2009):

When you talk about school progress, you are talking about people development. That’s the only aspect to improve schools unless you mean painting the buildings and fixing the floors. But that is not the school, that’s the shell. The school is people, so when we talk about excellence or improvement or progress, we’re really talking about the people who make up the building. (p. 589)

It was also concluded that if the assumption is true that people are the key to school improvement, then the fundamental role of the principal is to help create an environment in which staff are enabled to develop so the school can more effectively achieve its’ goals.

The research presented in this portion of the chapter emphasizes the important role of teachers in relation to student success. The evaluation of teachers has also been proven to increase student learning when its purpose is to improve the quality of instruction that students receive. The next portion of this chapter will outline barriers and obstacles that can hinder effective teacher evaluations.

Barriers and Obstacles to Conducting Teacher Evaluations

While there are numerous evaluation systems available for implementation, they all have their own deficiencies. Research has identified several barriers or obstacles to conducting effective teacher evaluation systems. These include qualification/quantity of evaluators, types of data used to evaluate, infrequency of evaluations, stress and nervousness on the part of the teacher, and teachers putting on a *dog and pony show* rather than a real picture of what happens from day to day.

Qualification/Quantity of Evaluators

Advice to policy makers regarding teacher evaluation systems was given by Mathis (2012) of the University of Colorado. One of his points was for districts to ensure they have a sufficient number of highly-qualified evaluators. He warned that lacking in this area can cause the system to be irregular, uneven, and inefficient. Teachers expressed this same concern in a study completed by Larsen (2009). Larsen reported teachers stated their performance appraisal was only as useful and meaningful as those performing the appraisal.

Knowledge of content, pedagogy, and students' developmental levels are all highly relative to teaching according to Danielson and McGreal (2000). Teachers may be much more knowledgeable in these areas than the administrator who evaluates them. These facts undermine the evaluation process, making it seem as if it has little value. Another concern, expressed by Stiggins (1986), was observations are subject to potential biases and opinions of just one observer. Some schools may only have one administrator conducting teacher evaluations; this limits the viewpoint and relies solely on that one person's thoughts and feelings. Evaluation instruments often appear to have precise,

numbered ratings, but these often are not matched to any specific reference outside of the instrument. This allows for the instrument to express subjective impressions based on the observer completing the instrument (Moxley, 1978).

Concerns about raters can affect the validity of observations was supported by Goe, Bell, and Little (2008). Proper training is necessary as evaluators are making judgments in the moment about what they see. There is no assurance that districts are using training systems to ensure validity and reliability in their evaluators. This means measuring teacher effectiveness through observations can be uneven, threatening the credibility of the protocols in use (Goe et al., 2008). In conclusion, no matter what system is used, its effectiveness depends on the quality of the person conducting it.

Types of Data Used to Evaluate

Another area of concern for Mathis (2012) was the use of student data in teacher evaluation systems. His concerns were the error of measurement in test scores is large and there are not relevant test scores available for all teachers. There is also the concern that using these scores creates undesirable consequences of narrowing the curriculum and teaching to the test and that scores earned by students are also influenced by factors that occur outside of the school setting. These factors include parent education and wealth or the lack thereof. Danielson and McGreal (2000) agreed that using standardized achievement data to evaluate teachers present technical challenges and are difficult to implement fairly.

Opposed to the issues presented regarding the inclusion of student achievement data are thoughts of concern that student achievement data are *not* used in teacher evaluation models (White, Makkonen, Vince, & Bailey, 2012). Danielson and McGreal

(2000) addressed this by suggesting evaluations included data from student work that can provide direct evidence of student learning. They recommended teachers select student work to represent the full range of skill and ability in their class. If collected over a period of time, this data can show student growth and improvement. Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012) claimed *value-added models* are designed to evaluate student test score gains from one year to the next can be used as evidence of the quality of teacher practices. Their belief was that measured achievement gains for specific teacher's students can define or reflect a teacher's effectiveness.

Infrequency of Evaluations

Often overlooked is the frequency of classroom visits by evaluators. Stiggins (1986) stated his concern that teacher evaluations only rely on a narrow sample of teacher performance, usually only one or two brief classroom observations. According to Marshall (2009), "The principal sees only a tiny fraction of teaching time" (p. 22). He stated on average, principals only thoroughly evaluate one lesson per year per teacher. This equates to only about 0.1% of their teaching. Besides the fact that the other 99.9% of the time goes unobserved, teachers' effectiveness is solely based on one lesson. Marshall stated for an evaluation to be credible, principals have to find a way to measure what is going on during that other 99.9% of the time.

Coinciding with this, teachers feel principals do not see them in action often enough, according to Larsen's (2009) study. Teachers who are seldom evaluated, typically once or twice a year, feel isolated and undervalued. Their preference is actually to be evaluated more frequently, even when those evaluations produce negative results. Teachers view this *spot-check* of their competence as holding no value. These visits do

not provide continuous data about the teacher's coverage of the curriculum or teaching methods (Cooper, Ehrensals, & Bromme, 2005). This is why Mathis (2012) recommended districts have sufficient numbers of highly qualified evaluators which would help increase the frequency of observations.

Stress and Nervousness

Larsen (2009) identified stress and nervousness on the part of the teacher as an obstacle to effective evaluations. Her study found teachers felt stressed and nervous, "even to the point of being sick" (p. 18). A contributor to this stress was the amount of time teachers were spending in preparation for the evaluation. A pattern seemed to emerge showing teacher's time was being restructured, so they were spending less time in contact with students and more time completing paperwork in order to meet requirements. The nervousness led to self-doubt in teachers. One teacher stated he experienced a lot of "double-thinking" (p. 22), a process through which he questioned everything he said and did during class throughout the evaluation. It is likely the evaluator did not get a *real* picture of what happens in the classroom in a day-to-day basis. In a study by Ong'ondo and Borg (2011), they identified fear as a primary impact on supervision. Teachers' practices were constrained by a perceived need to please their supervisors. They found, however, this fear motivated teachers to be vigilant at all times. However, the teachers did feel they taught in a more comfortable manner when the supervisor was not present. This more relaxed manner allowed them to try out new techniques they would not take chances on in the presence of a supervisor.

Another lens to view this through would be the idea of observer effects that occur when there is an outside observer presence in the classroom. McIntyre (1980) stated

there is a growing body of research that states the presence of an observer affects normal teaching behavior. This taken into consideration would lead to the conclusion that data collected during an observation are not a true reflection of a teacher's behavior. The report may only reflect the teacher behavior during an observation when an observer is present. He indicated this occurs most in relationship to a teacher's affective domain. The presence of an observer in the classroom resulted in positive teacher affective behavior, more praise and less criticism, and higher ratings of classroom attitude. The stress and nervousness of having an observer in the classroom contribute to this observer effect that is documented by McIntyre.

Marshall (2009) also addressed this concept of the observer effect stating "the principal's presence changes classroom dynamics" (p. 24). He stated this can often have a positive effect, causing the students to behave better creating a more orderly, cooperative class. However, he noted in some circumstances it can also have a negative effect. Teachers who are nervous their job is on the line or feel intimidated by the presence of the principal may be thrown off stride and botch their lesson.

Dog and Pony Show

Teachers often prepare their best lesson, putting on a "dog and pony show" just to impress the evaluator. Marshall (2009) said that evaluation visits are often announced ahead of time so teachers naturally want to put their best foot forward and showcase an especially well-planned lesson. While these lessons can give an administrator some sense of what a teacher is capable under the best circumstances, they know they are not observing the type of instruction that students receive on a day-to-day basis. By conducting evaluations in the announced way, Marshall claimed it sends the wrong

message to teachers: “You can put on a special show for me and that’s what counts” (p. 24). By doing this, the administrator is staying that it is okay to do something less-special the other 99.9% of the time.

The over-preparation teachers were doing to put their best foot forward for their evaluation was pointed out by Larsen (2009). Respondents to her survey noted teachers create lessons full of *bells and whistles* and *sparks* for their evaluations. The process is considered an exercise of “jumping through the hoops, not always realistic of what is really happening in the classroom” (pp. 22-23). The concern is that evaluations become about putting on a show rather than the teaching itself. Administrators are not seeing a true *day-in-the-life* of their students.

In a study by McIntyre (1980) during which teachers were asked to increase physical contact with students, it was noted that the presence of an observer in the classroom produced teacher behaviors different than those present during their normal routine. This was done by evaluators completing observations in the classroom, as well as behind a two-way mirror. McIntyre stated, “...if teachers know what an observer wants to see, they will perform that behavior” (p. 37). Ong’ondo and Borg (2011) found this same phenomenon. Teachers reported to them, “We teach plastic lessons to please them” (p. 521). Teachers even reported that learners were alerted and specifically prepared when an observation was to occur.

Obstacles and barriers to conducting teacher evaluations have been identified and include qualification and quantity of evaluators, types of data used in evaluation, the infrequency of evaluations, stress and nervousness of the teacher, and over-preparation by teachers. An effective evaluation system considers these factors and makes

adjustments to combat their negative impact on the evaluation's effectiveness. Research has also identified best practices in teacher evaluation that produce positive results in evaluation effectiveness. These practices are presented in the next section of this chapter.

Best Practices for Conducting Teacher Evaluations

There are multiple frameworks for teacher evaluations. Though there are a lot of similarities, they all have uniqueness about them. Frameworks are all created based on research conducted by the author, as well as their personal experiences. The following section of this paper will outline three such frameworks. The first is by Marzano, Frontier, and Livingston (2011), the second framework by Danielson (2010), and the third framework by Tucker and Stronge (2005). Following those sections, the paper will address specific pieces of teacher evaluation systems and their value. Elements addressed will cover classroom observations and feedback, teacher professional development, units of instruction, and student evaluation.

Marzano Framework

Marzano, Frontier, and Livingston (2011) created a framework that includes five conditions for developing teacher expertise. They were described as: (a) a well-articulated knowledge base for teaching, (b) focused feedback and practice, (c) opportunities to observe and discuss expertise, (d) clear criteria and a plan for success, and (e) providing recognition of expertise. Included under the first condition were also four domains that outline what a well-articulated knowledge base for teaching: classroom strategies and behaviors, planning and preparing, reflecting on teaching, and collegiality and professionalism. Marzano et al. (2011) believed that the more skilled a teacher is, the higher their student achievement will be. They stated, "the purpose of supervision should

be the enhancement of teachers' pedagogical skills, with the ultimate goal of enhancing student achievement" (p. 2).

The model presented by Marzano et al (2011) provides tools for data collection in regard to their teacher evaluation framework. These include an observation protocol (long and short form) and an observation snapshot. They presented rubrics to evaluate planning and preparing, evaluation of personal performance, implementation of a personal professional growth plan, promoting a positive environment, promoting exchange of ideas and strategies, and promoting district and school development. All protocols and rubrics rate teachers into five categories: innovating, applying, developing, beginning, and not using. Indicators are listed for each of the five categories (Marzano et al., 2011).

Danielson Framework

Goe et al. (2008) presented the Framework for Teaching (FFT) created by Danielson in 1996. This approach was based on research conducted by Danielson. The researcher worked to do the following:

- Define a holistic view of teaching.
- Describe the complex relationship of teachers and students.
- Examine the importance of tailoring teaching to the individual, developmental, and cultural differences of students.
- Consider the influence of the subject being taught on teaching.
- Spell out the implications of all this for teacher assessment (p. 22)

Danielson and McGreal (2000) explained in this research-based protocol, there are four domains of teaching responsibility: (a) planning and preparation, (b) classroom environment, (c) instruction, and (d) professional responsibilities.

Planning and preparation includes understanding of what is being taught, knowledge of the students' backgrounds, and designing instruction and assessment. Classroom environment refers to the teacher's ability to establish a classroom environment that is conducive to learning. Instruction is concerned with the teacher's skill in engaging students in learning by incorporating multiple instructional strategies. Professional responsibilities address a teacher's other responsibilities including self-assessment, communication, and professional development (Danielson & McGreal, 2000). The FFT includes 22 components, with 76 smaller elements, clustered under those four domains. The goal of the framework is to serve as a "foundation for professional conversations among practitioners as they seek to enhance their skill in the complex task of teaching" (Danielson, 2011, p. 1).

This framework provides several examples of checklists, protocols, and rubrics for schools to adapt to their needs for the purpose of data collection. The Danielson framework places teachers in either Track I (for beginning teachers), Track II (for tenured teachers), or Track III (for tenured teachers needing assistance). The purpose of these tracks is professional growth and development. Danielson and McGreal (2000) believed this to be a roadmap for an effective teacher evaluation system that combines quality assurance with professional development for all teachers.

Tucker and Stronge Framework

Tucker and Stronge (2005) introduced a framework for teacher evaluation that places an importance on student achievement in teacher evaluation systems. They presented different case studies in which districts used student achievement to evaluate teachers' performance. From those studies, they created a list of eight recommendations for districts to instruct them on the most appropriate ways to integrate student achievement into their evaluation systems.

1. Use student learning as only one component of a teacher assessment system that is based on multiple data sources.
2. Consider the context in which teaching and learning occur.
3. Use measures of student growth versus fixed achievement standards or goals.
4. Compare learning gains from one point in time to another for the same students, not different groups of students.
5. Recognize that gain scores have pitfalls that must be avoided.
6. Use a timeframe for teacher assessment that allows for patterns of student learning to be documented in a fair manner.
7. Use fair and valid measures of student learning.
8. Select student assessment measures that are most closely aligned with existing curricula.
9. Don't narrow the curriculum and limit teaching to fit a test unless the test actually measures what should be taught. (pp. 96-102)

The belief of these authors is "teacher performance directly impacts student learning and, therefore, measures of student learning should be included in the process of assessing

teacher quality” (p. 102). Data collection presented in this framework includes rubrics that judge the types of student evaluations performed, how they were used, and the results of student performance.

While it is important to study research-based frameworks for teacher evaluation, it is also vital to look at what research says about individual pieces of teacher evaluation systems that make up these and other frameworks. The following sections of this paper will address four aspects that play a role in the evaluation and supervision of teachers. These parts include classroom observation and feedback, teacher professional development, preparing units of instruction, and student evaluation. Research surrounding each of these topics will be presented.

Classroom Observation and Feedback

“A teacher’s classroom instructional practice is perhaps one of the most important yet least understood factors contributing to teacher effectiveness” (Danielson, 2010, p. 1). Wanzare (2002) stated classroom observation has been the most commonly used and most practical way of obtaining data on teacher instruction. The primary purpose to conducting classroom observations is to answer the following questions. “What is going on? How does this work? Can it be explained (by the data)? What laws and rules govern behavior in this context? How can classroom events be described accurately and vividly?” (Wanzare, 2002, p. 215). Data collected during these classroom observations can help to identify successful teaching practices and assess the effectiveness of teaching strategies and techniques. It also provides a picture of ongoing events in the classroom. Koops and Winsor (2006) stated the goal is to help faculty members identify strengths and weaknesses in their teaching and to improve their skill and craft of teaching.

The specific focus of a classroom observation may vary from one observer to another, or one teacher to another, depending on the goals set for that particular observation. These goals are usually determined in an observation pre-conference. During this time, the evaluator and the teacher discuss certain aspects of teaching that need to be monitored due to a need of improvement in that area, or based on school goals, or particular student needs (Wanzare, 2002).

“Instructional leadership means very little unless leaders are willing and able to observe teachers, offer advice about problems, and make formative evaluations that support and pinpoint areas to improve” (Khan et al., 2009, p. 585). Khan et al. provided the following tips for effective supervision:

- Classroom observations will be used to observe what is actually going on in the classrooms.
- What is the performance of the teacher and how are the students performing.
- Meets with teacher after each visit to discuss what was observed.
- Encourage teacher to express feelings and opinions about observational data and class activities.
- Offers teachers alternative teaching techniques and explanations of classroom events.
- Give praise for specific development of teacher’s skill if observed.
- Recommends resources and training programs in areas in which teacher needs to improve. (pp. 585-586)

Following these guidelines would lead to a more effective observation process.

While there is research supporting the positive effects on student achievement compared to teachers' scores on observation protocols, there is still a need for more research in this area. Goe et al. (2008) stated that little research is available to link scores on well-validated observation protocols to student outcomes. Observations of teachers may tell a lot about how a teacher's practice aligns with good practice, but without linking this information to student achievement, it is difficult to determine effectiveness (Goe et al., 2008).

Unfortunately, the domination of urgency forces administrators to complete a teacher's evaluation during one hour long observation. When this formal observation is the only data collected, this experience provides little in the way of a formative evaluation. Rather, this process becomes a summative evaluation and an opportunity is missed by the evaluator to promote growth in the craft of teaching.

According to Koops and Winsor (2006) observations are most valuable when they can be used as a formative tool to collect data to aid in teacher growth and improvement. Unfortunately this is often not the case. Often these observations are viewed as spot-checks on a teacher's basic competencies, thus providing little data about the teacher's compliance with curriculum or prescribed teaching methods. Teachers view these checks as mostly symbolic, done primarily to fill a contractual duty rather than to help teachers do a better job (Cooper et al., 2005).

Marzano (2012) suggested, in order to combat this problem, administrators must spend more time conducting observations. The current trend is to do unannounced, random observations. Some districts now require administrators to conduct about five of this type of observation on each teacher. These observations are short, walk-through

observations during which an administrator observes in a teacher's classroom for about three to five minutes (Marzano, 2012).

Koops and Winsor (2006) explained that multiple short visits allow an administrator to observe a variety of lesson types. This creates a greater opportunity for follow up conversations with teachers regarding what the students are doing in the classroom along with the teacher's goals. Used along with a more formal observation, this process allows the teacher and administrator to evaluate not only teaching methods and materials but also students' response to those items. While an exact number of observations required to meet the needs of teachers and evaluators was not uncovered after reading much literature, Marzano (2012) stated that far more than five observations would be needed to obtain an accurate representation of a teacher's true pedagogical skill.

Another benefit to these short, more frequent observations would be they address an earlier mentioned barrier to teacher observation in the observer effect. McIntyre (1980) addressed the effects an observer had on teacher and student performance. He believed more frequent visits by the observer would diminish the observer effects on the teacher as well as the students, therefore providing more accurate data of what is going on in the classroom on a day-to-day basis. McIntyre conducted a study that had observers visit a classroom every Monday morning for eight weeks. During the first four weeks, the presence of the observer influenced the verbal behavior more than it did during the last four weeks. This direction of change provided evidence that the observer effect diminished over a period of time (McIntyre, 1980).

Contrary to this, a study conducted by Ebmeier (2003) presented information concluding the frequency of observations is a positive thing when accompanied by other

factors. Ebmeier stated while the frequency of classroom observations is essential to teaching, it in itself does not directly influence a teacher's confidence, trust, or support of the principal. Their research on the affective reactions to principal supervision showed a teacher's level of confidence, commitment, and satisfaction are based on the teacher's belief in the principal's interest in and commitment to supporting teaching. These frequent classroom visits must be accompanied by other activities that focus on classroom teaching, such as rewarding sound teaching, providing technical and symbolic leadership, and placing greater emphasis on teaching.

It is a concern of Goe et al (2008) that observation protocols are often generic and do not take into account subject specificity. This can be an important missing piece as teachers try to present increasingly ambitious content. There is also some question about the validity of the protocols. They are often used in research only by the researcher who created them. Because protocols have not been used to improve practice, it is unknown whether a district can expect to see change in teachers' practice due to that protocol's implementation (Goe et al., 2008).

Taking all of this into consideration, it is also important to consider an evaluation, or observation has no meaning if it is not interpreted, questioned, discussed, and provides reflection. These processes, ultimately, lead to more effective decision making (Feeney, 2007). Quality feedback is an essential part of the observation process and should support teachers in making decisions about what to teach and about how to better help their students. Novice teachers require meaningful and instructive feedback at multiple times during the school year to facilitate the change needed to accomplish school or district goals (Roberson & Roberson, 2009). Murphy, Hallinger, and Heck (2013) stated

school leaders will be more likely to promote instructional quality if they direct their efforts in facilitative channels, one of which is providing actionable feedback.

School leaders need to be trained in how to provide feedback that investigates instructional change and how to implement improvement plans for struggling teachers (Range et al., 2012). Waring (2013) made the following suggestions for promoting reflection during the feedback stage. He recommended (a) talking less and being less directive, (b) withholding value judgments or unsolicited feedback, (c) asking questions such as ‘how do you think the lesson went?’, and (d) making open-ended statements about some aspects of teaching.

Hattie (2012) explained the best way to understand feedback is to consider the idea of the gap: feedback aims to reduce the gap between where someone is and where they should be. Reducing this gap would move someone’s current achievement closer to success. Hattie stated feedback can do this in various ways: it can provide cues that get a person’s attention and help them focus on what is needed to succeed; it can direct attention toward the steps needed to accomplish the task; it can provide information about previously misunderstood ideas; and it can be motivational so more effort and skill is invested in a task.

Looking at feedback in direct relationship to teaching, Feeney (2007) stated, “The feedback provided to teachers should be descriptive and based on what the teacher and students actually do, supported by evidence of student outcomes throughout the lesson” (p. 195). Feeney explained these authentic examples allow the teacher to replay events and evaluate the lessons effectiveness. This type of feedback is more meaningful to the

teacher because he or she just taught the lesson. Providing this kind of descriptive feedback can have a large impact on professional growth.

Preparing Units of Instruction

Curriculum materials are directly connected to the day-to-day work of educators. Teachers use these resources to help them make decisions about classroom practices. These resources help teachers define what students need to know and be able to do as a result of instruction (Beyer & Davis, 2009). Using these tools, and other professional judgments, help teachers to design units of instruction.

While most evaluation processes require teachers to write up a detailed lesson plan, few require any evaluation of a unit plan. Marshall (2009) believed this is a detriment to the education process because even a high-quality lesson plan can miss the bigger picture. By evaluating only one piece, a principal is only getting a glimpse of what is happening in the classroom. A unit-plan, which describes a teacher's plan for teaching concepts and skills over a three- to five-week period, tells more about the coherence and alignment of instruction. Marshall stated a broader unit plan puts individual lesson plans into perspective.

Teachers are held to high standards and are governed by stated education departments. It is often overwhelming to try to cover all of the standards required of them. Tomlinson (2001) created nine guidelines to ensure that these mandated standards are integrated with instructional best practices.

1. Reflect on the Purpose: Educators must know what information students must learn in order to become competent in a certain discipline.

2. Plan Curriculum to Include All Facets of Learning: Curriculum will include essential facts and vocabulary to be mastered in a unit of study. It will cover the basic concepts of the discipline and then show students how they will master those concepts.
3. Plan to Help Students Make Sense of Things: This ensures students make sense of the knowledge, the discipline, and life.
4. Organize Content so it is Manageable: Curriculum should come from a clear framework that helps all see the big picture of what is being taught and learned.
5. Design Instruction so that Learning is Invitational: “Learning should be joyful for young learners – or at the very least, intriguing and satisfying” (p. 43).
6. Design Instruction for Action: Learning experience should be focused on what students are to know, understand, and be able to do.
7. Design Instruction to Attend to Student Differences: Effective instruction should provide for variance in student readiness, culture, gender and interests.
8. Work for Safe, Respectful, and Trustworthy Environments: The most powerful messages come to students from the atmosphere of the classroom.
9. Teach for Success: “Most students can learn and do far more than educators believe they can.” (p. 46)

These guidelines point out the importance of the ability of a teacher to craft a plan-of-action when trying to implement standards required of them while integrating those into what they know about teaching and learning. Evaluating this skill is an important piece to teacher evaluation systems.

Kunzman (2003) conducted a study to identify the usefulness of formal teacher preparations for those teachers with prior experience. One of the areas he identified as an area of need for those teachers was designing curricula that aligned with learning and assessment goals. He found that teachers without this comprehensive vision were not able to conduct thoughtful unit planning but also lacked in the ability to address and assess skills and knowledge over the course of the year. He found "...too many students were not learning what was expected, and the lack of a coherent and comprehensive approach to curriculum made monitoring that learning and addressing its deficiencies tremendously difficult" (p. 246).

Teacher Professional Development

A large amount of research has indicated the importance of teacher professional development (PD) on the improvement of schools (Borko, 2004; Fenwick, 2004; Kalule & Bouchamma, 2013; McQuarrie & Wood, 1991; Payne & Wolfson, 2000; Wanzare, 2002). Professional development opportunities can be defined as, "opportunities that will help teachers enhance their knowledge and develop new instructional practices" (Borko, 2004, p. 3). In addition, Smith and Rowley (2005) believed PD is more than just a tool for improving the quality of teachers; it can also improve a teacher's commitment to the profession and reduce turnover. Participation in PD is an indicator of a teacher's commitment to the profession. Placing these avenues of thought together, the aim of PD is to bring about change in how teachers teach, in their beliefs and attitudes, and ultimately create increased student achievement (Kalule & Bouchamma, 2013).

The evaluation of teachers is directed at the improvement of teaching and learning in schools. "Teachers' professional learning has long been linked to educational reform"

(Fenwick, 2004, p. 259). Instructional supervision that supports the PD of teachers will contribute greatly to this effort (Kalule & Bouchamma, 2013). Wanzare (2002) believed teacher appraisal should provide directions for teacher development, provide the opportunity for teachers to develop new skills, and the outcome of the appraisal should inform further teacher development.

The role of the principal is to facilitate growth, development, and the success of new teacher or incompetent teachers. PD is the primary means for deepening content knowledge for these new or struggling teachers (Smith & Rowley, 2005). Intensive PD programs have the power to increase the knowledge base of teachers and improve their instructional strategies (Borko, 2004). According to Range et al. (2012), the most common strategy used with incompetent teachers is attempting remediation through PD plans that address areas needing improvement. These PD plans target specific skills identified by the teacher and principal as deficiencies and work towards remediation of the problem area. By providing information of varying complexity, the knowledge base of the teacher is increased incrementally at a natural developmental pace for the teacher (Roberson & Roberson, 2009).

Joyce and Showers (2002) presented four elements that must be present if staff development is to significantly affect student learning. The first is that a community of educators comes together to study together, put into practice what they learn, and to share the results. The second piece of advice is content of the staff development should center around curricular and instructional strategies that produce a high rate of student learning. Third, they suggested that the amount of change is significant enough that the student's gain in skills and knowledge is obvious. Lastly, staff development enables educators to

gain skills to implement what they have learned. The authors make the statement, "...we make the case that staff development and student achievement are crucially, causally linked and that the knowledge exists for designing and implementing programs that make a difference in the lives of students" (Joyce & Showers, 2002, p. 10).

In schools where teacher learning flourished, the link that tied them all together appeared to be a visible commitment to learning across all levels (Fenwick, 2004). This commitment was manifest by the time and psychological support provided by the administrator in regard to staff learning along with building collegial relationships. It also involved fostering a positive, creative environment where members felt accepted and trusted. An administrator's willingness to tackle tough issues and dilemmas of learning also added to the development of teacher trust (Fenwick, 2004). These things work together to create a culture of shared learning in a school.

For adults, learning looks different than it does for students. Marquardt (2011) suggested learning for adults must take place almost as a byproduct of work. According to Marquardt, work and learning are increasingly becoming one in the same. Learning is at the heart of any true productive activity. When knowledge work is performed, learning is taking place. "...you must learn minute by minute if you are to perform knowledge work effectively" (Marquardt, 2011, p. 13). Learning is becoming a life-long process and challenge for those working in education.

This learning comes in three types of learning (Marquardt, 2011). Adaptive learning happens when we reflect on past events and then change future actions based on those experiences. Anticipatory learning is the process of gaining knowledge from anticipating what could happen in the future. This type of learning seeks to avoid

negative things from occurring and plans for the best possible outcome. The last type of learning is action learning. This type of learning occurs while engaged in action. It is on-the-spot learning, based on a problem and relies on reflective inquiry (Marquardt, 2011). These types of learning are of high value to a learning organization.

In regard to individual learning, Koops and Winsor (2006) shared their conviction that every teacher deserves to have a professional growth plan. Many support the idea that this plan should be individual in nature, catering to a teacher's specific needs (Borko, 2004; Fenwick, 2004; Khan et al., 2009; Koops & Winsor, 2006; McQuarrie & Wood, 1991; Payne & Wolfson, 2000). Payne and Wolfson (2000) identified some ways in which principals can ensure staff receive the PD they need. First, they recommend each educator create a personal learning plan. This plan addresses his or her need to grow, stressing knowledge and skills related to improved student learning. When a need for improvement is identified in a school district, often a one-shot inservice is designed for the entire district. This type of workshop is generally ineffective and fails to address specified needs of individual teachers. Teachers, like all learners, require guided practice and feedback to refine new practices (McQuarrie & Wood, 1991).

Payne and Wolfson (2000) suggested principals hold faculty accountable for personal learning goals that are aligned with school strategic plans. This can be by designing individual growth plans that define a teacher's own goals for professional growth (Fenwick, 2004). Using artifacts such as instructional plans, lesson observations, and samples of student work bring the classroom into the professional development planning. Such artifacts enable teachers to examine their instructional strategies and

pinpoint ideas for improvement (Borko, 2004). Because individually created growth plans are teacher-created, they appear to honor professional autonomy (Fenwick, 2004).

Designing a PD plan with individuals gives supervisors opportunities to reinforce the strengths of teachers and make recommendations that can help teachers to develop new strategies (Koops & Winsor, 2006). Principals have to encourage teachers to acquire new skills, support them through frustrations, and recognize their efforts. Data should be gathered to assess the impact of staff development initiatives, and principals should publically celebrate indicators of improvement in order to sustain those initiatives (Khan et al., 2009).

Evaluation by Students

Can evaluations of teachers by students be effective? This type of evaluation can be traced to the middle ages, but the earliest formalized gathering of data seems to have begun in the 1920s when the University of Washington and Harvard students began publishing results from numerous evaluations to prospective students. While this became a relatively common practice for universities, few secondary or elementary schools felt it a necessary function until recently (Eastridge, 1976). Danielson and McGreal (2000) explained that some evaluation systems are based on the belief that an educator's skill may be viewed from multiple perspectives. Adding feedback from students can offer insight from another perspective beyond that of the principal or evaluator.

Stiggins (1986) stated there may be no more valid source of information about the learning environment than from the students who work in those environments. He believed that students can provide insights that no one else can offer. Ripley (2012) supported this by saying research has shown that if kids are asked the right questions,

they could identify with accuracy their most, and least, effective teachers. It turned out their answers were more reliable than any other measure of teacher performance, including observations and student test-score growth.

Teachers who are serious about professional growth are deeply interested in how his or her teaching affects students (Stiggins, 1986). “Surveys can offer highly valuable insights into a teacher’s performance, and they can provide feedback to teachers that is unavailable from any other source” (Danielson & McGreal, 2000, p. 51). Marzano (2013) concluded that student surveys should play a part in an array of assessments used to judge teacher quality.

Students’ opinions of teachers are important because students have the most contact with the teacher and are the direct consumers of a teacher’s services. Goe et al. (2008) reviewed multiple studies done to measure the reliability and validity of student ratings. They concluded students are the most direct clients of teachers, thus, they have deeper experiences with teachers. These experiences give the students more to draw from when completing a rating of the teacher. While there are still some validity concerns due to the possibility of an opinion on one trait influencing their ratings in other areas, this did not seem to affect student raters any more than it did adult raters.

Danielson and McGreal (2000) warned that evaluations should not solely rely on these students’ perceptions, because they are just that—perceptions. Goe et al. (2008) also cautioned that student ratings should not be used as stand-alone rating systems. Students are not qualified to rate teachers on curriculum, classroom management, content knowledge, or collegiality. Information gathered from students can be used along with other sources of information to paint an inclusive picture of the teacher’s overall

performance. Marzano (2013) recommended this data be collected in the form of a pre-test and post-test. This would allow for a growth measurement to be considered. Some, however, use it only as a post-test.

Existing frameworks, authored by various people, are available that include components those specific authors feel would provide an effective evaluation. In addition to established frameworks, research was presented on individual aspects, or pieces, of evaluations and their effectiveness. Classroom observations and feedback, teacher professional development, preparing units of instruction, and student evaluations are all included in the next framework presented in this chapter.

Marshall Framework

The conceptual framework for this paper was based on teacher supervision and evaluation as presented by Marshall (2009). His is a singular approach which addresses all of the above aspects but in a more narrow fashion. He takes pieces from each component (historical, teacher effect on student achievement, obstacles and barriers, and other theories on best practices) and outlines his framework. The framework presented includes four parts that a leader should implement in an effective teacher evaluation system.

The first piece is that of mini-observations. Marshall (2009) presented the idea of visiting teachers' classrooms frequently in small chunks of time. He believed that principals who make announced visits one or two times per year do not have a true sense of the quality of instruction that students receive day after day, month after month. By conducting evaluations in this manner, principals are trusting teachers to deliver the same quality of instruction all the time.

To address this problem, Marshall (2009) presented the idea of mini-observations. “The basic idea of a mini-observation is simple: the principal makes frequent, unannounced classroom visits and gives prompt feedback to teachers, and as a result, teaching and learning improve” (Marshall, 2009, p. 65). He realized these visits would have to be brief, but not so brief that he would not be able to focus on what was really happening. He also knew frequent visits would provide the best opportunity to offer praise along with criticism. Lastly, he recognized that visits would need to be unannounced to get a picture of every-day reality.

An important aspect of the mini-observation process is the feedback that teachers receive after each visit. Marshall (2009) knew that feedback would need to follow each visit so teachers were not left guessing what was thought of the lesson. He also felt these needed to be face-to-face conversations that would be delivered in a non-evaluative, non-threatening way. This would provide an opportunity for two-way communication through which the teachers felt comfortable participating. “Ideally, teachers leave feedback conversations with specific ideas for improving their practice—or a warm feeling that their work is appreciated by an intelligent and thoughtful colleague” (Marshall, 2009, p. 81).

According to Marshall (2009), feedback usually comes in one or more of the following ways: praise, reinforcement, suggestions, or criticism. Praise points out something the teacher has done well. Marshall stated for praise to be most effective, evaluators need to keep the emphasis on the positive outcome for students caused by the preferred behavior. Using statements such as, “I liked...” would put the focus on pleasing the boss rather than impacting students. Reinforcement comes in the form of

commenting on behaviors that are improvements resulted from earlier criticisms. This can occur following several observations after which constructive criticism has been offered. When corrected behavior is noticed, reinforcement helps to cement the preferred behavior in place.

Suggestions can be made when the administrator can give ideas about conveying content knowledge or classroom management. Administrators have their own personal experiences to draw from as well as what they observe from other classrooms. This information can be shared through suggestions so tried-and-true ideas can benefit all students. Last, feedback can include criticism. When problems are noted during mini-observations it is important to discuss those problems with the teacher. A criticism should be followed by a suggestion for improvement (Marshall, 2009).

The second aspect to Marshall's (2009) plan is that of monitoring and evaluating curriculum unit planning. While one detailed lesson plan is a fundamental building block for the classroom, it is only a small part of a teacher's efforts to convey knowledge and skills. A principal needs more information to assess the bigger picture. "What curriculum unit is this lesson part of? What are the unit's big ideas and essential questions? How does this unit align with state standards? How will students be assessed?" (Marshall, 2009, p. 26). All too often, an evaluation is overly focused on the one lesson that was observed and the bigger picture is unseen.

The third aspect is monitoring and evaluating interim assessments. Marshall (2009) believed that by monitoring interim assessments, student learning is placed at the center of conversations with teachers and teacher teams. Principals can do this by putting into place an effective interim assessment monitoring process through professional

learning community meetings. During these meetings, conversations are ongoing about real-time student achievement results, a continuous study of evidence of student learning.

Interim assessments are designed to provide information about student learning on a regular basis. They often mirror state tests. Some assessments are written by teachers, others are purchased from companies. If done correctly, studying interim assessments can produce high levels of achievement (Marshall, 2009).

The last piece to the system is that of end-of-year rubric evaluations (Marshall, 2009). These rubrics cover every aspect of the teacher's professional work and are used as a summative type evaluation to give the teacher feedback on their overall performance. Marshall (2009) believed these are best used to pull together all impressions gathered over the course of the school year. These rubrics can also improve teaching and learning in five distinct ways.

The first of the five ways in which the rubrics can improve teaching and learning is spotlighting aspects from the rubric that encourage professional learning community activities. Secondly, the rubric can be used to encourage teachers to set personal, measurable goals for improvement. Teachers can also be encouraged to fill out the rubrics for themselves, providing self-assessment. Studying input from advisory committees, or students and parents, can give teachers an idea about the quality of their customer service. This information can inform their practice and help them improve. Lastly, charting data from rubrics for the entire staff can help principals make decisions regarding professional development.

Marshall (2009) believed that by implementing these four components, "principals can build collaboration with and among teachers and foster continuous

improvement of teaching and learning aimed at closing the achievement gap” (p. xx). But for this to happen, evaluators have to rethink conventional teacher evaluation systems. Principals need to go beyond classroom inspection and supervise curriculum planning, interim assessment analysis, and follow-up. Through all this, student learning should be kept at the center of all conversations. With this expanded wealth of data, mini-observations can be used to spot-check classrooms and prompt meaningful conversations about what is observed. If a principal has a firm grasp on all these elements, the ultimate goal can be obtained and “all students graduating with the skills and knowledge they need to take the next step toward college and career success” (Marshall, 2009, p. 207).

The history of teacher evaluation systems presents theories and practices that have been tried and tested over time. Research found to emphasize the importance of teacher effectiveness on student achievement highlighted the importance of teacher evaluation and its relationship to student success. Many barriers and obstacles to conducting effective teacher evaluations have been identified. However, best practice for conducting teacher evaluations is also supported in research. This chapter of the paper presented research regarding all of these areas, ending with the Marshall framework that served as the conceptual framework for this research study. Chapter three will present the research design and methodology used in this study.

CHAPTER THREE

RESEARCH DESIGN AND METHODOLOGY

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. In January of 2002, the *No Child Left Behind* (*NCLB*) Act was implemented to provide an accountability system for all schools in the nation. Under this act, schools were required to have all students scoring at proficient or advanced levels on state assessments by the year 2014. As 2014 rapidly approached, many school districts found themselves in failing categories due to their inability to meet adequate yearly progress toward this goal.

Several states began applying for waivers to remove themselves from the accountability put in place by *NCLB*. In June of 2012, Missouri received approval of their waiver submitted to the United States Department of Education. Upon approval of this waiver, the state had to create its own accountability system to better identify schools in need of improvement as well as excelling schools. One specific guideline outlined in the waiver was the creation of a new teacher evaluation system. "If we can identify what those highly effective teachers do, then even more of the differences in student achievement can be accounted for" (Marzano et al., 2001, p. 3). This specific guideline, meant to improve the quality of teachers, led to the creation of the Network for Educator Effectiveness (NEE) system.

Seeing the need districts had for a new teacher evaluation system, the NEE system was created by partnering departments of the University of Missouri. The system was

created to provide districts with the components of an effective evaluation system, along with an online data storage site to house all of the data collected on teachers. The goal of this teacher evaluation system was to promote teacher growth and change in teacher practice.

Purpose of the Study

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. The study collected their perceptions on the impact NEE had on their teaching, specifically changes made. This information served as an evaluation of the program and will be used to guide its further development. It will also inform other states as they develop new teacher evaluation systems during this time of high accountability under *No Child Left Behind*. Lastly, the research will benefit school leaders as they implement NEE with educators under their supervision.

Research Questions

The research questions which guided this study were as follows:

1. What are teacher's perceptions of the change impacted by the Network for Educator Effectiveness based on the framework presented by Marshall?
 - a. What changes have been made in teaching as a result of the mini-observations?
 - b. What changes have been made in teaching as a result of the unit of instruction evaluation?
 - c. What changes have been made in teaching as a result of the professional development plan?

- d. What changes have been made in teaching as a result of analyzing the student survey data?

Design for the Study

Rossi et al. (2004) defined program evaluation as “the use of social research procedures to systematically investigate the effectiveness of social intervention programs” (p. 16). They additionally said that this social research is intended to improve social conditions. A valid description of the performance of a program includes a correct representation of what the program truly accomplishes. According to Rossi et al., program evaluations usually involve one or more of five program areas: need for the program, design of the program, program implementation and service delivery, program impact or outcomes, and program efficiency.

This research project was a program evaluation to measure teachers’ perceptions of the change that has occurred in their instruction due to the implementation of the NEE system. It measured the program’s impact, or outcomes, based on the teachers’ views of how it changed their instruction and made them better teachers. According to Rossi et al. (2004), a program’s impact theory consists of assumptions about the change that has occurred which was caused by the program and the improved conditions that are expected as a result. This study sought to do this by informing those creating and implementing the program about how effective the system has been in creating change in teacher performance.

When conducting a utilization focused evaluation, the researcher should recognize that the findings of the evaluation will be used by key decision makers to inform their day-to-day management or broader funding or policy decision making (Rossi et al.,

2004). Patton (2008) claimed identifying the primary intended users of the evaluation, the researcher can make better decisions about what to include in the evaluation. These primary intended users are referred to as stakeholders.

This qualitative case study collected data to be analyzed to determine if the Network for Educator Effectiveness created change in teacher behavior, thereby improving their performance. A case study consists of “inquiry in which the researcher explores in depth a program, event, activity, process, or one or more individuals” (Creswell, 2009, p. 13). A qualitative multi-site case study allowed the researcher to collect data through focus groups, and surveys. The data revealed changes, or lack thereof, in teacher behavior due to the implementation of the NEE.

Participants and Sampling Procedures

This qualitative multi-site case study focused on two school districts that participated for one year as a pilot school for the NEE system. During the study, the school districts were in their second year of implementation. These two districts served as a sample of teachers from a focused area of the state. This study did not include teachers from across the state in an effort to narrow the scope of the study. According to Field (2009), researchers collect data from a small subset of a population and then use that information to infer things about the whole population.

A maximum variation sample, as introduced by Hatch (2002), was used within these two districts. In this type of sample, participants are chosen based on differences in characteristics. While this research included all teachers within two school districts, it sought the maximum variation of the type of teachers included: male and female,

elementary through high school level, experienced and non-experienced, and teachers from all subject areas.

These districts were selected as a convenience sample due to the availability to the researcher and their participation in the NEE system. Hatch (2002) defined a convenience sample as individuals who are easily accessible to the researcher. This accessibility was due to the proximity of the districts in relation to the researcher's home district, as well as personal connections with several of the administrators in both districts.

School district A is located in the southwest portion of the state. This district is comprised of four schools that serve 1,394 students. They have a ratio of 14 students per full time equivalent teacher. Their dropout rate among 9-12 graders is 2%. Ten percent of their population is comprised of students identified in need of an individualized education plan.

School district B is also located in southwest portion of the state. This district is comprised of five schools that serve 3,191 students. They have a ratio of 15 students per full time equivalent teacher. Their dropout rate among 9-12 graders is 2%. Thirteen percent of their population is comprised of students identified in need of an individualized education plan.

To gain access to the sample districts, the researcher first made contact with the superintendents, who serve as the key gatekeeper for the district. According to Hatch (2002), the gatekeeper is the person who formally or informally controls access to the setting. In the case of most public schools, the superintendent is the key gate keeper. The researcher made an appointment to meet face-to-face with the superintendent to

explain the research project and what it would require of the district should they participate and to answer any questions. Permission was requested at this meeting, along with permission to meet with each of the building principals.

In the meeting held with each of the building principals, the project was again described and information was gathered on the teachers in those respective buildings. The focus groups were created using a purposeful selection (Creswell, 2009) in which participants were selected by the principals that would best help the researcher understand the problem or research questions. In this case, to ensure a maximum variation sample, teachers were selected to represent different grade levels, different departments, different experience levels, and different genders. This was done by each of the building principals. Two focus groups containing eight participants were formed within each district, one including elementary level teachers and one including secondary level teachers. This created a total of four focus groups, with 32 participants altogether. These participants were then invited to participate through email and supplied with an informed letter of consent. For the remaining piece of data, a survey was sent to all teachers in the district via email. Surveys were distributed to a total of 311 certified teachers. Informed consent was included with consent obtained through completion of the survey.

Data Collection

This research included data collected from focus groups and surveys. By using these two tools, the researcher could glean a large enough sample of qualitative data to make connections and identify recurring themes. This section of the paper will outline

the data collection methods used, along with the considerations for human subject protection and ethical considerations.

Data collection procedures. Rossi et al. (2004) contended focus groups are used to get information from a number of people efficiently. Krueger and Casey (2009) indicated focus groups are effective if “you are looking for the range of ideas or feelings that people have about something” (p. 19). They also stated that focus groups are purposed to promote self-disclosure among those participating. In order to do that the groups should include people with something in common. This tie causes participants to self-disclose more when they feel they are similar to other participants. Krueger and Casey also recommended creating a setting that is comfortable and non-judgmental because participants will disclose more in a non-threatening environment.

Two focus groups containing eight participants were formed within each district, one including elementary level teachers and one including secondary level teachers. This created a total of four focus groups, with 32 participants altogether. Focus groups were about one hour in length and were guided by the focus group protocol (See Appendix A).

A large net was cast to collect data from a large group through an open-ended question survey. Fink (2009) stated, “Surveys are information-collection methods used to describe, compare, or explain individual and societal knowledge, feelings, values, preferences, and behavior” (p. 1). They can be used when researchers want to evaluate a program’s effectiveness to change people’s knowledge or attitudes. The survey gave participants the anonymous freedom to express themselves without feeling any judgment. Being able to submit responses to a survey anonymously will yield true data.

Concerns about return rate prompted the researcher to include focus groups along with a survey. Surveys were sent to all teachers employed in the district for a total of 311. The time requirement for their completion was approximately 20 minutes. The survey included demographic questions. It also included four Likert scale questions inquiring about the extent to which change had occurred due to the implementation of each piece of NEE. These Likert questions were followed by an open-ended question requesting the respondent list the changes made due to that implementation. (See Appendix C)

Human subjects protection and ethical considerations. When involving human subjects in a research project, researchers should gain approval from their local Institutional Review Board (IRB). The purpose of the IRB is to ensure that research “is done with ethical regard to the rights and welfare of human participants” (Seidman, 2006, p. 59). The researcher submitted a proposal and all appropriate documents for approval from IRB before beginning the research.

Prior to each focus group, participants were given an informed letter of consent which explained the research project and requested their participation (See Appendix B). Informed consent was also provided to participants conducting the survey (See Appendix D). Subjects were given the right to refuse participation or to stop participation at any time. No identifying information was gathered from participants in order to protect their identity. Some demographic information was collected for the purpose of identifying the population sampled, but the information was kept secure to ensure confidentiality of participants’ responses. Participants were informed of any risks and benefits that could result from taking part in this project.

Data Analysis

The qualitative data were studied using an interpretive analysis. Hatch (2002) stated that interpretive analysis is “about giving meaning to data” (p. 180). He suggested that researchers read the data through to get a sense of the whole. After this step, read the data to identify impressions and record those. Next, the data should be reread while the researchers coded the data where themes exist. Creswell (2009) recommended the researcher read through all data collected, looking for what each piece is about and making notes in the margins. Once each piece of data has been analyzed this way, the researcher can use the notes to find any similarities that can be placed together to create themes. These themes are then given codes. Once these have been established, the researcher can return to each piece, marking applicable information with the codes. At this point, a draft summary can be written to bring the pieces together to make meaning of the themes in the data.

Role of the Researcher

In a qualitative study, the researcher is the key instrument according to Creswell (2009). The researcher may use an instrument, such as a protocol, but the researcher is the one actually collecting the data. This makes the researcher’s positionality and bias an important factor in a research project. The position taken by the researcher for this project was the constructivist paradigm. Constructivists tend to view reality as unknowable (Hatch, 2002). Their viewpoint on inquiry sees it as made of individual perspectives or constructions of reality. This positionality had an effect on this project in that it guided the type of study, types of data collection, and type of data analysis completed.

“Because the researcher is the instrument in qualitative inquiry, a qualitative report must include information about the researcher” (Patton, 1999, p. 1198). Bias that may affect the study include the researcher’s role as an elementary principal along with current participation in the NEE system. As an elementary principal, consideration was given to not focus on only elementary information. All data were examined equally and included in the research to include the maximum variation needed. Another key factor was the current participation in the NEE system in the researcher’s home school district. Information gathered from personal experience was not included in this study so results are a true representation of what is found in the two southwest Missouri school districts being examined for this case study.

Trustworthiness

Establishing validity, or trustworthiness, is central to qualitative research (Roulston, 2010). According to Roulston, showing evidence credibility by looking closely at the sources and procedures by which it was gathered is important for qualitative studies. The researcher disclosed as much information as possible about how the data were gathered and where it came from. Participants were allowed to review the data collected to ensure correctness.

To ensure confirmability and transferability, the researcher attempted to include enough detail and rich information from the data collected to allow the reader to make a judgment of the findings. Accounts should “include enough contextual detail and sufficient representation of the voices of the participants that readers can place themselves in the shoes of the participants at some level and judge the quality of the findings” (Hatch, 2002, p. 16). This rich description of the data will allow the reader to

confirm the validity of the findings and transfer the information gained to other possible areas of practice.

Limitations and Assumptions

Limitations and assumptions are a part of every research project. Limitations can affect the quality of research or impact its credibility. It is also important for the researcher to explore their own assumptions about the research, but this is not an easy task (Hatch, 2002). Assumptions, by nature, are generally unexamined because they can be personal views held by the researcher. This requires the researcher to reflectively consider all possible assumptions.

Limitations

While “qualitative researchers argue that no direct relationship exists between the number of participants and the quality of the study” (Hatch, 2002, p. 48), there was concern that the study only collected data from two school districts. The NEE system is being implemented across the state, but for the purpose of narrowing the research only two schools were selected. To address this concern, the researcher attempted to use multiple forms of data collection and made sure to obtain a maximum variation sampling.

Rossi et al. (2004) stated that at various times policy makers and program managers need to distinguish worthwhile programs from ineffective ones and launch new ones or revise existing ones to achieve desirable results. Considering this, the newness of the NEE system makes it a prime candidate for an evaluation. However, this could also be viewed as a limitation due to the possibility of participants responses skewed by the initial implementation of a new program.

Assumptions

Three major assumptions were identified for the methodology portion of the project. The first assumption being the NEE system was being fully and correctly implemented in the two districts studied. According to Rossi et al. (2004), implementation failure occurs when outcomes are poor because the necessary activities thought to bring about improvements did not occur. Implementation was discussed with the superintendent and the building principals during the initial meetings to ensure full implementation. By sampling two school districts, the researcher hoped to combat this concern.

The second assumption questions the transferability of the findings if the voices of the teachers from these two districts do not mirror those of teachers across the state. The researcher understands that there could be other factors going on in the districts to affect the teachers' attitudes and opinions about the evaluation system. By ensuring a maximum variation sample, the researcher hoped to gather enough data from a large enough variation of people to overcome any personal attitudes or biases.

Lastly, the researcher assumed any changes in teacher behavior found in the data did result from the NEE system. It should be considered that the districts have other professional development systems in place that may also affect teacher change. Questions were structured to attempt to keep the teachers' focus on change that occurred due to the NEE system and not from other means.

Summary

This qualitative case study was conducted to measure the perceptions of teachers in two small southwest Missouri school districts regarding the NEE teacher evaluation

system. Data were collected in regard to changes that have occurred in their teaching following the implementation of the program. The utilization focused program evaluation will guide the creators of the new evaluation system in their further development of the system.

Participants were selected from two school districts that are involved in the NEE system for the second year. A maximum variation sample was selected to represent all types of teachers from all levels and subject. Focus groups and surveys collected the data needed for the researcher to make conclusions about the effectiveness of this new program. Several measures were taken to ensure the validity of the study, including human subject protection and other ethical considerations.

CHAPTER FOUR
ANALYSIS OF DATA

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. The study collected their perceptions on the impact NEE had on their teaching, specifically changes made. The research questions which guided this study were as follows:

1. What are teacher's perceptions of the change impacted by the Network for Educator Effectiveness based on the framework presented by Marshall?
 - a. What changes have been made in teaching as a result of the mini-observations?
 - b. What changes have been made in teaching as a result of the unit of instruction evaluation?
 - c. What changes have been made in teaching as a result of the professional development plan?
 - d. What changes have been made in teaching as a result of analyzing the student survey data?

This qualitative multi-site case study focused on two school districts that participated for one year as a pilot school for the NEE system. During the study, the school districts were in their second year of implementation. These two districts served as a sample of teachers from a focused area of the state. This research included data collected from

focus groups and surveys. By using these two tools, the researcher could glean a large enough sample of qualitative data to make connections and identify recurring themes.

Demographic Information

A maximum variation sample, as introduced by Hatch (2002), was used within the two districts that participated in this research. In this type of sample, participants were chosen based on differences in characteristics. The research included all teachers within two school districts, it sought the maximum variation of the type of teachers included: male and female, elementary through high school level, experienced and non-experienced, and teachers from all subject areas. Demographic data were collected from the survey and the focus group interview to ensure a maximum variation sample was collected.

Survey Demographics

The survey was distributed to 311 certified teachers via email. Twenty-seven percent of those responded for a total of 84 responses. Twenty-seven responses, or 32%, were received from school district A. Fifty-six responses, or 67%, were received from school district B. One respondent, or one percent, chose to leave their district unidentified.

Table 1

Survey Responses by District

School District	Participants	Percent
School District A	27	32
School District B	56	67
Unidentified	1	1

Note. $N = 84$.

Level of experience was measured in the next demographic question of the survey. Respondents were asked to place themselves in one of three categories to identify their years of experience. They were given the option to select 1 – 5 years, 6 – 15 years, or 16+ years. Twenty-one respondents, or 25%, selected 1 – 5 years of experience. Forty-three respondents, or 51%, selected 6 – 15 years of experience. Twenty respondents, or 24%, selected 16+ years of experience.

Table 2

Years of Experience of Survey Participants

Years of Experience	Participants	Percent
1 – 5 Years Experience	21	25
6 – 15 Years Experience	43	51
16+ Years Experience	20	24

Note. $N = 84$.

Information was also collected regarding the level of students participants taught. Respondents were given the option to choose one or more of the following: Elementary (Pre-k – 5), Middle School/Junior High (6 – 8), or High School (9 – 12). Elementary teachers were represented by 36 respondents, or 43%. Middle school/junior high teachers were represented by 27 respondents, or 32%. High school teachers were represented by 19 respondents, or 23%. Two respondents chose multiple levels including both elementary and middle school/junior high, representing 2% of responses.

Table 3

Level Taught by Survey Participants

Level Taught	Participants	Percent
Elementary (Pre-k – 5)	36	43
Middle School/Junior High (6 – 8)	27	32
High School (9 – 12)	19	23
Multi-level	2	2

Note. $N = 84$.

The last piece of demographic data collected was in regard to gender. Sixty-nine of the 84 respondents identified themselves as female, making up 82% of the responses. Only fifteen respondents were male, accounting for 18% of responses. This combined demographic data provided a picture of the participants who contributed to this study.

Table 4

Gender of Survey Participants

Gender	Participants	Percent
Female	69	82
Male	15	18

Note. $N = 84$.

Focus Group Demographics

The focus groups were created using a purposeful selection (Creswell, 2009) in which participants were selected by the principals that would best help the researcher understand the problem or research questions. In this case, to ensure a maximum variation sample, teachers were selected to represent different grade levels, different departments, different experience levels, and different genders. This was done by each of the building principals. Two focus groups containing eight participants were formed within each district, one including elementary level teachers and one including secondary level teachers. This created a total of four focus groups, with 32 participants asked to participate. These participants were then invited to participate through email and supplied with an informed letter of consent.

The elementary focus group from district A had six participants in attendance while the secondary group from the same district had four. In district B, the elementary focus group had five participants and the secondary group had six participants. This created a total of 21 focus group participants for this study.

Table 5

Focus Group Participants by District

District	Participants	Percent
School District A	10	48
School District B	11	52

Note. $N = 21$.

The following tables express the demographic information of the focus group members including participants by district, years of experience, level taught, and gender. The first piece of demographic data identified the participants' level of experience in teaching. Four of the 21 participants' level of experience was in the 1 – 5 year range, making up 19% of the population 10 were in the 6 – 15 year range, or 48% of the population. And 7 were in the 16+ year category, completing the other 33% of the population. While a majority of the participants were in the mid-range of experience, it added to the study to have some new teachers as well as seasoned veteran teachers.

Table 6

Years of Experience of Focus Group Participants

Years of Experience	Participants	Percent
1 – 5 Years	4	19
6 – 15 Years	10	48
16+ Years	7	33

Note. N = 21.

The second piece of demographic data identified the level of students participants worked with. There were 11 participants, or 52%, who represented the elementary (Pre-k – 5) category, four, or 19%, represented the middle school/junior high (6 – 8) category, and six, or 29%, represented the high school (9 – 12) category. While it was an almost even split between elementary and secondary participants, the study could have gained more insight from the mid-level teachers (6-8) with greater participation from that group.

Table 7

Level Taught by Focus Group Participants

Level Taught	Participants	Percent
Elementary (Pre-k – 5)	11	52
Middle School/Junior High (6 – 8)	4	19
High School (9 – 12)	6	29

Note. N = 21.

The last piece of demographic data represents the gender of the focus group participants. Fourteen, or 67%, of the 21 participants were female. Seven, or 33%, of the focus group participants were male. The reason for this imbalance is unknown. The groups were chosen using a purposeful selection, but those that chose to participate did not balance in this demographic piece.

Table 8

Gender of Focus Group Participants

Gender	Participants	Percent
Female	14	67
Male	7	33

Note. N = 21.

Findings

After collecting the data, the researcher divided the analysis of the survey and focus group data into four parts in relationship to the research questions regarding the mini-observations, the unit of instruction evaluation, the professional development plan, and the analysis of student survey data. The Likert scale items from the survey were tabulated and the qualitative data were taken from the open-ended survey questions and focus group responses. After analyzing the survey and focus group data, the researcher was able to develop findings around each of the four research questions.

Research Question 1a: Mini-observations

Research question 1a asked, what changes have been made in teaching as a result of the mini-observations? Survey respondents were asked to rate the extent of change that has taken place in their instruction as a result of the mini-observations conducted as part of the NEE system. They could rate their amount of change according to the following scale: 0 – no change, 1 – some change, 2 – moderate change, and 3 – significant change. The results of the Likert scale items are represented in Table 9.

Table 9

Likert Responses to Amount of Change as a Result of Mini-observations

Likert Responses	Participants	Percent
0 – No change	30	36
1 – Some change	23	27
2 – Moderate change	22	26
3 – Significant change	7	9
No Response	2	2

Note. $N = 84$.

The qualitative data were studied using an interpretive analysis. Hatch (2002) suggested that researchers read the data through to get a sense of the whole. After this step, data should be read to identify and record impressions. Next, the data should be reread while the researcher coded the data where themes existed. While taking this approach, the researcher identified five themes that emerged from the data collected from the survey and the focus groups about the mini-observations. These themes included the increased use of technology, an increase in student engagement, increased use of assessment, improvement based on specific administrator/system recommendations, and explanations regarding no changes made in instruction.

Increased use of technology. There were nine comments in the survey and focus group responses related to the increased use of technology in classroom instruction.

Statements were made that technology was used more intentionally and in more

meaningful ways. Awareness was heightened regarding the use of other means to teach during lessons, including technology. One respondent stated, “I have added technology to everything, even when I do not think that is the best route to achieve a goal for a specific student. We are graded on use of technology, whether it fits or not.” The respondent went on to clarify that they felt their building administrator was able to identify when the use of technology is advantageous and when it is not, so they felt fortunate in that regard. It was also noted by one respondent that they did not feel using technology in every subject level was appropriate. A further negative was identified in that the administrator had made recommendations to increase the use of technology, but the teacher was frustrated because the district had not allocated resources for newer equipment.

Increase in student engagement. One teacher stated, “I put more thought into how to improve student engagement (use of manipulatives, discussion with peers, choral response, etc.)” The observations made teachers more aware of strategies to increase student engagement when planning lessons. Teachers referred to changing what students need to do when they finish work and are waiting for others to finish. There were comments teachers are doing more to make sure students know they are expected to do their job at all times, because they score lower points if the students are not engaged. Teachers stated they use more student led activities and give students greater opportunities to communicate with one another about the content they have learned in class.

One respondent stated small group instruction had increased and guided reading had taken the place of independent reading to a large degree in her classroom, thus

increasing student engagement. Changes also took place in higher level questioning and deeper class discussions. Three teachers alluded to focusing on questions that require more than a yes or no answer and requiring students to explain their thinking more often. Lastly, goal setting was referred to by multiple responses to help students stay focused and motivated.

Increased use of assessment. A teacher stated, “because I am evaluated on how many times I assess students and how many varieties of assessment are used, I increased the amount and variety of assessments I use.” Six combined survey respondents and focus group participants referred to increasing the use of assessment in everyday lessons. The mini-observations encouraged teachers to use more formative assessments during class time to assess students’ understanding. “Due to the walk-throughs, I have tried to implement as much formative assessment as I can into my lessons, and I try really hard to focus on the standards and goals that I am being evaluated on.”

Improvement based of specific administrator/system recommendations.

While this is a more general category, eleven of the focus group and survey respondents referred to making changes based on specific feedback received from their administrator. It was stated by one survey respondent that any area where scores were low became an area of focus for them. They shared that in a previous walk-through, they had received a low score in critical thinking. In turn, they were able to go back and adjust their teaching strategy and lesson accordingly. Other respondents referred to changes in instruction occurring due to specific suggestions made by their administrator as part of the comments/feedback portion of the mini-observations.

I try to take what my principal lets me know I can work on and integrate it into my lesson, usually as soon after our follow-up as possible. As a first year teacher, this feedback is worth its weight in gold.

The criteria and standards outlined in the NEE system itself have also affected change in instruction. Five teachers referred to including certain standards into their lessons because they knew they would be evaluated on them. They wanted to focus on these strategies because they would improve student learning.

The changes that I have made with my classroom instruction have been more conscious to include the different aspects of the NEE criteria that the district expects me to do. I believe that these changes have made a positive change in my instructional practices.

Explanations regarding no changes made. If teachers indicated that no changes in their instruction had taken place, they were asked to explain. A couple of responses stated that their teaching *is what it is*, regardless of the presence of an administrator. They did not feel the observations changed the way they taught because they were already doing the best they could on a day-to-day basis. One responded the feedback they had received from their administrator was inadequate to warrant any changes. They stated they were often just asked if the teacher had any questions, and if not, there was no follow-up discussion.

Another cited that the short amount of time spent in the classroom was not enough for the evaluator to really know what was happening. The teacher stated they had to explain why a certain activity was taking place because the administrator was not there long enough to get the whole picture. It was also noted any changes that had taken place

for one teacher were based on best practices, district professional development, and the needs of their students rather than any feedback from their administrator. Lastly, one teacher felt the mini-observations had not affected her instruction in a positive way. She felt frustrated because she could not do anything *fun* with her students anymore for fear that she might receive an observation.

Table 10

Summary of Themes for Research Question 1a: Mini-observations

Indicated Change	62%	Indicated No Change	36%
Increased use of technology		Explanations regarding no change:	
Increase in student engagement		-already doing their best	
Increased use of assessment		-inadequate feedback	
Improvement based on specific administrator/ System recommendations		-lack of time in classroom	

Note. 2% of participants made no response to Likert item.

Mini-observation effected change in 62% of survey and focus group participants. These changes were grouped into four themes. These themes included increased use of technology, increase in student engagement, increased use of assessment, and improvement based on specific administrator or system recommendations. The 36% that indicated no change in their instruction explained the lack of improvement was due to their belief that they were already doing their best, they received inadequate feedback from their administrator and observers did not spend enough time in their classroom. Two percent of participants did not respond to this item.

Research Question 1b: Unit of Instruction Evaluation

Research question 1b asked, what changes have been made in teaching as a result of the unit of instruction evaluation? Survey respondents were asked to rate the extent of change that has taken place in their instruction as a result of the unit of instruction evaluation conducted as part of the NEE system. They could rate their amount of change according to the following scale: 0 – no change, 1 – some change, 2 – moderate change, and 3 – significant change. The results of the Likert scale items are represented in Table 11.

Table 11

Likert Responses to Amount of Change as a Result of Unit of Instruction Evaluation

Likert Responses	Participants	Percent
0 – No Change	47	56
1 – Some Change	19	23
2 – Moderate Change	10	12
3 – Significant Change	3	4
No Response	5	5

Note. $N = 84$.

Again, the qualitative data were studied using an interpretive analysis to identify themes in the responses about the unit of instruction evaluation. The researcher identified four emerging themes in this portion of the data. The themes identified were regarding

increased rigor, greater collaboration, positive impact on daily instruction, and explanations regarding no changes being made.

Increased rigor. The survey responses indicated there had been an increase in rigor following the evaluation of the unit of instruction. It was noted that, as a grade level team, the rigor of assignments had been raised. There was more attention placed on hands on activities and creating more in-depth lessons. Lessons were modified to include questioning techniques and justification to increase the rigor of lessons. More emphasis was placed on students struggling to achieve knowledge while continuing to raise the rigor. One respondent stated they had been more aware of the depth-of-knowledge level at which they assessed students.

Greater collaboration. Several of the comments showed that collaboration had taken place in creating these units of instruction. There were references to grade level teams or content area teams. One respondent stated, “I’ve tried to work more with the other social studies teacher to have the same activities/lesson plan for specific areas we are evaluated on.” Another commented that their grade level team was trying to stay closer to what each of them is teaching, and they are conducting more assessments. Even when some teams had their curriculum in line, they were able to evaluate it and make some changes. The teams were able to work together to make decisions to improve lessons and increase rigor. On the negative side, frustration was shown by one respondent that the team was evaluated on one unit of instruction with all team members achieving the same score regardless of the amount of work individuals put into it.

Positive impact on daily instruction. Another theme that emerged was that of the impact the evaluation of the unit of instruction had on teachers’ daily instruction.

According to one respondent, “I look more closely at how my students are learning the information taught. I monitor their progress closely and make changes in their instruction based on their needs.” Teachers shared they added more movement and more interactive lessons because of this evaluation piece. For some, just knowing they were going to be evaluated caused them added pressure to make lessons exemplary in the “eyes of those who are observing.” It also prompted teachers to be more aware of teaching what is on their unit of instruction and not get off track. The guidelines set forth by the NEE system, along with administrator feedback, led some to be more conscious to include these aspects into their daily teaching.

Explanations regarding no changes made. The evaluation of a unit of instruction was referred to by two respondents as a “jumping through hoops” exercise. Frustration was expressed at the amount of time it took teachers to complete, thus taking time away from what they perceived to be more useful tasks.

I have been teaching for a long time. I know what I’m doing. I know the Course Level Expectations. I don’t need to waste my time writing out some big elaborate lesson plan. Besides, just because I can write out some big elaborate lesson plan, does not mean I am a competent teacher. All of this work I had to do on this unit of instruction just took away my precious time from doing things that are more important.

A focus group participant shared they feel, “we are too busy trying to prove we are doing our job, to do our job!” Another expressed the frustration that it took away the creativity and teachable moments they felt are a necessary part of teaching and learning. Some

stated they had not made any changes because their scores were adequate or they felt they did not need to change what they were already doing.

Table 12

Summary of Themes for Research Question 1b: Unit of Instruction Evaluation

Indicated Change	39%	Indicated No Change	56%
Increased rigor		Explanations regarding no change:	
Greater collaboration		-time consuming	
Positive impact on daily instruction		-lessened creativity	
		-saw no need for change	

Note. 5% of participants made no response to Likert item.

Unit of instruction evaluations effected change in 39% of participants. These changes were grouped into three themes: increased rigor, greater collaboration, and positive impact on daily instruction. For 56% of participants, the unit of instruction evaluation failed to create change in their instruction. Reasons were cited as it was too time consuming, it lessened teacher creativity, and teachers saw no need for change because they had received adequate scores. Five percent of participants gave no response to this item.

Research Question 1c: Professional Development Plan

Research question 1c asked, what changes have been made in teaching as a result of the professional development plan? Survey respondents were asked to rate the extent of change that has taken place in their instruction as a result of the professional development plan conducted as part of the NEE system. They could rate their amount of

change according to the following scale: 0 – no change, 1 – some change, 2 – moderate change, and 3 – significant change. The results of the Likert scale items are represented in Table 13.

Table 13

Likert Responses to Amount of Change as a Result of the Professional Development Plan

Likert Scale	Participants	Percent
0 – No Change	43	51
1 – Some Change	21	25
2 – Moderate Change	8	10
3 – Significant Change	4	4
No Response	8	10

Note. $N = 84$.

Study of the qualitative data was again conducted to look for emerging themes on participants’ responses. These represent an overall picture of what participants were sharing regarding the implementation of the professional development plan as part of the NEE system. The four themes identified in this section were sharper instructional focus, implementation of learned skills, teachers as learners, and explanations regarding no changes made.

Sharper instructional focus. This appeared to be the largest area affected by the implementation of the professional development plan. The plan provided focus to teachers regarding improvements they needed to make in their classrooms. Data

collection was identified as one way to focus instruction. Data collection provided teachers with information to identify areas in which their instruction needed improvement, or in which their students needed more instruction. Other teachers referenced specific criteria that they were working on or implementing more because it had been included in their professional development plan. These skills included critical thinking, level of taxonomy, alphabet and literacy skills, lesson planning and assessments. An additional way the professional development plan had provided instructional focus is by serving as a reminder to teachers of what areas they are working on:

...since we wrote it at the beginning of the year I have been conscious of what was put in it and make every attempt to make any changes that need to be made in order to meet the goals set in the professional development plan.

The plan helps the teachers have a goal of what they are working toward. The plans are checked by some on a regular basis to make sure the teachers are addressing what they stated they would address.

Implementation of learned skills. Teachers referred to the implementation of skills they had learned during district professional development. It was stated that any strategies learned in professional development were tried in the classroom with students. “I like to use the things we talk about in our professional development in the classroom. Any strategies we discuss or activities we participate in during those development meetings I try to play with in the class.” Teachers stated they like when the professional development addressed best practices, and they learned things that could be implemented in their instruction. Some shared specific skills that had been acquired and therefore

implemented into their instruction such as clickers, Evernote, chrome books, and other tools.

Teachers as learners. There were general, broad comments about teachers being learners themselves and their desire to constantly improve their craft. “I have always been honest with myself regarding what I need to work on or learn more about. I love learning and professional development is a part of that.” There were statements made about wanting to grow and learn as a teacher and their desire to work to better themselves as teachers. One respondent stated that their professional development had included more cross-curricular collaboration as well as building level collaboration. This collaboration helped them learn from each other as well as what they were studying.

Explanations regarding no changes made. This section appeared to have fewer negative comments regarding their lack of change in this area. However, they were similar as in other sections in that teachers feel this process is busy work and a waste of time. One stated they had not changed because they felt they were always looking for better ways to teach and that had not changed because of the implementation of the NEE system. Another respondent said the plan was written for them by their administrator and they had not even seen the new plan. It was also stated, “We meet about the PD plan but I collect data that would have been collected anyway. It does not make me a better teacher.” During the focus group one participant shared that their professional development plan had not been based on any data collected from their mini-observations.

Table 14

Summary of Themes for Research Question 1c: Professional Development Plan

Indicated Change	39%	Indicated No Change	51%
Sharper instructional focus		Explanations regarding no change:	
Implementation of learned skills		-waste of time	
Teachers as learners		-saw no need for change	
		-written by administrator	
		-not based on data collection	

Note. 10% of participants made no response to Likert item.

Professional development plans affected change for 39% of participants. These positive changes were grouped into three themes: sharper instructional focus, implementation of learned skills, and teachers as learners. The professional development plan did not affect change for 51% of focus group and survey participants. Reasons were cited as it was a waste of time, teachers saw no need for change, plans had been written by the administrator, and they were not based on data collection from observations. Ten percent of participants chose no response to this item.

Research Question 1d: Student Survey Data

Research question 1d asked, what changes have been made in teaching as a result of analyzing the student survey data? Survey respondents were asked to rate the extent of change that has taken place in their instruction as a result of analyzing student survey data conducted as part of the NEE system. They could rate their amount of change according to the following scale: 0 – no change, 1 – some change, 2 – moderate change,

and 3 – significant change. The results of the Likert scale items are represented in Table 15.

Table 15

Likert Responses to Amount of Change as a Result of Analyzing Student Survey Data

Likert Responses	Participants	Percent
0 – No Change	53	63
1 – Some Change	12	15
2 – Moderate Change	5	6
3 – Significant Change	1	1
No Response	13	15

Note. $N = 84$.

This portion of the system seemed to have the least amount of implementation. However, the smaller amount of qualitative data was again analyzed for emerging themes. The four themes identified in this section were positive comments regarding student survey information, misunderstood data-driven decision making, concerns about evaluations by children, and explanations regarding no changes made.

Positive comments regarding student survey information. Of the qualitative data studied in this area, there were a few positive comments in relationship to evaluating student survey data. One teacher stated, “It is important to get to know all students and build a community with all your different classes.” Another responded that they believe

their teaching is more effective than before. And lastly, one commented that projects were shaped based on feedback from their students.

Misunderstood data-driven decision making. This topic occurred in multiple comments showing a lack of understanding about what the NEE includes in the student survey data. Teachers misinterpreted this portion of the NEE system as collecting student assessment data. There were statements by several regarding their ability to collect more student data. They made comments about how they checked student work and assessments to drive their instruction. “Assessments drive my teaching. I constantly check my students work to determine what I need to research.” Teachers mentioned using more pre- and post-testing to gather data on students’ progress.

Concerns about evaluations completed by children. There were concerns expressed about allowing children to conduct evaluations of teachers. A teacher claimed that most high school students are not mature enough to evaluate a teacher. This teacher worried that students would not be able to separate their feelings about a teacher from the facts about the effectiveness of that teacher.

As a professional, it upsets me that 12 to 13 year old students are evaluating me. I do not agree with this aspect of NEE. How many other professions do you know in which preteens/early teenagers are evaluating the employee?

It was expressed that students’ opinions were not worthwhile since the only good evaluations would come from students who like the teacher.

Explanations regarding no changes made. Six teachers expressed that this portion of the system had not yet been implemented in their school. “This has not been done yet. My students are not surveyed.” For some, effectiveness was lost because they

had not seen any results of the survey. “I was never made aware of the results of my student surveys.” A few mentioned seeing the results but not making changes because it was not relevant or there had been no emphasis placed on it. “I have yet to use the student survey data to guide my instruction.”

Table 16

Summary of Themes for Research Question 1d: Student Survey Data

Indicated Change	22%	Indicated No Change	63%
Positive comments -builds community -teaching is more effective -projects based on student feedback		Misunderstood data-driven decisions Concerns about evaluations by children Explanations regarding no change: -no implementation -had not seen or used results -no emphasis placed on this data	

Note. 15% of participants made no response to Likert item.

Student survey data affected change in 22% percent of participants. Three positive comments were made in regard to this portion of the system. The process of collected student surveys resulted in no change for 63% of participants. Themes emerged as misunderstood data-driven decision making, concerns about evaluations completed by children, and other explanations for not making changes as a result of this process. Of focus group and survey participants, 15% chose no response to this Likert item.

General Thoughts about the Effectiveness of NEE

The survey and focus group protocol both included a final, general question regarding the respondents' general thoughts about the Network for Educator Effectiveness. This broad question prompted great response from those participating in the focus group and those completing the survey. These comments reflect teachers' perceptions of the program as a whole. Upon analyzing this data, the researcher was able to organize it into ten themes. These themes included: teaching changed for the better, good system format, value in increased administrator presence, subjective and arbitrary system, the unattainable 7, teacher stress and anxiety, administrator stress and anxiety, length of evaluations too short, frequency of evaluations too high, and excess busy work. The first of the themes identifies positive aspects of the system that promote change. The last six themes emerged as barriers to change. The following sections will outline the findings in each of these categories.

Teaching changed for the better. While a majority of comments focused on the areas of NEE that teachers felt frustration with, there were several comments that put it in a positive light. Comments were made that NEE had, in fact, changed their teaching for the better. Teachers stated they conducted more self evaluation on their students' achievement, assessments, and lessons. It was said NEE encouraged them to think more critically about their lessons and to keep focused on goals they were striving to achieve. One teacher said it helped her to be more aware to design lessons that promote more student engagement and cooperative learning. It has forced teachers to look more closely at their instructional strategies.

Good system format. Along with the changes that NEE has encouraged in classroom instruction, there were also positive comments about the format of the system.

I like the format of NEE. I appreciate it is on a continuum and the goal isn't to try and catch teachers as much as find a few areas of improvement as well as celebrate what you as a teacher are already doing.

Teachers can see the growth in their instruction by seeing the growth in their evaluation scores. Respondents stated they appreciate the immediate feedback. One shared they felt it gave administrators a way to give feedback in a helpful way, without feeling they are “attacking” teachers. This feedback helped the teacher identify and address areas they did not even realize were of concern. It was also expressed that the quick availability of scores online was a helpful piece.

Value in increased administrator presence. Teachers also expressed the value they found in having the administrator in their classroom on a more frequent basis. “I like the mini-observation walk throughs. I think the administration can learn a lot through these.” It was stated that by having the administrator in the room more often, they were able to see “real-life” scenarios. Teachers thought it was important for administrators to be in classrooms more than one time per year. “I am happy to have my administrator in my classroom on a regular basis. It allows him to see what I do every day!”

Subjective and arbitrary system. In responses from teachers, the system was often referred to as subjective and arbitrary. Teachers shared concerns that the scores were subjective in nature and could differ depending on multiple variables.

Teacher/administrator relationship was one influence teachers presented that could cause

a deviation in scores. “On several occasions teacher scores have either been inflated or deflated depending on the relationship with the administrator.” Some expressed that an administrator’s feelings about a particular teacher also effects scoring.

Another area of concern was the administrator’s lack of understanding on how to assess and score teachers. It was stated that even though administrators were “trained” it seemed they did not understand the scale and therefore did not evaluate consistently. “I think in our building it may just be a lack of understanding for how the evaluations should really be scored and used.” Along with a lack of understanding regarding the scoring, there was also concern expressed about the administrators’ lack of content knowledge. A teacher expressed that the person responsible for evaluating a teacher may not know anything about the subject matter being taught.

Buildings with multiple administrators caused teachers concern about the subjectivity due to the discrepancy in scores between two different administrators. They felt scores varied depending on who the evaluator was, even when they were doing the exact same thing during their evaluations. “We have two administrators in our building that score, and their scores are always drastically different even on similar lessons. One administrator always gives much higher scores than the other and this seems to be true throughout our district.” Another scenario was shared in which two teachers taught the exact same lesson and one earned very high scores while the other received very low scores. This discrepancy in scoring is frustrating to teachers.

The unattainable 7. The mini-observation evaluations are scored using a rubric that contains a scale from 1 to 7 depending on what the administrator actually observes during their time in the classroom. The NEE system is set up to promote growth in

teachers, so administrators are trained to begin on a score of 3 and then move up or down the continuum depending on their observation of that standard. The level of the score varies depending on the frequency of which the standard is observed and if it is being used with all students or only part of the class.

This specific piece of the evaluation process seems to cause the most unrest among teachers. It is felt, “they have been told not to give perfect scores. The inability to get a perfect score is frustrating and makes the scoring process hard to understand.” Over and over, teachers shared their dismay that they felt unable to achieve the highest score level.

This process was compared to telling students, even the brightest and most industrious, that the best they could expect is a C and that A’s are unattainable. The students would stop trying. Another example shared was telling the basketball team that there would be no way they could make it to the state championship. This would lower their moral, and they would have no incentive to work harder. Teachers feel the NEE does not allow them to aim for excellence because that is considered impossible.

This scoring leads to the premise that it is okay to be an average teacher. By not being able to score at the highest level, teachers are given the message that average is acceptable. One veteran teacher stated,

If I have taught for 30 years and have post-graduate school hours, my entire quality as a teacher is evaluated in a few, 10-minute segments a few times a year. I am told to be content with a score of three to four out of seven possible points with three being average. I am not an average teacher nor do I want my permanent file to contain evaluations that rate me as such. I have watched new

and older teachers perplexed by their numbers and I have yet to meet one person who inspired to do better because of the NEE. My motto is, 'By the grace of God I can get to Heaven but on the NEE it is impossible to score a seven!' What could be more absurd?

Being reduced to a scale of 1 – 7, with no likelihood of achieving the top score, leaves teachers feeling intimidated, demoralized, and unmotivated.

Teacher stress/anxiety. The NEE put a lot of stress on teachers. "They seem to rate themselves as not as good as they really are when they look at scores received by their principal." One teacher expressed the feeling of never being good enough, that there were always some areas that needed improvement. They did admit this caused them to constantly improve instruction, "which is positive as long as my ego can handle the slams." The subjectivity caused undue stress on the teachers as they attempted to please the observer. Teachers stated this system has created great anxiety and taken away any fun teachers had remaining in their jobs. They feel they are under a microscope and cannot truly teach to the needs of their students for fear of not making the desired grade.

Administrator stress/anxiety. A few teachers even noted the stress the NEE system caused for administrators. One commented they noticed the administrators feeling pressure to see improvement. They felt scores had been rated low early in the year in order for them to improve throughout the year despite little to no change in actual instruction. Teachers also noted it was difficult for administrators to get all of the required observations done. "I have noticed more stress among the principals. It seems to require a great deal of work on their part." It was said they were often scrambling at the end to conduct all of the required observations, taking away from their

meaningfulness. The amount of work this system placed on administrators seems to be a too great.

Length of evaluations too short. Teachers shared openly regarding their lack of change due to the length and or frequency of the mini-observations. One stated,

I believe the short walk-through evaluations can be compared to taking something out of context. Things can be misinterpreted or misunderstood and incorrect assumptions all can be made by taking things out of context and I think this is often the case with the walk-through.

Teachers shared they were often penalized if something was not observed, when often times that specific thing just did not happen during the small amount of time the administrator was in the room. Another concern addressed is that, “some observations are so short that the content observed is not enough to make a fair evaluation of the teacher’s ability to provide quality instruction.” The short amount observed is not enough for an administrator to make any sound recommendations regarding that skill or standard.

Frequency of evaluations too high. The frequency of evaluations was mentioned by several. Besides the challenge of having six to eight mini-observations per year, teachers were frustrated that veteran teachers or exemplary teachers were required to have the same number of observations as new or struggling teachers. One focus group participant shared, “I am offended that NEE thinks it is necessary to observe veteran teachers every month. I am a professional and I conduct my teaching accordingly.” Teachers perceive this to be a lack of trust in their ability to do what they are supposed to be doing. A survey respondent stated, “It seems that they are just trying to complete

them and they are not always as thorough or as beneficial in their comments.” This serves as a barrier to receiving the feedback in a helpful light.

Excess busy work. “I believe the NEE is a good tool for evaluation but it is a lot of extra work for all parties involved and to some degree repeated work.” Teachers resounded that the NEE seemed to create a lot of busy work for them, making things more complicated and time consuming. It was stated that this busy work did not necessarily lead to improved instruction. The NEE system was referred to as jumping through hoops on several responses. In some cases, teachers felt the administrator had assigned this busy work, but he did not seem to really value the product. It was felt that if the administrator did not place value on it, then it was a waste of time.

Teachers spoke of frustration with the time involved in creating a unit of instruction for evaluation. This process was viewed as a waste of time and as something more appropriate for student teachers or new teachers. The suggestion was made that teachers needed a model to guide the creation of the units of instruction so that they were not guessing at what was expected of them. The time spent on this was time they needed to “focus on students instead of keeping administrators happy”.

Table 17

Effectiveness of NEE

Promotes Change	Barriers to Change
Teaching changed for the better	Subjective and arbitrary system
Good system format	The unattainable 7
Value in increased administrator presence	Teacher stress/anxiety
	Administrator stress/anxiety
	Length of evaluations too short
	Frequency of evaluations too high
	Excess busy work

These ten themes reflect survey and focus group participants' perceptions of the NEE system as a whole. Three themes identified positive aspects that would promote change while the remaining seven identified barriers to change. All ten themes paint a picture of participants' perceptions of the effectiveness of the NEE system to create change in their instruction.

Summary

Chapter four outlined the responses of teachers surrounding the research questions which guided this study. Their perceptions of the amount of change that occurred in their instruction were presented using their Likert scale responses, followed by their qualitative responses to the open ended questions included in the survey and focus group interviews. Responses to questions regarding each of the four research questions were examined and placed into emerging themes. Results were reported in each of these themes.

The latter part of the data came from an open ended question posed to participants regarding their overall thoughts about the effectiveness of the NEE in producing change in their instruction. Seven themes emerged from this data. The first of the themes identified positive aspects of the system that promote change. The last six themes emerged as barriers of changes. Data were organized around these seven themes.

Chapter five will provide conclusions made by the researcher based on these presented findings. Implications and recommendations for practice will be presented in this portion of the paper. The limitations of the study will be identified. This chapter will also address recommendations for further study followed by a summary.

CHAPTER FIVE

CONCLUSIONS

Under the requirements of the *No Child Left Behind (NCLB) Act*, all public schools must meet adequate yearly progress (AYP) every year in raising student achievement. The end result was supposed to be 100% of students reaching proficient or advanced by the year 2014. As 2014 quickly approached, *NCLB* came under much scrutiny due to recent reports on the high number of *failing* schools. In studying AYP reports from schools across the nation for the 2010-2011 school year, it was reported only 48% of schools met AYP. This percentage of failing schools has slowly increased from 29% failing in 2006, to the most current percentage of 48% in 2012 (Usher, 2012). In 2009, President Obama and Education Secretary Arne Duncan announced a \$4.3 billion dollar education initiative called *Race to the Top*. This initiative would provide money to schools who agreed to complete certain reforms (White House, n.d.).

These initiatives prompted education leaders to look into other ways to hold schools accountable for student performance. One of the identified avenues needing improvement was that of teacher evaluation systems. While DESE had outlined seven principles of effective evaluation along with three frames for collecting data, districts would be required to create their own systems for collecting and analyzing this data. To alleviate the work districts would put into creating and maintaining a new evaluation system, the Network for Educator Effectiveness (NEE) system was created.

The NEE system was born from two auxiliary units of the College of Education at the University of Missouri: the Heart of Missouri Regional Professional Development

Center (RPDC) and the Assessment Resource Center (ARC). The NEE system is based on the Missouri educator standards and indicators. The comprehensive system is web-based, providing a place for storing and managing data on each teacher in the evaluation system. NEE includes five sources of data that can be accumulated for each teacher in the system: mini-observations, units of instruction, professional development plans, student survey data, and student assessment data.

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system as presented by the University of Missouri. The study collected their perceptions on the impact NEE had on their teaching, specifically changes made. The research questions which guided this study were as follows:

1. What are teacher's perceptions of the change impacted by the Network for Educator Effectiveness based on the framework presented by Marshall?
 - a. What changes have been made in teaching as a result of the mini-observations?
 - b. What changes have been made in teaching as a result of the unit of instruction evaluation?
 - c. What changes have been made in teaching as a result of the professional development plan?
 - d. What changes have been made in teaching as a result of analyzing the student survey data?

This chapter will provide conclusions made by the researcher and discussion based on findings presented. Implications and recommendations for practice will be presented in

this portion of the paper and the limitations of the study will be identified. This chapter will also address recommendations for further study.

Conclusions

The following conclusions were drawn by the researcher after analyzing the data gathered throughout the study. Data were collected from two southwest Missouri school districts that served as a sample of teachers from the state. The data were gathered using surveys and focus group interviews and were analyzed using interpretive analysis. Hatch (2002) stated that interpretive analysis is “about giving meaning to data” (p. 180). Upon analysis, data were organized into themes under each research questions. From this, the researcher was able to make the following conclusions.

Research Question 1a: Mini-observations

After collecting all data, it is apparent that this portion of the NEE system is the most deeply implemented across the two districts that were studied. The body of data collected in regard to this piece was greater than the other three areas of the program indicating a greater knowledge and experience level with this portion over the others. This is also supported by the data that only two participants chose no response on this item indicating they had not had any experience with this piece.

According to the Likert scale ratings, 62% of respondents made some amount of change in their instruction due to the implementation of the NEE system, 36% indicated no change, and 2% chose to not respond. While the levels of change varied, it should be noted that an amount of change occurred due to the implementation of this piece of the system. This data indicate a success of mini-observation evaluations as a tool for teacher evaluation.

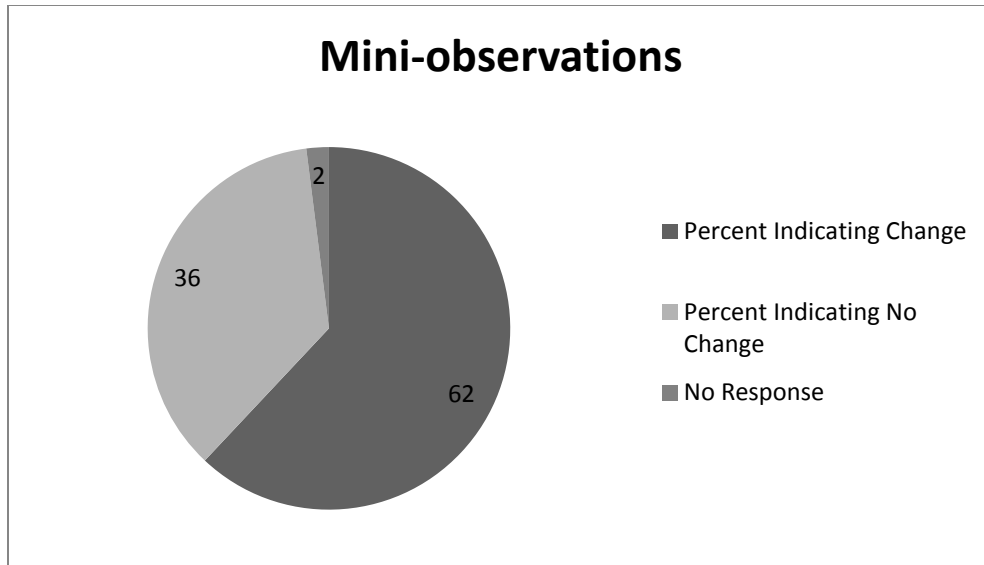


Figure 1. Percent of change indicated based on implementation of mini-observation evaluations.

During the focus groups and in the open-ended responses to the survey questions, teachers indicated several components to their classroom instruction that had been altered based on the implementation of the mini-observations. They were aware of the criteria being evaluated and were more conscious to include those pieces into their lessons. They also noted that following these mini-observations, changes were often made due to the feedback that was received from their administrator.

It was also apparent, however, that teachers had great concern about several aspects of this system. Some of their arguments were legitimate concerns that hampered the effectiveness of the program and served as a barrier to change. Some concerns however, were subjective and not shared by all participants. Most concerns surrounded the subjectivity of the scoring along with the frequency of evaluation.

Research question 1b: Unit of Instruction Evaluation

The unit of instruction evaluation appeared to have second highest level of implementation. This was determined again by the body of data collected and the number of participants who chose no response indicating their lack of experience with

this piece of the system showing only five. According to the Likert scale responses, 39% of participants made some amount of change due to the implementation of the unit of instruction evaluation, 56% indicated no change, and 5% chose no response to the Likert item. These numbers indicated a deficiency in this portion's ability to create change.

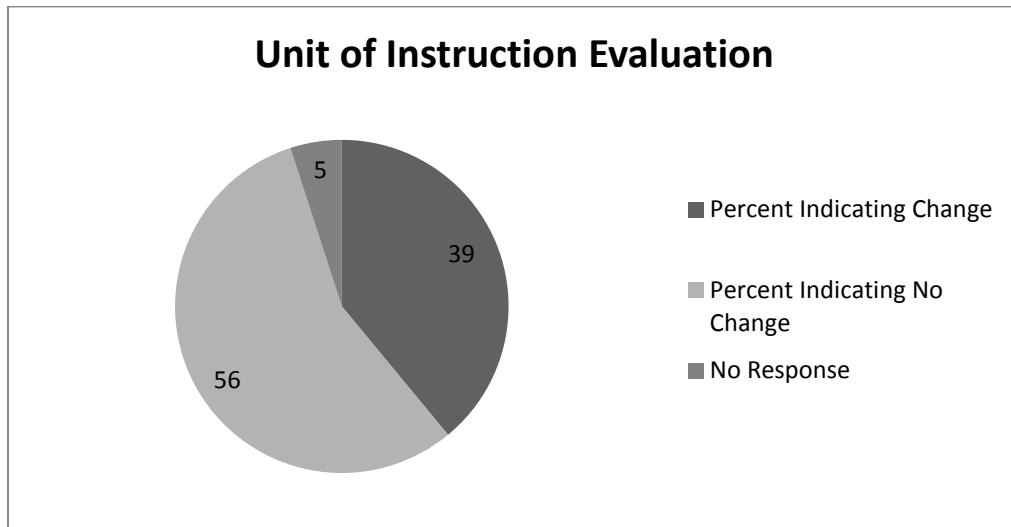


Figure 2. Percent of change indicated based on implementation of unit of instruction evaluation.

Upon analyzing the data, it can be concluded that this may be due to the improper implementation of this piece. It was indicated by several teachers that this unit of instruction had been completed by teams rather than individually. This had been done because upon initial implementation, districts felt it would help teachers to work together. While they did appreciate this, it often led to teachers working on units that were not in their realm of teaching. Another concern regarding this was that some teachers did most of the work while others did little. Several did indicate this collaboration as a positive aspect. They felt the learning that came from working as a team or with another teacher had a positive impact on their instruction. However, a majority viewed this process as an exercise to fulfill a requirement rather than a tool to improve instruction.

The researcher is led to believe that upon proper, or full, implementation of this piece there is potential to create change. This assumption is based on the data collected from those who indicated that it had influenced their instruction. They shared that the rigor of their lessons and assessments had increased. Levels of questioning, teaching, and assessment had been raised due to the development of the unit of instruction. They also felt it narrowed their focus on what they were supposed to be teaching and kept them from getting off track. Teachers were more conscious to cover all standards that had been identified as part of their unit.

Research Question 1c: Professional Development Plan

This portion of the NEE system would rank third in the level of implementation. There were minimal responses to the open-ended questions and not much discussion was offered during the focus group interviews. Eight participants indicated the lack of implementation by not responding to this item. The Likert scale responses showed 39% of respondents indicated any amount of change occurring in their instruction due to the implementation of the professional development plan, 51% indicated no change, and 10% chose no response. These numbers indicated a deficiency in this portion's ability to create change.

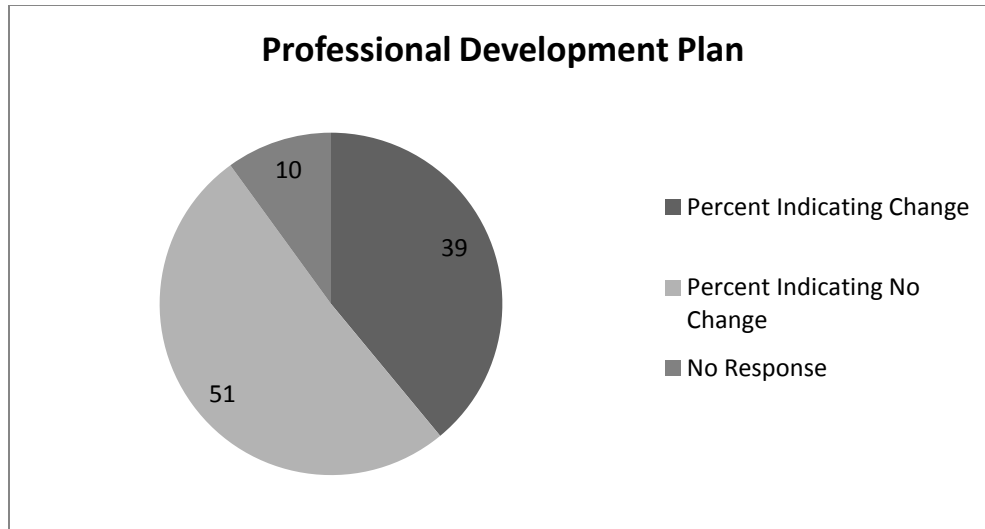


Figure 3. Percent of change indicated based on implementation of professional development plans.

Based on the knowledge of the implementation and the responses by participants, it is again determined this lack of effectiveness may be due to the improper implementation. Several responses indicated that these plans had been completed, but no survey data or focus group responses indicated that their professional development plan had been based on any data collected from their mini-observations, unit of instruction evaluation, or student survey data. These plans should be written on an individual basis with standards selected based on data collected from the other pieces of the system. This ensures individuals are working on areas in which they may be lacking skills or knowledge. This individualization would likely increase the possibility of creating impact on instruction.

There is great potential in this piece as several participants indicated they viewed themselves as constant learners. They shared they liked to learn and implement new things. It was the desire of most participants to improve themselves as teachers and to perfect their craft. The further implementation and development of the professional development plans will cause an increase in the amount of change it has on instruction.

Research Question 1d: Student Survey Data

This portion of the system revealed the least amount of data, showing the least amount of implementation. Thirteen participants chose no response to this item and there was little discussion in relationship to it. This may be due to the fact that these student surveys are not given to students below the fourth grade. Given that 43% of survey responses and 52% of the focus group participants were elementary teachers, this may have impacted the results. The Likert responses pointed out only 22% of respondents indicated any amount of change in their instruction due the implementation of student surveys, 63% indicated no change, 15% chose no response. At this level of implementation, this is determined to have little effect on teachers' instructional practices.

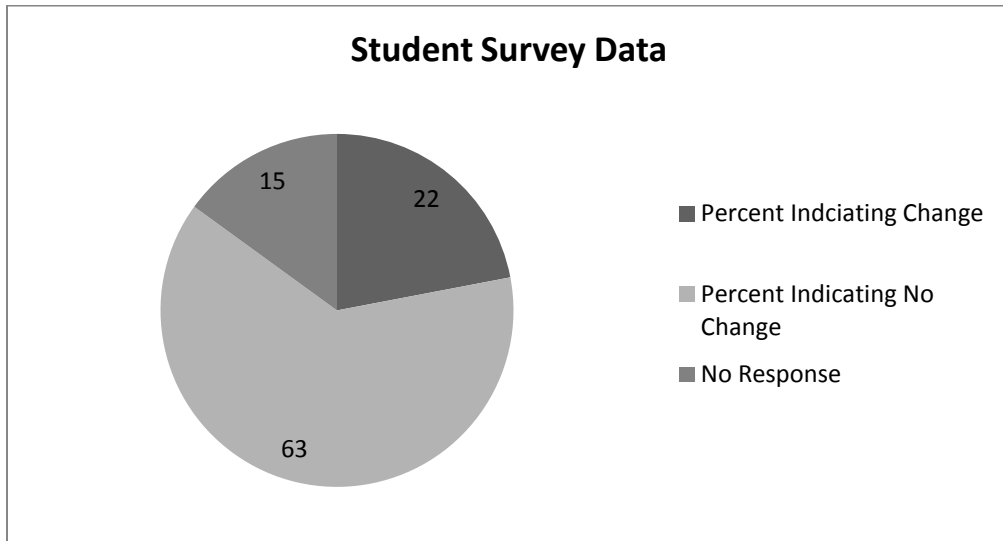


Figure 4. Percent of change indicated based on implementation of student survey data.

A majority stated this piece had not been done yet or had been planned to take place later in the school year but had yet to occur. Some indicated students had taken the survey, but the results were never shared or there was no emphasis placed on the results. Aside from the low-level of implementation, there is also another factor blocking the

effectiveness of this piece. Teachers have the mindset that students are not equipped with the skills to evaluate teachers. They view it as an unfair assessment and unreliable data.

Table 18 outlines a summary of these conclusions. Positive and negative aspects were identified surrounding each of the research questions as presented in this portion of the paper. These findings are summarized into these categories in the following table. These conclusions were based on the data collected through this study.

Table 18

Summary of Conclusions

Evaluation Piece	Positive	Negative
Mini-observations	-Raised awareness of standards -Good feedback	-Subjective -Frequency/Length
Unit of Instruction	-Increased rigor -Narrowed focus -Raised coverage of standards	-Waste of time -Not relevant
Professional Development Plan	-Teachers are constant learners -Teachers desire improvement	-Not based on data -Not individualized
Student Survey Data	-Student feedback shaped Instruction	-Low implementation -Results not shared with teachers -Viewed as unreliable

Discussion

This program evaluation serves to assess the effectiveness of the NEE system. Therefore, it is not comparable to any previous studies. However, as presented in chapter two, much research has been presented on each of the four aspects of NEE that are evaluated in this study. The following section of this paper will connect the findings of this study to the research already existing on each of the four pieces of NEE.

Data showed the mini-observation piece to be the most effective of the four. In a previous study it was stated, “Instructional leadership means very little unless leaders are willing and able to observe teachers, offer advice about problems, and make formative evaluations that support and pinpoint areas to improve” (Khan et al., 2009, p. 585). The data from this program evaluation supported this. Teachers noted that following the mini-observations, changes were often made due to the feedback that was received from their administrator. By providing the instructional leadership through observations, administrator feedback created an avenue for change and focus on specific skills and standards.

The unit of instruction evaluation was found to be less effective, but still created some amount of change. Data pointed to a lack of individualization in this area, teachers were required to write units in a collaborative way. This often took away some of the meaning because the work was not relevant to a teacher’s specific curriculum. Beyer and Davis stated that these resources help teachers define what students need to know and be able to do as a result of instruction (2009). If this unit of instruction planning is not done in a manner that requires teachers to define specifically what their students need to know and do, then it will not reach its full potential.

The third piece, also the third in terms of effectiveness, was the professional development plan. Again, the largest downfall to the implementation of this portion was the lack of individuality. Koops and Winsor (2006) shared their conviction that every teacher deserves to have a professional growth plan. Many support the idea that this plan should be individual in nature, catering to a teacher's specific needs (Borko, 2004; Fenwick, 2004; Khan et al., 2009; Koops & Winsor, 2006; McQuarrie & Wood, 1991; Payne & Wolfson, 2000). The data in this study supports this research. Teachers shared their frustration in creating professional development plans that were not individual in nature or that were not created by themselves, but had them done for them by an administrator.

Last, and least effective in creating change, was the collection and analysis of student survey data. There were only a few teachers in this study that found any merit in accepting feedback from their students. However, Ripley (2012) stated research has shown if kids are asked the right questions, they could identify with accuracy their most, and least, effective teachers. It turned out their answers were more reliable than any other measure of teacher performance, including observations and student test-score growth. Evaluation of student survey data found the most discrepancy in data from this study compared to previous research.

Implications and Recommendations for Practice

The following section of this paper will outline implications and recommendations for practice. These are based on the study of related literature in conjunction with the data collected from this study. Based on these pieces of

information, it is believed these recommendations would make the NEE system have an even greater effect on classroom instruction.

Research Question 1a: Mini-observations

As these mini-observations have shown effectiveness in creating change, it is recommended they continue in regularity and length. Koops and Winsor (2006) explained that multiple short visits allow an administrator to observe a variety of lesson types. This creates a greater opportunity for follow up conversations with teachers regarding what the students are doing in the classroom along with the teacher's goals. This system provides the multiple visit criteria to accomplish this. Administrators could enhance this practice by possibly staying in the classroom a little longer to be able to observe more parts of a lesson, but the researcher understands that time restraints may inhibit this due to the number of classroom visits required.

One recommendation for improvement in this area would be to keep the same number of visits but make some of them less formal than others. Teachers expressed that they liked the consistency of what was observed, so that would need to stay consistent. However, half of the visits could be conducted and communicated to the teacher as a form of formative assessment but not be officially scored in the system. The remaining observations would receive scoring and be documented.

In addition to this, teachers expressed a desire to see their scores before they were officially submitted. The system could include a way to "hold" scores until the feedback portion had taken place and could then be submitted by the administrator after the follow-up conversation had taken place. This would also increase the frequency by which the follow-up conversations would occur. This feedback is a vital part of the program that is

often times skipped due to restraints of time. Feeney (2007) stated, “The feedback provided to teachers should be descriptive and based on what the teacher and students actually do, supported by evidence of student outcomes throughout the lesson” (p. 195). Feeney explained these authentic examples allow the teacher to replay events and evaluate the lessons effectiveness. Providing this feedback must occur to make the mini-observation piece most effective.

Research Question 1b: Unit of Instruction Evaluation

In order to increase the effectiveness of the unit of instruction evaluation piece, full implementation has to occur. Teachers need to be completing these in areas that are relevant to their job. These resources help teachers define what students need to know and be able to do as a result of instruction (Beyer & Davis, 2009). These units would provide the most benefit if they were related to the subject matter or level of their particular students.

Another recommendation is that the units be completed in a format that is useful and in line with other district expectations. Teachers shared that they often had to transfer from one format to another simply to satisfy the system. Districts should be allowed to accept the format they use as their district curriculum to satisfy this requirement. This would alleviate teachers’ frustration with repeated work and the task would become more meaningful.

In addition to these suggestions, education regarding what the units should contain and the expectations held by administrators should be conducted. The study participants expressed confusion and dismay at not knowing what was expected and spent too much time trying to understand the requirements rather than focusing on the content.

By providing more guidance, the focus would shift back to the content rather than the structure of the document. Tomlinson (2001) stated curriculum will include essential facts and vocabulary to be mastered in a unit of study. It will cover the basic concepts of the discipline and then show students how they will master those concepts. By making the shift toward content, essential facts and vocabulary, this piece would improve in effectiveness.

Research Question 1c: Professional Development Plan

To be effective in creating change, the professional development plans have to be written individually by teachers based on needs identified by data. Wanzare (2002) believed teacher appraisal should provide directions for teacher development, provide the opportunity for teachers to develop new skills, and the outcome of the appraisal should inform further teacher development. By utilizing the data provided from the mini-observations, administrators and teachers already have the tools needed to provide direction for teacher development. This information is simply not being used at this point. By utilizing this already existing data, the effectiveness of the professional development plan would improve exponentially.

Administrators should require teachers to construct their professional development plan based on areas of need identified through their mini-observations, unit of instruction, or student surveys. Data collected in this study reveal teachers' willingness to learn and become better teachers. If this piece is developed to its fullest extent, this would be a very powerful tool to improve instruction that teachers would embrace.

Research Question 1d: Student Survey Data

Education is the key to making this piece more effective. As it stands, teachers undervalue the power feedback from their students could hold. Administrators need to take time to present research, share examples, and promote productive comments that students can offer regarding a teachers' instructional practices. Stiggins (1986) stated there may be no more valid source of information about the learning environment than from the students who work in those environments. He believed students can provide insights no one else can offer. When teachers learn to filter out some of the data from students that is not constructive in nature, they would reveal the usefulness of most students' comments. Administrators should commit to this process by providing research and spending time with teachers sorting through data collected from students to help them identify what is useful and what is not.

Overcoming Barriers

The data collection uncovered multiple barriers to change that are inhibiting the NEE system from reaching its fullest potential for effectiveness. The following sections will address implications and recommendations for practice to reduce these counter-productive barriers. Recommendations made will be in relationship to the tool itself, as well as, toward the administrators who implement the system.

Subjectivity. Subjectivity is going to be a concern as long as humans are conducting the evaluations. This can be minimized by implementing or improving upon practices in administrator training. Administrators are required to attend training on a yearly basis in order to participate in the NEE evaluation system. They are required to

complete evaluations of videos and score consistently with other administrators in order to become *calibrated*.

The purpose of this *calibration* is to reduce subjectivity. This process needs to be conducted on a more frequent basis. The other area that is lacking is the training provided has not covered many of the standards that are included in the system. There is the possibility that administrators are evaluating standards on which they have not received training. The NEE system administrators need to further develop their training modules to include more standards.

Range et al. (2012) stated school leaders need to be trained in how to provide feedback that investigates instructional change and how to implement improvement plans for struggling teachers. This is another area the NEE system could add to ensure administrators are consistent in how feedback is provided to teachers. Training needs to include not only when to give feedback, but how to do it so it has the greatest impact. This could improve some of the subjectivity in feedback received by different administrators.

The unattainable 7. The idea that scoring a seven on the evaluation tool was unattainable was found throughout the data. This is merely due to a lack of communication. Survey participants stated that, “no person is supposed to be able to score at the highest level”, and “we were told that NO ONE will earn a 7.” Koops and Winsor (2006) believed observations are most valuable when they can be used as a formative tool to collect data to aid in teacher growth and improvement. This supports the idea of a multipoint scale with room for growth. It was communicated to administrators and teachers through NEE training that scoring somewhere in the middle

was a good starting point. It was also expressed that most teachers would not get sevens during every evaluation. This does not mean NO ONE can ever get a score of seven.

Administrators need to do a better job communicating to teachers that the intent is not to keep teachers from reaching their fullest potential. The intent is to provide information to them regarding what is seen during their observation period. If the standard is being met, then a seven can be achieved. However, to expect to receive a seven on multiple observations is not realistic. It is also important to communicate the expectations so teachers know what to do to reach the high score of seven.

Stress and anxiety. This area of need should be addressed by building administrators implementing the program. Evaluating how requirements are fulfilled and eliminating any duplicate or unnecessary steps would be a benefit to teachers. Removing obstacles such as confusion and unneeded paperwork would lessen frustration and make teachers more open to improvement. Expectations and requirements should be made easily understood so teachers are not guessing at what to submit or what is being evaluated. This would also eliminate the barrier of teachers feeling there is a lot of busy work that is not truly valued by administrators.

While the program is about promoting growth in teachers, it also needs to provide them with praise for work that is done well. Hattie (2012) explained that good feedback can be motivational so more effort and skill are invested in a task. Everyone has areas that are in need of improvement, but they also have areas of strength. By taking time to acknowledge those strengths, teachers would feel more supported and less under fire, thus removing some of the stress and anxiety.

Limitations

The study was limited to two small school districts in order to narrow the focus for the purpose of this particular study. However, the small number of schools participating limited the findings. Generalizations about teachers across the state were made based on the data collected from only two schools. Collecting data through two forms, survey and focus groups helped to gather a wealth of data points from which the researcher could make conclusions about the NEE system.

Rossi et al. (2004) stated that at various times policy makers and program managers need to distinguish worthwhile programs from ineffective ones and launch new ones or revise existing ones to achieve desirable results. This validates a need to evaluate new programs. However, the recent implementation of the NEE system limited the findings to a degree. There were pieces of the system that appeared to have no influence on classroom instruction because they had yet to be fully or correctly implemented. Some of the data collected may have been skewed by participants' lack of experience or frustration with the initial implementation of this new program.

The lack of participation in the focus group interviews served as a slight limitation as well. While there were a sufficient amount of participants to collect relevant data, the researcher hoped to gain more insight into teachers' perceptions of change through these face-to-face interviews. The good response rate of the surveys helped to overcome the lower participation in this area.

Recommendations for Further Study

The purpose of this study was to examine teachers' perceptions of change impacted by the Network for Educator Effectiveness teacher evaluation system. It served

as a program evaluation of this new system. School districts across the state are implementing new teacher evaluation systems; this evaluation provides them with research on the NEE system's effectiveness.

For future study, it is recommended that this program be evaluated after it has shown deeper implementation. To study this system again after multiple years of implementation would reveal its true effectiveness. Participants would be better able to address all pieces of the system if they had more experience with them. It is also more likely that all pieces would have proper implementation.

The focus of this study was to evaluate whether or not change took place in teachers instruction as a result of the implementation of the NEE system. Participants were asked if they had made changes and if so, explain those changes. A future study would garner more rounded feedback if it placed more focus on the positives, as well as, the negatives in relationship to change. The data in this research often showed no change in teachers' instruction but there were often no explanation as to why. Future research should inquire as to why change was made or why it was not. This would provide more data from which to draw conclusions about improvement of the program.

Another probe that would enhance a future study would be to ask teachers what they would suggest to improve the program. Some provided this information in their open ended response. However, most did not. A targeted question asking for suggestions or ideas to enhance this evaluation system would provide a large portion of ideas that could then be compared to prior research to identify strategies that could be put into place for program improvement.

A final way to further study the NEE system would be to broaden the research to include more than two school districts. Including districts across the state, or samples from different parts of the state, would garner a more complete picture of what teachers' perceptions are throughout the state. By broadening across the state, the research data would include a more diverse population from which to collect data. The data would also cover more variance in implementation levels and experiences.

Summary

This qualitative case study was conducted to measure the perceptions of teachers in two small southwest Missouri school districts regarding the NEE teacher evaluation system. Data were collected in regard to changes that have occurred in their teaching following the implementation of the program. The utilization focused program evaluation will guide the creators of the new evaluation system in their further development of the system. This would also inform administrators on recommendations for improving implementation.

Participants were selected from two school districts that are involved in the NEE system for the second year. A maximum variation sample was selected to represent all types of teachers from all levels and subject. Focus groups and surveys collected the data needed for the researcher to make conclusions about the effectiveness of this new program. Data was analyzed to identify the effectiveness of the NEE system. Effectiveness was identified in one of four areas at the current level of implementation. Recommendations were presented to aide in continued development and implementation of the system to further the success of the program.

References

- Beyer, C. J., & Davis, E. A. (2009). Using educative curriculum materials to support preservice elementary teachers' curricular planning: A comparison between two different forms of support. *Curriculum Inquiry, 39*(5), 679-703.
doi:10.1111/j.1467-873X.2009.00464.x
- Bolman, L. G., & Deal, T. E. (2008). *Reframing organizations: Artistry, choice, and leadership* (4th ed.). San Francisco, CA: Jossey-Bass.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15. doi:10.3102/0013189X033008003
- Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the 21st century*. New York, NY: Carnegie Corporation.
- Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.
- Cogshall, J., Rasmussen, C., Colton, A., Milton, J., & Jacques, C. (2012). *Generating teacher effectiveness: The role of job-embedded professional learning in teacher evaluation*. Retrieved from
www.gtlcenter.org/sites/default/files/docs/GeneratingTeacherEffectiveness.pdf
- Cooper, B. S., Ehrensals, P. A., & Bromme, M. (2005). School-level politics and professional development: Traps in evaluating the quality of practicing teachers. *Educational Policy, 19*(1), 112-125. doi:10.1177/0895904804272231
- Corcoran, S. P. (2007). Long-run trends in the quality of teachers: Evidence and implications for policy. *Education Finance and Policy, 2*(4), 395-407.

- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Danielson, C. (2010). *Danielson's framework for teaching for classroom observations*. Retrieved from <http://metproject.org>
- Danielson, C. (2011). *The framework for teaching*. Retrieved from <http://www.danielsongroup.org>
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional learning*. Princeton, NJ: Educational Testing Service
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), 8-15.
- DuFour, R., & Marzano, R. (2009). High-leverage strategies for principal leadership. *Educational Leadership*, 66(5) 62-68.
- Eastridge, H. E. (1976). Student evaluation and teacher performance. *NASSP Bulletin*, 60(401), 48-54. doi:10.1177/019263657606040110
- Ebmeier, H. (2003). How supervision influences teacher efficacy and commitment: An investigation of a path model. *Journal of Curriculum and Supervision*, 18(2), 110-141.
- Feeney, E. J. (2007). Quality feedback: The essential ingredient for teacher success. *Clearing House*, 80(4), 191-198.
- Fenwick, T. J. (2004). Teacher learning and professional growth plans: Implementation of a provincial policy. *Journal of Curriculum and Supervision*, 19(3), 259-282.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.

- Fink, A. (2009). *How to conduct surveys: A step-by-step guide* (4th ed.). Thousand Oaks, CA: Sage.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Retrieved from <http://www.gtlcenter.org/sites/default/files/docs/LinkBetweenTQandStudentOutcomes.pdf>
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Retrieved from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>
- Hatch, J. A. (2002). *Doing qualitative research in education settings*. Albany, NY: State University of New York Press.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York, NY: Routledge.
- Heck, R. H. (2009). Teacher effectiveness and student achievement: Investigating a multilevel cross-classified model. *Journal of Educational Administration*, 4(2), 227-249.
- Hughes, V. M. (2006). *Teacher evaluation practices and teacher job satisfaction* (Doctoral dissertation). Retrieved from <https://mospace.umsystem.edu>
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: ASCD.
- Kalule, L., & Bouchamma, Y. (2013). Supervisors' perceptions of instructional supervision. *International Studies in Educational Administration*, 41(1), 89-104.
- Kennedy, M. M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*, 39(8), 591-598. doi:10.3102/0013189X10390804

- Khan, Z., Khan, U. A., Shah, R. U., & Iqball, J. (2009). Instructional leadership, supervision and teacher development. *The Dialogue*, 4(4), 580-592.
- Konold, T., Jablonski, B., Nottingham, A., Kessler, L., Byrd, S., Imig, S., & McNergney, R. (2008). Adding value to public schools: Investigating teacher education, teaching, and pupil learning. *Journal of Teacher Education*, 59(4), 300-312.
doi:10.1177/0022487108321378
- Koops, J. B., & Winsor, K. A. (2006). Creating a professional learning culture through faculty evaluation. *The Journal of Education*, 186(3), 61-70.
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks, CA: SAGE Publications.
- Kunzman, R. (2003). From teacher to student: The value of teacher education for experienced teachers. *Journal of Teacher Education*, 54(3), 241-253.
doi:10.1177/0022487103054003006
- Larsen, M. A. (2009). Stressful, hectic, daunting: A critical policy study of the Ontario teacher performance appraisal system. *Ontario Journal of Educational Administration and Policy*, (95), 1-44.
- Marquardt, M. J. (2011). *Building the learning organization: Achieving strategic advantage through a commitment to learning* (3rd ed.). Boston, MA: Nicholas Brealey.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation*. San Francisco, CA: Jossey-Bass.
- Marzano, R. J. (2012). Art & science of teaching: Reducing error in teacher observation scores. *Educational Leadership*, 70(3), 82-83.

- Marzano, R. J. (2013). Art & science of teaching: How to show student learning. *Educational Leadership*, 71(2), 82-83.
- Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD.
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: ASCD.
- Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD.
- Mathis, W. (2012). *Research-based options for education policy making: Teacher evaluation*. Retrieved from <http://nepc.colorado.edu/publication/options-teacher-evaluation>
- McIntyre, D. J. (1980). Teacher evaluation and the observer effect. *NASSP Bulletin*, 64, 36-40. doi:10.1177/019263658006443408
- McQuarrie, Jr, F. O., & Wood, F. H. (1991). Supervision, staff development, and evaluation connections. *Theory Into Practice*, 30(2), 91-96.
- Merriam-Webster. (2014). *Dictionary*. Retrieved from <http://www.merriam-webster.com/dictionary>
- Missouri Department of Elementary and Secondary Education. (2013). *Overview of essential principles of effective evaluation*. Retrieved from <http://www.dese.mo.gov/eq/documents/EssentialPrinciplesOverview-July2013.pdf>

- Missouri Department of Elementary and Secondary Education. (2012A) *Professional frames of the educator*. Retrieved from <http://www.dese.mo.gov/eq/documents/eq-ees-professional-frames.pdf>
- Missouri Department of Elementary and Secondary Education. (2012B) *State of Missouri ESEA flexibility request*. Retrieved from <http://dese.mo.gov/qs/documents/qs-esea-waiver-mo-flexibility-request.pdf>
- Missouri Department of Elementary and Secondary Education. (2012C) *Statutory authority*. Retrieved from <http://dese.mo.gov/eq/documents/eq-ees-statutory-authority.pdf>
- Moxley, Jr., R. A. (1978). Teacher evaluation: Images and analysis. *Journal of Teacher Education*, 29(6), 61-66. doi:10.1177/002248717802900619
- Murphy, J., Hallinger, P., & Heck, R. (2013). Leading via teacher evaluation: The case of the missing clothes? *Educational Researcher*, 42(6), 349-354. <http://dx.doi.org/10.3102/0013189X13499625>
- National Board for Professional Teaching Standards. (2014). *The beginnings of a movement*. Retrieved from <http://www.nbpts.org/beginnings-movement>
- Normore, A. H. (2005). Integrating personnel evaluation in the planning and evaluation of school improvement initiatives. *American Journal of Evaluation*, 26(3), 348-351. doi:10.1177/1098214005278757
- Nye, B., Hedges, L., & Konstantopoulos, S. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257. doi:10.3102/01623737026003237

- Ong'ondo, C. O., & Borg, S. (2011). 'We teach plastic lessons to please them': The influence of supervision on the practice of English language student teachers in Kenya. *Language Training Center, 15*(4), 509-528.
doi:10.1177/1362168811412881
- Owings, W. A., Kaplan, L. S., Nunnery, J., Marzano, R., Myran, S., & Blackburn, D. (2006). Teacher quality and troops to teachers: A national study with implications for principals. *NASSP Bulletin, 90*, 102-131.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research, 34*(5), 1189-1208.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Payne, D., & Wolfson, T. (2000). Teacher professional development; The principal's critical role. *NASSP Bulletin, 84*(618), 13-21. doi:10.1177/019263650008461803
- Range, B. G., Duncan, H. E., Scherz, S. D., & Haines, C. A. (2012). School leaders' perceptions about incompetent teachers: Implications for supervision and evaluation. *NASSP Bulletin, 96*(4), 302-321. doi:10.1177/0192636512459554
- Rice, J. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Ripley, A. (2012, September 19). Why kids should grade teachers. *The Atlantic*. Retrieved from www.theatlantic.com/magazine/archive/2012/10/why-kids-should-grade-teachers/309088/#
- Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 471-458.

- Roberson, S., & Roberson, R. (2009). The role and practice of the principal in developing novice first-year teachers. *Clearing House*, 82(3), 113-118.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1), 43-74. doi:10.1162/EDFP_a_00022
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.
- Roulston, K. (2010). Considering quality in qualitative interviewing. *Qualitative Research*, 10(2), 198-228. doi:10.1177/1468794109356739
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Schmoker, M. (1999). *The key to continuous school improvement* (2nd ed.). Alexandria, VA: ASCD.
- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (3rd ed.). New York, NY: Teachers College Press.
- Smith, T. M., & Rowley, K. J. (2005). Enhancing commitment or tightening control: The function of teacher professional development in an era of accountability. *Educational Policy*, 19(1), 126-154. doi:10.1177/0895904804270773
- Stiggins, R. J. (1986). Teacher evaluation: Accountability and growth systems--different purposes. *NASSP Bulletin*, 70(51), 51-58. doi:10.1177/019263658607049015

- Taylor, F. W. (1916). The principals of scientific management. In J. M. Shafritz, J. S. Ott, and Y. S. Yang, (Eds.), *Classics of organization theory* (7th ed., pp. 65-76). Boston, MA: Wadsworth.
- Tomlinson, C. A. (2001). Standards and the art of teaching: Crafting high-quality classroom. *NASSP Bulletin*, 85(622), 38-47. doi:10.1177/019263650108562206
- Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- University of Missouri. (2013). *Network for educator effectiveness*. Retrieved from nee.missouri.edu
- U.S. Department of Education. (1983). *A nation at risk*. Retrieved from <http://www2.ed.gov/pubs/NatAtRisk/risk.html>
- Usher, A. (2012). *AYP results for 2010-2011--November 2012 update*. Retrieved from <http://www.cep-dc.org/displayDocument.cfm?DocumentID=414>
- Wanzare, Z. O. (2002). Rethinking teacher evaluation in the third world: The case of Kenya. *Educational Management Administration & Leadership*, 30(2), 213-229. doi:10.1177/02611X02030002511
- Waring, H. Z. (2013). Two mentor practices that generate teacher reflection without explicit solicitations: Some preliminary considerations. *RELC Journal*, 44(1), 103-119. doi:10.1177/0033688212473296
- White, M. E., Makkonen, R., Vince, S., & Bailey, J. (2012). *How California's local education agencies evaluate teachers and principals*. (REL 2012-023). Washington, D.C.: U.S. Department of Education.

White House. (n.d.) *Race to the Top*. Retrieved from

<http://www.whitehouse.gov/issues/education/k-12/race-to-the-top>

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher & classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*, 57-67.

APPENDIX A

Focus Groups Protocol

Opening: Tell me your name, how long you have been teaching, and what you teach.

Transition: Share your thoughts on the purpose of teacher evaluations.

Key Questions:

1. How many mini-observation walk-throughs did your administrator conduct in your classroom during the 2012-2013 school year?
2. How quickly did they give you face-to-face feedback on that walk-through?
3. Was that feedback helpful? Why or why not?
4. What changes did you make in your teaching, if any, as a result of those walk-throughs and any feedback received?
5. How has your district implemented the evaluation of a unit of instruction?
6. What changes did you make in your teaching, if any, as a result of that evaluation?
7. How has your district implemented the evaluation of a professional development plan?
8. What changes did you make in your teaching, if any, as a result of that evaluation?
9. How did your district implement the collection of student survey data?
10. Were you evaluated by students using this tool?

If not, why?

If so, explain any changes you made in your teaching as a result of this evaluation.

Closing: What are your overall thoughts on the Network for Educator Effectiveness?

APPENDIX B

Informed Consent – Focus Group

Identification of Researchers: This research is being conducted by Jennifer Katzin, a doctoral student with the Department of Education Leadership & Policy Analysis, College of Education, at the University of Missouri.

Purpose of the Study: The purpose of this qualitative study is to understand the perspectives of pre-kindergarten through twelfth grade teachers working in two small school districts located in southwest Missouri who are enrolled in the NEE system.

Request for Participation: We are inviting you to participate in a focus group discussion forum which will provide feedback on the changes you have or have not made in your teaching as a result of the implementation of the Network for Educator Effectiveness. If you decide not to participate, you will not be penalized in any way. You can also decide to stop at any time without penalty. If you do not wish to answer any of the questions, you may simply ask to skip them.

Exclusions: You must be an employee of the participating school districts. You must have participated in the Network for Educator Effectiveness.

Description of Research Method: The research involved will utilize a focus group discussion interview and sharing of anecdotal stories. The focus group interview will take approximately one hour to complete. The discussion will be recorded for accuracy.

Privacy: All the information we collect will be confidential. I will create a code for your name or any information which could be used to identify you. Neither the recordings nor the focus group notes will be shared with any organization internally or externally to the University of Missouri program. Once the recording has been transcribed, it will be deleted.

Explanation of Risks: The risks to this study are similar to the risks of everyday life.

Explanation of Benefits: The results from this research project will benefit the study of effects of the Network for Educator Effectiveness.

Questions About Your Rights: If you have any questions about your rights as a research participant, please contact my chairperson Dr. Robert Watson, Missouri State University, by email at RobertWatson@MissouriState.edu.

Your participation signifies your informed consent.

APPENDIX C

SURVEY

What district do you work in?

How long have you taught?

- 0 – 5 years
- 6 - 15 years
- 16+ years

What level do you teach?
grade)

- Elementary (Pre-K – 5th grade)
- Middle School/Junior High (6th grade – 8th grade)
- High School (9th grade – 12th grade)

What is your gender?

- Male
- Female

For the following questions, a three point scale will be used:

- 0 - No Change
- 1 - Some Change
- 2 - Moderate Change
- 3 - Significant Change

If your district has not implemented a portion of the evaluation system, you may skip that question.

1. To what extent have you made changes in your instruction as a result of the mini-observation walk-throughs conducted as part of the Network for Educator Effectiveness?

- | | | | |
|-----------|-------------|-----------------|--------------------|
| 0 | 1 | 2 | 3 |
| No change | Some Change | Moderate Change | Significant Change |

If any amount of change was indicated in the above scale, please describe those changes made. If no change was indicated please explain.

2. To what extent have you made changes in your instruction as a result of the evaluation of a unit of instruction as part of the Network for Educator Effectiveness?

0	1	2	3
No change	Some Change	Moderate Change	Significant Change

If any amount of change was indicated in the above scale, please describe those changes made. If no change was indicated please explain.

3. To what extent have you made changes in your instruction as a result of the evaluation of a professional development plan as part of the Network for Educator Effectiveness?

0	1	2	3
No change	Some Change	Moderate Change	Significant Change

If any amount of change was indicated in the above scale, please describe those changes made. If no change was indicated please explain.

4. To what extent have you made changes in your instruction as a result of the collection of student survey data as part of the Network for Educator Effectiveness?

0	1	2	3
No change	Some Change	Moderate Change	Significant Change

If any amount of change was indicated in the above scale, please describe those changes made. If no change was indicated please explain.

Please share your overall thoughts of the Network for Educator Effectiveness:

THANK YOU!

APPENDIX D

Informed Consent Form - Survey

Purpose of the Study:

The purpose of this qualitative study is to understand the perspectives of pre-kindergarten through twelfth grade teachers working two small school districts in southwest Missouri, who are enrolled in the NEE system.

What will be done:

You will complete a survey, which will take 20 minutes to complete. The survey includes questions about your involvement in the Network for Educator Effectiveness. Survey questions will address your perceptions of the change that has occurred in your teaching as a result of the implementation of NEE.

Benefits of this Study:

You will be contributing to knowledge of the effectiveness of the Network for Educator Effectiveness.

Risks or discomforts:

No risks or discomforts are anticipated from taking part in this study. If you feel uncomfortable with a question, you can skip that question or withdraw from the study altogether. You are free to quit at any time before you have finished the survey.

Confidentiality:

Your responses will be kept completely confidential. I will NOT know your name from the survey. No part of the survey will be shared with any organization internally or externally to the University of Missouri program. Only the researcher will see your individual survey responses and the results of our content analysis from the survey. At the end of the survey, I ask your permission to use quotations from your survey responses. If you agree to let me use quotations, I will NOT include any names or nicknames you use, nor will I include identifying names along with the quotations.

Decision to quit at any time:

Your participation is voluntary; you are free to withdraw your participation from this study at any time. If you do not want to continue, you can simply not respond to the survey. You also may choose to skip any questions you do not wish to answer.

How the findings will be used:

The results of the study will be used for scholarly purposes only. The results from the study will be presented as a dissertation paper and will benefit the study of the Network for Educator Effectiveness.

Consent Form – Survey (cont’d.)

Contact information:

If you have any questions about your rights as a research participant, please contact Dr. Robert Watson, Missouri State University, by email at RobertWatson@MissouriState.edu.

By beginning the survey, you acknowledge you have read this information and agree to participate in this survey, with the knowledge you are free to withdraw your participation at any time without penalty.

VITA

Jennifer L. Katzin is currently the elementary principal at Logan-Rogersville Elementary. Prior to this, she served as an assistant principal at the Upper Elementary and a second grade teacher in the same district. She hopes to continue her growth as an educational leader so she can better serve her staff and students.