UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

OKLAHOMA TEACHERS' PERCEPTIONS OF THE QUALITATIVE PORTION OF THE TEACHER EVALUATION SYSTEM

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF EDUCATION

By

IVANA BEATTY
Norman, Oklahoma
2016

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/48168

OKLAHOMA TEACHERS' PERCEPTIONS OF THE QUALITATIVE PORTION OF THE TEACHER EVALUATION SYSTEM

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF EDUCATIONAL LEADERSHIP AND POLICY STUDIES

BY

_____
Dr. Hollie Mackey, Chair

_____
Dr. Joyce Brandes

_____
Dr. William Frick

_____
Dr. Jeffrey Maiden

_____
Dr. Ronna Vanderslice

© Copyright by IVANA BEATTY 2016
All Rights Reserved.

DEDICATION

I dedicate this dissertation to all the guardian angels who made me see my stumbling blocks as stepping stones

ACKNOWLEDGEMENTS

Reflecting on this moment, I am overwhelmed by the amount of family and friends in my life who never stopped encouraging and believing in me. This moment of accomplishment belongs to you too! Special thanks to the following people:

Stacy Shaneyfelt, my dearest friend. Thank you for not allowing me to give up. I am the most thankful for you as you have walked every step of this program with me No more late nights of writing, editing, crying, and you convincing me to keep going. WE DID IT.

Matthew Beatty, my brother, my rock. Your belief that I can do anything often makes me believe that I can. In life, people may come and go, but I am thankful that you are my constant. You give me wings.

Dr. Hollie Mackey, Committee Chair. Thank you for being my biggest cheerleader at the University of Oklahoma. From my very first class, I looked up to you and knew you were the mentor I wanted to push me beyond measure and you did not disappoint. Thank you for being you as I would have never completed this dissertation without you. EACS will never be the same with you… a true loss.

Joyce Brandes, William Frick, Jeffery Maiden, and Ronna Vanderslice, Committee Members. Thank you for your guidance and support. I will always be appreciative to each of you for agreeing to serve on my committee. Thank you for sharing this journey with me.

To my family and friends who survived, thank you for putting up with me when I was grumpy, tired, and could not participate in the fun. I could not have done this without your understanding and patience.

TABLE OF CONTENTS

Acknowledgements iv

List of Tables x

List of Figures xi

Abstract xii

CHAPTER I: INTRODUCTION 1

Introduction 1

Historical Roots of Teacher Evaluation 4

Purpose of Teacher Evaluation 6

Before Race to the Top in Oklahoma 11

Teacher Training for Implementation 12

Background of the Study 13

Research Problem 15

Purpose Statement 17

Statement of Intent 18

Conceptual Framework 19

Organizational Theory 21

Schools within Organizational Framework 25

Teacher evaluation through an Organizational Lens 27

Research Questions 29

Significance of the Study 30

Review of the Research Method Used 31

Methodological Assumptions 32

Limitations and Delimitations 33

Summary 34

Operational Definitions 35

Overview of Chapters 37

CHAPTER II: LITERATURE REVIEW 39

Introduction 39

History of Teacher Evaluation 39

Relationship between the administrator and teachers 41

Why Do We Have Teacher Evaluation? 43
Promote Growth of Teachers 45
Removal of Ineffective Teachers 45
Improvement of Professional Development 48

Capturing Teacher Effectiveness 49
The Different Types of Evaluation Sources 51
Summative Versus Formative Evaluation 52
Best Practices 53
Is the System Irreparable or in Need of Repair? 55

Policies 56
National Policy 56
Oklahoma State Policy 58
2010 School Laws of Ok. Chapter 1- Ok. School Code Article V: School Districts and Boards of Education 60

Qualitative Measures: The Evaluation System Chosen 62
Tulsa TLE Framework 62
Marzano Framework 63

Quantitative Measures 63
Value-Added Measures 64
Student Learning Outcomes/Student Objective Outcome 64
Other Academic Measures 66

Oklahoma Data Collections 66
Gaps in the Research 69

The Continuously Changing Policy Landscape: Oklahoma TLE System 71

Teacher Training for Implementation 72

Perceptions 74
Administrators' Perception of Teacher Evaluation 75
Teachers' Perception of Teacher Evaluation 76

Summary 80

CHAPTER III: METHODOLOGY 81

Purpose 81

Research Questions Re-stated 81

Sample 82
Superintendent Data Collection 83
Teacher Data Collection 83

Confidentiality 84

Research Design 84
Questionnaire Development Process 85
The Teacher Evaluation Profile (TEP) 86
Researcher Questions 89

Pilot Study 92

Instrumentation 94
Validity and Reliability 95

Data Collection 97

Data Analysis 99

Validity 100
Internal Validity 100
External Validity 101

Methodological Assumptions 101

Limitations and Delimitations 102

Summary 103

CHAPTER IV: RESULTS 104

Overview 104

Review of Research Methods, Design, and Instrumentation 105

Research Questions Restated 105

Data Collection 106

Data Analysis 107
Official Questionnaire Reliability Analysis 107

Descriptive Statistics 108

Means 109

Comparison Means 113

Inferential Analysis: Multiple Regression 115
Implementation 115
Improving Teacher Growth 117
Removal of Ineffective Teachers 118
Professional Development 119
Policy 120

Summary 121

CHAPTER V: DISCUSSION 123

Summary of the Purpose 123

Summary of the Literature 124

Research Questions 124

Summary of the Methodology. 125

Summary of the Findings 125
Descriptive Statistics 126
Means Scores 126
Comparison Means 126
Multiple Regression 127

Research Questions Answered 128
What are Teachers' Perceptions Regarding Teacher Training for Implementation? 128

What are Teachers' Perception s about the Purpose of the Evaluation in Regards to improving Professional Growth? 128
What are Teachers' Perceptions about the Purpose of the Evaluation in Regards to Removing Ineffective Teachers? 129
What are Teachers' Perceptions about the Purpose of the Evaluation in Regards to Receiving Professional Development? 129
What are Teachers' Perceptions about the Purpose of the Evaluation In Regards to Policy? 130

Connections to the Literature 131
Implementation 131
Professional Growth 131
Removal of Ineffective Teachers 132
Professional Development 132
Policy 133

Recommendations 133

Suggestions for Further Research 135

Summary 135

REFERENCES 137

APPENDICES 146

A. Questionnaire 146

B. Permission to use Teacher Evaluation Profile (TEP) 154

C. Model Section Criteria for Tulsa Teacher and Leader Effectiveness Framework and the Marzano Causal Teacher Evaluation Model 156

D. Teacher Evaluation Profile (TEP) 161

E. Oklahoma Criteria for Effective Teachers Performance Evaluation 167

F. Superintendent Letter of Approval 169

G. Teacher Recruitment Letter 171

H. Doherty's Teacher Evaluation Profile (TEP) for Teachers 172

LIST OF TABLES

Table 3.1: Teacher Evaluation Profile for Teachers aligned to my Questionnaire Constructs 87

Table 3.2: Content Validity 91

Table 3.3: Reliability Scores 97

Table 4.1: Frequencies and Percent of Questionnaire Demographics 109

Table 4.2: Questionnaire Statistics and Frequencies 110

Table 4.3: Total Possible Point and Means Report 113

Table 4.4: Comparison Means 114

Table 4.5: Categorical Independent Variables 115

Table 4.6: Sim. Multiple Regression Model Summary for Implementation 116

Table 4.7: Sim. Multiple Regression Results for Regression Equation 116

Table 4.8: Sim. Multiple Regression Model Summary for Improving Teacher Growth 117

Table 4.9: Sim. Multiple Regression Results for Regression Equation 118

Table 4.10: Sim. Multiple Regression Model Summary for Removal of Ineffective Teachers 118

Table 4.11: Sim. Multiple Regression Results for Regression Equation 119

Table 4.12: Sim. Multiple Regression Model Summary for Professional Development 120

Table 4.13: Sim. Multiple Regression Results for Regression Equation 120

Table 4.14: Sim. Multiple Regression Model Summary for Policy 121

Table 4.15: Sim. Multiple Regression Results for Regression Equation 120

LIST OF FIGURES

Figures 1: Evaluation Theory Tree 20

Figures 2: Evaluation Organization Flowchart 26

ABSTRACT

This quantitative study explores teacher evaluation from a teachers' perspective through an organizational lens. Political reforms such as Race to the Top has ignited state reforms within the teacher evaluation framework. Since 2012 Oklahoma has gone from the initiation phase to the incorporation phase of implementing the qualitative portion of the new Oklahoma Teacher and Leader Effectiveness (TLE) System. The study is a non-experimental quantitative analysis that uses questionnaires to collect data. There are two purposes for the study: 1) to quantify Oklahoma teachers' perceptions to determine levels of favorableness towards the new teacher evaluation frameworks, which is the qualitative portion of the Oklahoma TLE; and 2) to quantify Oklahoma teachers' perceptions to determine levels of favorableness towards district teacher training for its implementation. With these questions, the study found that teachers were mostly unfavorable or neutral. Five multiple regression were run and all showed statistical significance except for the removal of ineffective teachers. The study identified the variance among teachers' perceptions across frameworks, teacher experiences, locations, and school levels. From the results, various stakeholders will be able to utilize the data as a communication tool to improve as the Oklahoma TLE system continues towards full implementation.

Keywords: Teacher evaluation, Teacher perceptions, Organizational change

CHAPTER I

Introduction

Teacher accountability is a highly controversial topic in the field of PK-12 schooling, in both theory and practice. Teachers earn degrees and certifications that demonstrate their ability to understand and apply specific content knowledge from their chosen fields, and then transfer that knowledge to students (Nolan & Hoover, 2005). Therefore, earning certification in content areas that provide teachers the designation of being 'highly qualified' and serve as an entrance requirement into the profession begs the question; why are teachers increasingly asked to "prove" their mastery of content knowledge and instruction despite their demonstration of skills required to earn the designation of 'highly qualified', presumably meaning these teachers are effectively prepared to teach children?

Teacher evaluation is a "function designed to make comprehensive judgments concerning teacher performance and competence for the purpose of personnel decisions such as tenure and continuing employment" (Nolan & Hoover, 2005, p. 26). Traditionally, teachers have been required to annually demonstrate that they are highly qualified and effectively prepared to teach children through assessment based on administrators' observation. Teacher accountability is facilitated through teacher evaluations comprised of observations, portfolios, rating scales, collaborative efforts, and other forms. In addition, teachers are assessed through other academic measures and value added methods, which ranks educators by students' test scores. The observation is measured with an evaluation tool comprising a portion of the process of teacher evaluation. If teachers are following the protocol to enter into the field of

education and undergo annual observational evaluation, why are critical stakeholders asking for a so-called *more* realistic and effective evaluation tool to determine teacher effectiveness? Further, what are teachers' perceptions of the purpose for this new evaluation requirement? What are teachers' perceptions of the teacher training that prepared them for the implementation of new measures included in the new evaluation tool?

With the declining public perception of both teachers and education, policymakers are enacting policies, regulations, and mandates in an attempt to improve teacher effectiveness and "fix" education (Alexandrov, 1989**;** Clark**,** 1993; Danielson & McGreal, 2000; Derrington & Campbell, 2013, Sawchuk, 2015). With an awareness of the continuously changing policy landscape, this study seeks to address educational reform based on the American Recovery and Reinvestment Act of 2009 (ARRA). The federal policy includes an embedded state level grant opportunity titled, "Race to the Top" (RTTT) (US Department of Education [USDOE], 2009b). The Race to the Top grant is a $4.35 billion dollar endowment intended to support the work of innovative states (USDOE, 2009b). To qualify for the grant award, the United States Department of Education requires states and districts to implement major revisions to their educational policies and procedures to meet numerous criteria, such as: (a) state success factors, (b) standards and assessment, (c) data systems to support instruction, (d) great teachers and leaders, (e) turn around the lowest achieving schools, and (f) general selection criteria (USDOE, 2009b). Oklahoma and other competing states have and continue to address educational reform aligned with these criteria to qualify for receiving RTTT funds. The purpose of this study is to examine the degree of

favorableness in Oklahoma teachers' perceptions of teacher evaluation reform and implementation aligned with Criteria D, Section two (CDS2) of the Race to the Top Grant within ARRA that addresses teacher evaluation.

Oklahoma intends to fully implement a new teacher evaluation system called the Oklahoma Teacher and Leader Effectiveness Evaluation (TLE) (Oklahoma State Department of Education [OSDoE], 2012). The process began in academic year 2013-2014 when Oklahoma policy mandated all schools completely integrate one of three preselected evaluation frameworks (Marzano Causal Teacher Evaluation Model, Tulsa Teacher and Leader Effectiveness Frameworks, or Danielson's Evaluation Framework) to meet the CDS2 of the RTTT grant and the qualitative portion of the Oklahoma TLE system. As such, Local Education Agencies (LEA) across the state began the initial three-year process of revamping policy, training educators, and incorporating the TLE system into their districts. Oklahoma has 43,840 teachers and 516 school districts (C. Hassell, personal communication, October 25, 2013; K. Isenhour, personal communication, April 14, 2015). The Tulsa framework was selected by 483 school districts, 50 school districts selected the Marzano framework, and no school districts selected the Danielson Framework. The calculation include some of the 24 charter schools that have chosen to participate.

The literature base for teacher evaluation is extensive, yet the breadth of literature concerning teachers' knowledge is limited in terms of perceptions of teacher evaluation and its implementation through the lens of quantitative analysis. Most teacher evaluation research is rooted in qualitative analysis. Research specifically identified as qualitative case studies when studying teacher perceptions from the few

existing studies (Wormmeester, 2005). This trend may change as states approach full implementation of their evaluation systems; the RTTT initiatives further require teachers' perspectives of the system and its implementation (USDOE, 2015; Jiang, Sporte, & Luppescu, 2015). The lack of quantitative studies about teachers' perspectives of teacher evaluation validate the need for further analysis of quantitative data (Milanowski & Heneman, 2001; Tuytens & Devos, 2009; Wormmeester, 2005; Jiang, Sporte, & Luppescu, 2015). This empirical study intends to partially fill the literature gap within the current scholarship.

Historical Roots of Teacher Evaluation

Evaluation of teachers is a process that originated as early as the pre-civil war era (Danielson & McGreal, 2000; Nolan & Hoover, 2005). Nolan and Hoover found evaluations were completed by clergymen or school masters acting as inspectors to address the upkeep of facilities and management. Collaboration between inspectors and teachers was nonexistent. Personnel had specific job requirements and executed those specified responsibilities with precision (Nolan & Hoover, 2005). Supervisor and superintendent positions were developed in the late 19th century (Nolan & Hoover, 2005). Evaluations were primarily aimed at assessing the effectiveness of personnel management more than evaluating the effectiveness of teachers' instruction abilities.

It was not until the 20th century that evaluations shifted toward teachers and their ability to teach effectively (Danielson & McGreal, 2000; Nolan & Hoover, 2005). The Industrial Revolution spurred focus on increasing efficiency, and the influence of the movement resulted in greater efficiency in evaluating teachers. As a result, multiple rating scales were developed to guide administrators in determining teachers' level of

effectiveness (Danielson & McGreal, 2000; Nolan & Hoover, 2005). Rating scales varied in purpose through the years and continue to be used today (Danielson & McGreal, 2000; Nolan & Hoover, 2005). These evolving scales provided administrators the requisite data for guiding administrative decisions and making judgments about the qualities of a successful teacher (Nolan & Hoover, 2005), recognizing viable candidates for the school of pedagogy (Alexander, 1957; Danielson & McGreal, 2000), establishing teachers' salary (Alexander, 1957), and recognizing identifiers leading to teachers' level of effectiveness to determine projected student success (Beller, 1971; Danielson and McGreal, 2000; Derrington & Campbell, 2013; Nolan & Hoover, 2005). As the rating scales were incorporated into the educational system, their objectives varied between management of staff and reflection for teacher improvement. As policies changed and better practices developed, a debate of purpose between administrators and teachers became increasingly evident resulting in confusion as the role of administrators fluctuated between inspector, helper, evaluator, and counselor (Danielson & McGreal, 2000; Nolan & Hoover, 2005).

The relationship between administrators and teachers became more ambiguous as rating scales for differing purposes advanced (Nolan & Hoover, 2005). Administrative role ambiguity emerged between the administrator's role as the supervisor for accountability and the role as supervisor for improvement. This caused a breakdown in expectations among teachers who questioned why administrators were entering the classroom at any given time. Teachers found collaborating with administrators difficult because they were wary of exposing their strengths and weaknesses about teaching and learning when they did not fully trust the administrator.

Teachers did not know if an administrator was observing them to encourage growth or remove them from the teaching force (Nolan & Hoover, 2005). Administrators were also often unclear of the purpose for entering the classroom. At times the focus was to encourage teachers and promote growth and other times it was to gather evidence that supported removing a teacher. Consistency was lacking as the administrator's role vacillated between collaborative to stringent or staying distant to give a true evaluation while working towards maintaining a level of trust to continue a positive working relationship (Nolan & Hoover, 2005). Mixed perceptions between administrators and teachers created what Nolan and Hoover call a "tug-of-war between evaluative and helping functions of the supervisor" (2005, p. 23). The historical roots that created confusion and mixed perceptions of teacher evaluation still permeate modern evaluation schemas. There is still a disconnection between administrators and teachers views regarding the purpose of teacher evaluation (Nolan & Hoover, 2005).

Purpose of Teacher Evaluation

Teacher evaluations are used to measure educators' performance and to promote professional development (Danielson, 2010). Traditionally, using teacher evaluations is a long-term method for assessing teacher effectiveness. Depending upon the Local Education Agency (LEA), teachers formally worked within a one page checklist completed bi-annually or annually to establish teacher effectiveness (Derrington, 2011; Papay, 2012). The concern was, based on a given teacher's school district policy and school culture, once evaluations were completed a status quo seemed to perpetuate due to minimal administrative and teacher interaction to create a catalyst for teacher growth or removal of ineffective teachers (Donaldson, 2009; Waintroob, 1995). Poor evaluation

systems did not seem to separate the effective teacher from the ineffective teacher and administrators rarely objectively recognized either side of such a spectrum (Donaldson, 2009; Waintroob, 1995).

School administrators must use the evaluation tools for their intended purposes of promoting teacher growth, removal of ineffective teachers, and improving professional development. In other words, teacher accountability and professional development are central to effective evaluation (Weems & Rogers, 2010). What benefit is a teacher evaluation system if it is not performed well? Administrators have a role in determining how a program is implemented, used, and sustained; a critical element for performing evaluations well. Administrators must be familiar with what is happening in the classroom and aware of any challenges teachers are facing. This requires greater collaboration and distributive leadership between administrators and teachers. Administrators and teachers working with joint effort provides administrators the opportunity to engage in more instructional leadership and less personnel management (Elmore, 2000; Goldrick, 2002; Kyriakides & Demetriou, 2007). This effort serves to decrease confusion through effective collaboration and role clarification between the administrator as the supervisor and the administrator as the evaluator.

Leadership is in transition from a top-down approach towards a distributed approach as administrators spend more time in classrooms (Elmore, 2000; Goldrick, 2002; Kyriakides & Demetriou, 2007; Range, Scherz & Holt, 2011; Rowan, 1990; USDOE, 2015). Administrators have assumed a more active role within instructional leadership as accountability demands have increased. As teacher evaluations become more sophisticated and increased administrative accountability includes identifying a

teacher's level of effectiveness, administrators must be cognizant of teachers' attitudes and perceptions of the process (Derrington & Campbell, 2013). This recognition creates the opportunities for dialogue between the two, assisting educators in using the evaluations to improve their teaching (Wagoner & O'Hanlon, 1968). Greater administrative awareness facilitates administrators and teachers working as a unified team because administrators can approach individual teachers with their unique needs in mind. Wagoner and O'Hanlon explain, "Reaction towards an evaluation can range from threat (a negative reaction) to challenge (a positive reaction) (1998, p. 472)". Accordingly, administrators should approach a teacher who has a negative attitude towards teacher evaluation differently than a teacher whose attitude is positive if they want to encourage improved instructional performance (Derrington & Campbell, 2013; Wagoner & O'Hanlon, 1968). Differing attitudes among teachers result in varying levels of effort and involvement with the evaluation tool. These differing attitudes and perceptions make it difficult to use teacher evaluation systems to improve teacher effectiveness (Danielson & McGreal, 2000; Derrington & Campbell, 2013).

The struggle to understand the relationship between administrators and teachers is not new (Nolan & Hoover, 2005). Principals strive to create a positive working relationship with their staff, but serve as the ultimate advocate for student learning at the same time. Principals are responsible for student success within their buildings (Derrington & Campbell, 2013). The process of evaluation can cause teachers to become more leery of administrators' intentions as greater partnership and collaboration are required to meet other accountability needs (Nolan & Hoover, 2005). Is the principal the inspector or the helper? Is she the evaluator or the counselor? It can be

difficult for teachers to develop and maintain positive relationships and trust with their supervisors when they do not know which administrative hat she is wearing when she walks through the door.

Lack of trust may be attributed to teachers' expectations being skewed (Weisberg et al., 2009). Most teachers believe that they are working at the highest level of performance rankings on an evaluation scale. This perception results in anything less than the highest score available as a personally- directed insult or attack rather than an opportunity for professional growth (Weisberg et al., 2009). "Changes in teacher evaluation policies have the potential to significantly affect every part of a school, including all aspects of adults' behaviors, skills, attitudes, beliefs, and working relationships" (Derrington & Campbell, 2013, p. 236). Administrators must keep this in mind as they work to maintain positive school culture. West and Derrington state since the principal serves as the overseer and implementer of these changes, a positive supervisory relationship based on trust is necessary for developing an effective evaluation process and maintaining positive teacher relationships (as cited in Derrington & Campbell, 2013, p. 237).

Why is there so much disequilibrium with teacher evaluation systems when it is universally accepted that evaluations provide the data necessary for determining teachers' strengths and weakness and developing strategic professional development to improve instruction? It is not the general concept or purpose of evaluation that causes frustration among educators, but the often ambiguous implementation of evaluation systems as a whole (McGreal, 1983). Further, evaluating teachers with an evaluation system intended to maintain effectiveness is highly complex (McGreal, 1983). One

evaluation system cannot capture the full picture of educator effectiveness, just as one visit into a teacher's classroom cannot measure the strengths and weaknesses of a teacher for the purpose of summative evaluations. Even if a single evaluation could capture the full picture, understanding the collected data presents an equally complex concern. Maslow and Kelly found, "The heightened attention on teacher evaluation policy, researchers and policy makers rarely identify the systematic use of teacher evaluation data to improve organizational performance as a goal of a strengthened teacher evaluation system" (2012, p. 601). Policy-makers usually opt for overhauling the whole evaluation system before improving specific areas of weakness within the current system as determined by longitudinal data analysis. These challenges give opportunist voices to paint teachers or the evaluation framework as ineffective when in reality the evaluation frameworks are used inappropriately. Meanwhile, scholars' produce sound, yet largely ignored studies that use empirical evidence to identify flaws within the implementation of evaluation systems and disconnection between stakeholder perceptions and the reality of teacher effectiveness.

Dismal standardized test scores are attributed as the primary reason for public dissatisfaction with teacher effectiveness (Goldrick, 2002). When test data are released and compared by district, state, and nation, stakeholders make judgments about students' performance without scrutinizing the whole picture. Erroneous causal connections are inferred where poor student performance is directly attributed to ineffective teachers, increasing the demand for better teacher evaluation. Ironically, most evaluators give *all* teachers positive ratings on summative evaluations despite the evaluation tool used, making it difficult to distinguish between effective and ineffective

teachers (Goldrick, 2002). For example, Donaldson (2009) found 100% of Chicago teachers had longitudinally received a satisfactory or above rating on summative evaluations for over four years. Similarly, 96% San Bernardino, Californian educators met or exceeded expectations for the 2002-2003 and 2003-2004 school years; only .1% (1 in every 930 teachers) of Illinois teachers received an unsatisfactory rating. Donaldson's findings demonstrate high numbers of teachers performing exceptionally well on the teacher appraisals, yet, similarly high numbers of students in the schools where these teachers worked were unprepared to pass their state achievement tests (2009). These examples offer stakeholders the opportunity to label the evaluation system as ineffective because administrators appear to be failing to identify the most and least effective teachers. Administrators are often unfamiliar with the specific criteria for measuring teacher effectiveness (Kyriakides, Demetriou & Charalambos, 2006) and fail to follow through with administrative decisions to either promote growth or remove ineffective teachers.

Before Race to the Top in Oklahoma

Oklahoma was already leading multiple states in addressing teacher evaluation improvements prior to Race to the Top (RTTT) and the resulting implementation of the state's most current evaluation initiative (National Council on Teacher Quality [NCTQ], 2011). The state legislature had mandated adopting the Oklahoman Criteria for Effective Teaching in 1987 (Oklahoma State Department of Education [OSDOE], 1999). The measures for evaluation have been modified since that time with varying results of effectiveness, but the criteria have been consistent. The 2012 update to the Oklahoma Criteria for Effective Teaching satisfied the federal requirement that teachers

meet the highly qualified status for No Child Left Behind (NCLB) measures (OSDOE, 1999).

The Oklahoma Criteria for Effective Teachers Performance generally includes visual indicators for administrators to quickly determine teacher effectiveness (OSDOE, 1999). It utilizes two categories: 'practice' and 'products' with specific indicators under each category. The practice category includes the indicators 'teacher management' and 'teacher instruction'. Each indicator includes observable actions that guide administrators in determining teacher effectiveness. Examples of 'teacher management' observable actions include: 'preparation', 'routine', 'discipline', and 'learning environment'. The evaluation provides several suggestions of observable actions to assist administrators in establishing teacher effectiveness under this indicator The 'product' category includes teachers' artifacts supporting instructional strategies (OSDOE, 1999) (see Appendix E).

Teacher Training for Implementation

Teacher training for implementation cannot be overlooked when designing teacher evaluation systems. There are two stages for successful implementation during teacher training: awareness of the instrument innovation (Tuytens & Devos, 2009) and understanding the complexity of the instrument (Sartain, Stoelinga, & Brown, 2011). Teachers will construct knowledge and develop their own perceptions towards teacher evaluation innovation as they develop an understanding for the new implementation. Therefore, it is essential that trainers help teachers understand the teacher evaluation innovation's importance with clarity. In addition, it is important that collaboration and

communication through the process is optimized to increase the likelihood of buy-in and stability of the innovation.

Background of the Study

Oklahoma has been unsuccessful in obtaining the RTTT grant having received scores ranging from 211 to 399 out of 500 possible points from grant evaluators (OSDOE, 2014). This poor performance has not deterred Oklahoma policymakers from advancing their goals to improve educational reform initiatives, particularly in teacher evaluation. The Oklahoma State Board of Education continues to adopt policies intended to meet various criteria within the RTTT grant recommended by the Teacher and Leader Effectiveness Evaluation System Commission. Oklahoma TLE is one of these policies.

Oklahoma passed legislation establishing the creation and authority of the TLE system in 2011 (Oklahoma Law on Oklahoma State Department of Education [OLOSDE], 2013). Under Oklahoma Policy, Section 118 (70 O.S. § 6-101.10), Evaluation of Teachers and Administrators section A-1, the teacher evaluation requirement was established, which states:

> Every policy of evaluation adopted by a board of education shall be based upon a set of minimum criteria developed by the State Board of Education, which by no later than the 2013-2014 school year, shall be revised and based upon the Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE) system developed by the State Board of Education as provided in Section 6 of this act. (OLOSDE, 2013)

Currently, every district in Oklahoma, in collaboration with teachers' unions, have adopted a policy for implementation of a new teacher evaluation framework within the TLE**.**

Information about the adoption and development of the TLE was distributed across districts in a top-down manner while the State Board of Education (SBoE) collaborated with the TLE commission to develop the teacher evaluation system. The SBoE determined the parameters of district decision-making for incorporating the new program. Oklahoma's SBoE chose to not require all districts to adopt a single evaluation system, allowing for some flexibility, but did limit local flexibility towards which evaluation systems could be selected (The Center on Great Teachers and Leaders [COGTL], N.D.). In Oklahoma, every district was given the opportunity to implement one of three acceptable evaluation programs fulfilling Oklahoma's TLE commission's standards: Marzano Causal Teacher Evaluation Model, Tulsa Teacher and Leader Effectiveness Frameworks, and Danielson's Evaluation Framework. Each evaluation framework exclusively represented the qualitative measure of the TLE system.

The Oklahoma Teacher and Leader Effectiveness Evaluation system is divided into two measures: qualitative and quantitative, which accumulate to a score of 100%. Implementation of the new teacher evaluation system has been cumulative, adding one measure at a time. The first phase of teacher evaluation reform began in the 2012-2013 school year with focus on the qualitative measure. Districts and administrators had the opportunity to choose one of the three pre-selected evaluation frameworks, offer training for teachers, provide administrators time to manipulate the software, and either pilot the new teacher evaluation framework or continuing with the current evaluation procedures. The second phase occurred during the 2013-2014 school year, when all districts in the state of Oklahoma were required to fully implement the chosen teacher evaluation framework as a qualitative measure of the TLE (Goldrick, 2002). This

implementation strategy appeared to reflect a trend across states when researching Race to the Top winners' State Department of Education websites and analyzing trends within the States of the States (NCTQ, 2011). In the same school year, all districts in Oklahoma participated in a no-stakes pilot for Other Academic Measures (OAM) and started training over Value-Added Models (VAMs), which, at the time, made up the quantitative measure of the teacher evaluation system (Glisson, 2014). The third phase occurred during the 2014-2015 school year where the term 'student growth measures' replaced the term 'quantitative portion' representing 35%, measured by either value added measures or student learning objective /student outcome objective (SLO/SOO) (OSDOE, 2014). During this phase, the objective was for school districts to start collecting data from the qualitative measures, OAMs, and student growth measures (OSDOE, 2014). The SLO/SOO was in the initial stage, therefore stayed in the training and implementation phase (AIR & OSDE, 2014) and later was revamped due to teacher push back (OSDOE, 2015). The TLE system will continue with incremental implementation until deemed fully operational.

Research Problem

The American Recovery and Reinvestment Act mandates states across the nation re-address teacher effectiveness and provides opportunities for additional funding through the Race to the Top grant. A portion of this political reform is improving the requirement that administrators use teacher evaluations as a tool to distinguish between teachers who are effective and ineffective. The result is to provide opportunities for growth, remove ineffective teachers, and establish better alignment with professional development and teacher needs. It is essential that administrators

understand the purpose of an evaluation tool attached to such high-stakes and recognize if teachers are mastering the listed objectives (USDOE, 2015). Administrators must also comprehend how teachers perceive the evaluation tool to gain a deeper understanding how to foster buy-in for the evaluation system (Derrington, 2011; Jiang, Sporte, & Luppescu, 2015). Gaining an awareness of teachers' perceptions is essential for true educational change in classroom practices. This change is possible through successful implementation of the new teacher evaluation policy (Tuyten & Devos, 2009). Knowledge of teachers' perception has the potential to overcome previous flaws and prompt positive teacher reactions to a system (Milanowski & Heneman, 2001). Moreover, teachers' acceptance of a system can help to identify factors that account for differences in teachers' favorableness (Milanowski & Heneman, 2001; Peterson & Comeaux, 1990; Tuytens & Devos, 2009) gaining increased buy-in of a program (Nolan & Hoover, 2005).

State policy required all Oklahoma schools to have fully implemented one of the preselected evaluation frameworks as part of the qualitative portion of the Teacher and Leader Effectiveness Evaluation system by the 2013-2014 academic year. This decision was based on future plans to apply for the Race to the Top grant. Teachers had no voice in the decision, yet it required teachers' full participation. Not knowing teacher's perceptions of and attitudes toward the qualitative portion of the TLE system is problematic for effective implementation and utility. This study seeks to quantify teachers' perceptions of the implementation of the new TLE system and the level of favorableness towards teacher evaluation's three purposes.

This study is bracketed within RTTT policy and could be encouraging set in new policy in the future.

Purpose Statement

There are two purposes for this study: (1) to quantify perceptions held by Oklahoma teachers to determine levels of favorableness towards the new teacher evaluation framework (the qualitative portion of the TLE); and (2) to quantify perceptions held by Oklahoma teachers to determine levels of favorableness towards teacher training of the TLE implementation. Do teachers believe that the new evaluation system will effectively meet the purpose of evaluations including improving teacher quality, removing ineffective teachers, and aligning professional development? In addition, are teachers' perceptions favorable or unfavorable toward the implementation of the new program? The concomitant intent was to identify perceptions and recognize variance among the identified perceptions.

Variance will demonstrate the relationships between perceptions, which encompass dependent factors and the independent factors including frameworks, teacher experiences, locations, and school levels. For this study, when looking at each of the perceptions listed in the research questions, is there variance in teachers' perceptions between the Marzano Casual Teacher Evaluation Model and the Tulsa Teacher and Leader Effectiveness Frameworks? Is there variance in teachers' perceptions between tenured and non-tenured teachers? Is there variance in teachers' perceptions between urban, suburban, and rural areas? Finally, is there variance in teachers' perceptions between elementary and secondary levels?

Zepeda & Ponticelli recommends, "Until researchers and decision makers know and understand more about teachers' beliefs, assumptions, values, and perceptions, the

theoretical perspectives are of little use" (1998, pg. 86). The purpose of the study is relevant because an awareness of teachers' perceptions has the potential to overcome previous flaws and help instill positive teacher reactions toward a system (Milanowski & Heneman, 2001). Further, as this study is set in the ARRA policy, teachers' acceptance of a system can further recognize factors that account for differences in teachers' favorableness gaining increased buy-in of a program (Milanowski & Heneman, 2001; Peterson & Comeaux, 1990; Tuytens & Devos, 2009). Therefore, teachers' advocates can use the study's results to persuade the state department of education, local school boards, as well, central office personnel who make decisions about teacher evaluation procedures in Oklahoma. These advocates negotiate with various stakeholders regarding how the new system should be implemented and how it will influence teachers professionally. Answers to the research questions can assist identified leaders and administrators in making more objective judgments about the TLE system as it advances toward implementation of the quantitative aspect of the full TLE system. In addition, they (the research questions) can help district administrators with professional development to further ensure successful implementation and school building administrators with teacher acceptance and eventual ownership to gain buy-in and sustainability of the TLE system.

Statement of Intent

The purpose of this quantitative study is to provide data about teachers' perceptions towards the Oklahoma Teacher and Leader Effectiveness Evaluation framework (TLE) and its implementation. Results of the data will increase awareness of how Oklahoma teachers perceive the newly applied teacher evaluation framework as

it continues to achieve full implementation of the system. To do this study, I collected and analyzed data on teachers' perceptions about the teacher evaluation framework and its implementation. The results show level of favorableness towards the research questions, relationships between variables and create predictions with the findings.

Conceptual Framework

Teacher evaluation can be embedded within multiple theories. Yet, when researching other dissertations and studies, I found that scholars have avoided any attachment to a theoretical lens aside from a specific evaluation framework. An example of how evaluation can fit within multiple theories is represented in Figure 1.1. Alkin and Christie (2004) developed an evaluation theory tree that portrays a trunk and its primary branches, including prominent theorists. From their demonstration, they illustrate:

> The trunk is built on a dual foundation of accountability and systematic social inquiry. These two areas have supported the development of the field in different ways. The need and desire for accountability presents a need for evaluation. The importance of accounting for actions or for resources used in the conduct of programs is particularly evident for programs supported by government entities. Accountability and control is not a limiting activity, but, rather, is designed to improve and better programs and society. The social inquiry root of the tree emanates from a concern for employing a systematic and justifiable set of methods for determining accountability. While accountability and control provides the rationale, it is primarily from social inquiry that evaluation models have been derived. (p.12)

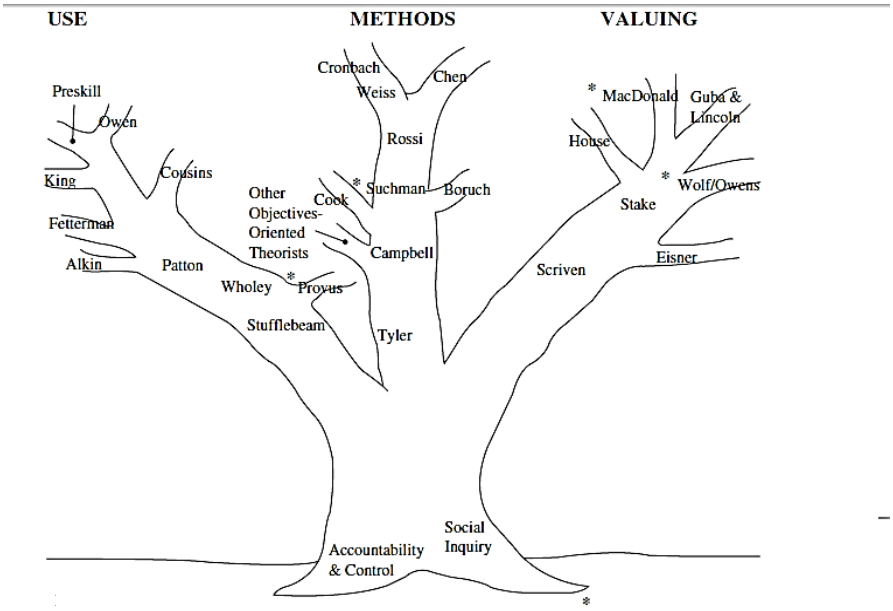


Figure 1.1. Evaluation theory tree. Reprinted from [adapted from] M.C. Alkin, & C.A Christie, 2004, An evaluation tree. In evaluation roots (Chapter two). Retrieved from http://www.sagepub.com/upmdata/5074_Alkin_Chapter_2.pdf.

All listed theories have a prominent place within different aspect of a teacher evaluation system, however, for the purpose of this study, I have selected an organizational theory lens, which is not a part of the evaluation tree, but looks at the whole tree from a broader scope. Alkin and Christie explain the three branches as discrete categories. M*ethods* deals with obtaining generalizability, or understanding the framework of teacher evaluation (internal factors). My study is not trying to analyze a specific evaluation framework and its methods. *Valuing* establishes the role of the evaluator to give value or to make judgements on data (or teachers) to determine teacher effectiveness (external factors). This study can portend valuing, but there are two main reasons the valuing framework is not a best fit. Valuing is from the administrators' perspective and it goes no further than making judgements towards teacher's effectiveness. Several scholars have used the valuing framework to investigate teacher efficacy by determining teacher motivation or level of ability to teach which is not what

I am investigating. *Use* originally was focused on decision making, but has broadened to include how information will be used and who will be using the information (external factors) (2004). All three branches address the design of the evaluation framework and how it is being manipulated either internally or externally. This study avoids the design of the chosen frameworks and focuses on the recent implementation of new frameworks triggered by federal and state policies. Additionally, it seeks to identify teachers' perceptions of the framework improving the purpose of evaluation from the previous evaluation framework. The focus of this study is on Oklahoma teachers and their perceptions towards the implementation of the new qualitative portion of the teacher evaluation system and the level of favorableness concerning its purpose.

The purpose of teacher evaluation is to: (a) promote growth of teachers (b) remove of ineffective teachers, and (c) improve of professional development. All of these actions overlap departments of a school's organizational system to improve the efficiency of the teaching workforce. Therefore, teacher evaluation works closer with organizational theory than other theoretical lenses.

Organizational Theory

Organizational theory is a field of study that investigates the impact of individuals, groups and structures upon behavior within an organization (Robbins, 1987). It works within a political system that concerns itself with hierarchy, control, authority, coordination, cooperation, and efficiency (Moe, 1995). *Reframing organizations* by Bolman and Deal addresses four frames of management within organizations: structural, human resource, political, and symbolic (1997). Each frame has its strengths and weaknesses, neither more capable than the other. As managers, the

goal involves acknowledging which frame(s) best fits an organization and recognizing the opportune moments to reframe as situations arise.

Structural Frame

The structural frame rooted by Max Weber's theories exemplifies the design of an organization and its productivity. The metaphor for this frame is a machine or a factory where everyone or thing has its job to make the organization run smoothly. The central concepts focus on rules, roles, policies, technology, among others. Awareness of this frame is essential to an organization's success because it not only helps to avoid misdirection of energy and resources, but intervenes and identifies places for enhancements and constrains for organizational accomplishment. There are six assumptions that define the structural frame (p.40):

1. Organizations exist to achieve establish goals and objectives.
2. Organizations work best when rationality prevails over personal preferences and external pressures.
3. Structures must be designed to fit an organization's circumstances.
4. Organizations increase efficiency and enhance performance through specialization and division of labor.
5. Appropriate forms of coordination and control are essential to ensuring that individuals and units work together in the service of organizational goals.
6. Problems and performance gaps arise from structural deficiencies and can be remedied through restructuring.

Human Resource Frame

Douglas McGregor laid the foundation for the human resource frame. This frame focuses more on the relationship between people and the organization. The human resource frame focuses on ways to establish a positive culture and uses symbols and/or mottos to give employees something to believe in. The metaphor for this frame signifies family as the central concepts are needs, skills, and relationships. It recognizes that both the organization and people need one another and with best fit, needs are met. There are four assumptions that define the human resource frame (p.102):

1. Organizations exist to serve human needs rather than the reverse.
2. People and organizations need each other: organizations need ideas, energy, and talent; people need careers, salaries, and opportunities.
3. When the fit between individual and system is poor, one or both suffer: individual will be exploited or will exploit the organization- or both will be victims.
4. A "good fit" benefits both: individuals find meaningful and satisfying work, and organizations get the talent and energy they need to succeed.

Political Frame

The political frame views organizations as active political arenas that host a complex web of individual and group interests. Unlike the before mentioned frames, this frame embeds struggle for power, which can shift depending on bargaining and negotiation agreements. The metaphor for this organization is a jungle where the central concepts are power, conflict, competition, and organization politics. There are five assumptions that define the political frame (p.163):

1. Organizations are *coalitions* of various individuals and interest group.
2. There are *enduring difference* among coalition members in values, beliefs, information, interest, and perceptions of reality
3. Most important decisions involve the allocation of *scarce* resources- who gets what.
4. Scarce resources and enduring differences give *conflict* a central role in organizational dynamics and make *power* the most important resource.
5. Goals and decisions emerge from *bargaining*, *negotiation*, and *jockeying for position* among different stakeholders

Symbolic Framework

The symbolic framework's goal is to maintain support in the eyes of its constituents. Symbolism creates a positive climate by using symbols, ceremonies, and beliefs to establish a culture of success. The metaphor is a carnival and or theater having central concepts around culture, meaning, ceremony, stories, among others. This framework aims to inspire by pulling at ones 'heartstrings.' Under this perspective, organizations are judged by their appearance. There are six assumptions that define the symbolic frame (p.216):

1. What is most important about any event is not what happened but what it means.
2. Activity and meaning are loosely coupled: events have multiple meanings because people interpret experience differently.
3. Most of life is ambiguous or uncertain-what happened, why it happened, or what will happen next are all puzzles.

4. High levels of ambiguity and uncertainty undercut rational analysis, problem solving, and decision-making.
5. In the face of uncertainty and ambiguity, people create symbols to resolve confusion, increase predictability, provide direction, and anchor hope and faith.
6. Most events and processes are more important for what is expressed than what is produced. They form a cultural tapestry of secular myths, rituals, ceremonies, and stories that help people find meaning, purpose, and passion.

Schools within Organizational Framework

School systems can be embedded within the organizational framework because they are complex political organizations that involve multiple systems collaborating internally and externally among departments (Giacquinta, 1973, Rowan, 1990). As far as frames, teacher evaluation aligns with three of Bolman and Deal's frameworks. First, from a general view, schools are, "Mechanistic management systems," state Rowan (1990). Leaders are making decisions, policies, and procedures to promote efficiency by routinizing teacher's input, output, and behaviors (Rowan, 1990). Second, teacher evaluation systems work within the human resource department to recruit and retain effective teachers that work with individual schools and districts to produce the output of effective teaching in the workforce to facilitate student success. The last frame, the symbolic framework, assesses the performance of individuals, which is the main idea of teacher evaluation. Bolman and Deal maintain that "evaluation is necessary to ensure a responsible, serious, and well-managed image" (1997, p. 244). The OCED (2005)

developed a conceptual framework, shown in Figure 1.2, of how teacher evaluation works with and around other organizations below to ensure the effectiveness of a teaching workforce:

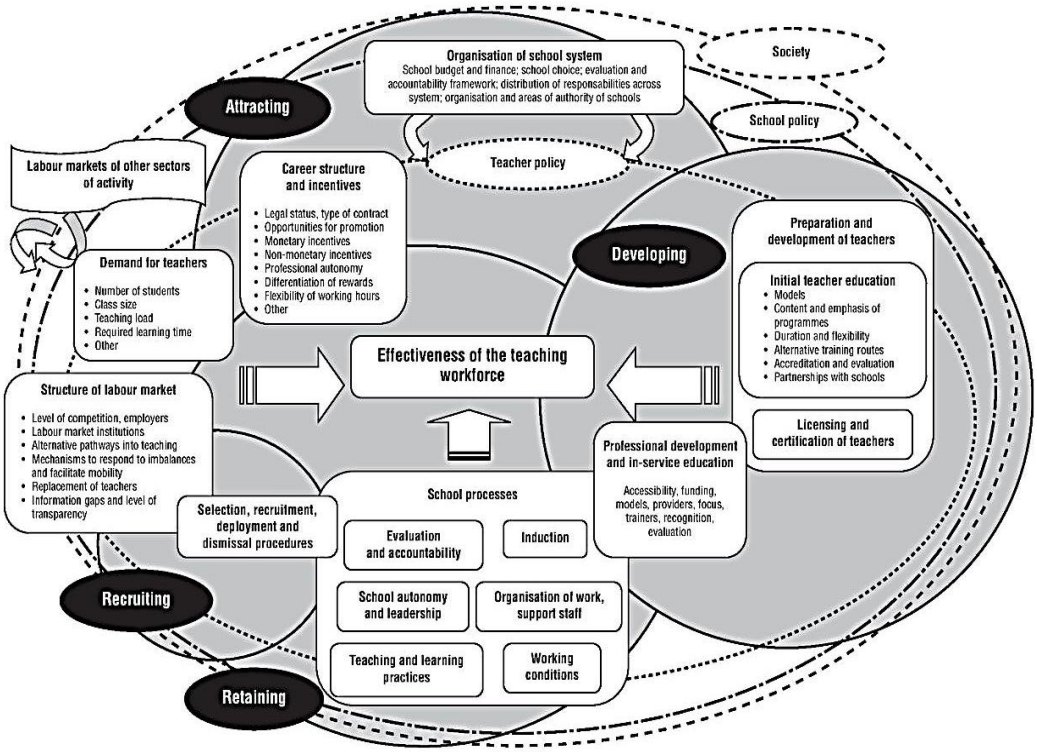


Figure 1.2. Evaluation Organization Flowchart. Reprinted [or adapted] from "OECD," 2005, *Teachers Matter: Attracting, Developing and Retaining Effective Teachers,* OECD, Paris. Doi: 10.1787/9789264018044-en.

Maslow and Kelley (2012) have labeled organizational theory as the best fit for research on teacher evaluation. They recognize the need to analyze evaluation data to identify expert knowledge of individuals, departments, or groups that can support school improvement. Analyzing data helps practitioners recognize areas of weakness to inform changes in organizational processes and systems that support teacher development, such as induction, mentoring, professional development, structured

professional collaboration time, and providing training and support for department chairs in an organizational framework.

Teacher Evaluation through an Organizational Lens

Schools are situated within an organization framework. There are multiple lenses to approach teacher evaluation. Two lenses closest to this study are: organizational control and organizational change. Darling- Hammond et al. (1983) and Rowan (1990) discuss organizational control. Through the organizational control lens, behaviors, e.g. staff development, curriculum, and teacher evaluation are manipulated externally from the classroom to influence teachers with the intention to acquire higher student learning outputs. In this case, teacher evaluation systems exemplify the manipulated behavior. Hence, the more accountability measures that the legislature mandates and require administrators to implement and assess through evaluation frameworks, the more control external stakeholders have to unify teachers to given behaviors (Darling-Hammond et al., 1983; Rowan, 1990; Tuytens & Devos, 2009). Making organizational control work as a mechanical device in a well-oiled organization (Darling-Hammond, et al., 1983; Rowan, 1990).

Through the organizational control lens, teacher evaluation is a behavior being implemented in schools in hopes of developing and increasing a better organization. The study could situate itself within this lens as it investigates a policy driven reform to control an organization internally. Yet, this is not the primary focus of this study. The major emphasis is to see teachers' behaviors to create change and not the evaluation framework. This study addresses teachers' perceptions towards the behavior to

determine the level of favorableness not how the behavior is influencing teachers and their output. This makes the organization control lens not appropriate for this study.

For this study, teacher evaluation works best under the lens of organizational change. Some early pioneers were Durkheim, Parsons, and Weber in the 1950s (as cited in Fullen, 1998), but it did not solidify until the 1960s when equity and civil rights reforms were ignited (Fullen, 1998). Other prominent theorist in this field include Schein & Bennis (1961); Gross, Giacquinta, & Bernstein (1971); and Hage and Aiken (1971) discussed by Giaquinta (1973) and Fullen (1998). In the beginning, organizational change was "Laboratory based, detached from the day-to-day instructional issues and functioning of schools" (Fullen, 1998, p. 202), but has since expanded. Hage and Aiken explains organizational change:

> Is not concerned with changes in individuals, such in their abilities, interest, behavior and motive. Instead, in jobs and their arrangements and how these relate to changes in the functioning of the organization. (1970, p. 13)

The stages of organizational changes have changed slightly over time, but currently, organizational change in schools encompasses three stages: Initiation, implementation, and incorporation. *Initiation* is a process that, when successful, leads to the introduction of organizational innovation (Giaquinta, 1973). Innovations can be a product, service, technology, or administrative practice. For this study, the Race to the Top Grant is the administrative innovation. States, in an attempt to receive the RTTT grant, began to revamp their teacher evaluation systems to meet Criteria D Section 2. Innovations are essential, but before they can accomplish its intentions, it must be recognized and implemented by organizational members (Gross, Giacquinta, & Bernstein, 1971). *Implementation* is a process that when successful, results in the

alteration of organizational members' behaviors and attitudes so they can conform to the expectation of the innovation (Giaquinta, 1973). This is the implementation of the teacher evaluation frameworks, which aims to increase teacher accountability. Further, it addresses the study's research question to quantify teachers' perceptions towards teacher training for implementation and the purposes of the new evaluation system. *Incorporation* is a process leading to stability of the new behavior so the innovation becomes a regular part of the school's organization (Giaquinta, 1973). This process is addressed in the results of the study when quantifying the level of favorableness teachers feel towards the new teacher evaluation framework and its implementation. This study aligns most closely within organizational change as the lens is not trying to manipulate teachers, but allow, because something has happened, to assess the organizational members' attitudes and perceptions towards the behavior.

Research Questions

The primary purpose for this study is to quantify teachers' perceptions towards the new qualitative measure of the Teacher and Leader Effectiveness Evaluation system. The following illustrates the study's conceptual structure and organizes the information related to the identified areas of interest:

1. What are teachers' perceptions regarding teacher training for implementation?
2. What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth?
3. What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers?

4. What are teachers' perceptions about the purpose of the evaluation in regards to receiving professional development?

Significance of the Study

Research that focuses on teachers' perception of teacher evaluation systems is scarce (Tuyten & Devos, 2009; Jiang, Sporte, & Luppescu, 2015). In fact, Milanowski & Heneman identified only three reputable studies in 2001 that were conducted prior to their study (2001). There are numerous studies with a concentration on the history and purpose of teacher evaluation; administrator as evaluators and their perception; the reasons that the evaluation system is broken and ways to improve; however, there is limited research on teachers' perceptions of evaluation systems. This disparity makes available a broad range of areas for research, a gap, because teacher perception epitomizes a vital key to the acceptance of an evaluation system. As McGreal endorsed back in 1983, evaluation focuses on partnerships between administrators and teachers (McGreal, 1983). A school district can adopt a great evaluation tool, but administrators and teachers must ultimately be willing and able to collaborate together to safeguard effective teaching and student learning outcomes. This partnership cannot develop without the trust and awareness of teachers' perceptions and attitudes (Derrington, 2011; Minnici, 2014). The results will not only help stakeholders in establishing evaluation procedures that are productive for teachers, but will further enable administrators to use the tool to improve the school's overall organizational system and student success.

The desired outcome is the availability of data to determine if teachers in Oklahoma feel that the new teacher evaluation frameworks were implemented

effectively, and if teachers attribute the framework for their professional growth improvement, the district's professional development and removal of ineffective teachers. The secondary desired outcome is to recognize if there is or is not significant variance among teachers' perceptions across frameworks, teacher experiences, locations, and school levels.

From the results, various stakeholders will be able to utilize the data as a communication tool. Policy makers can utilize the results when given the opportunity to share positive or negative findings to inform public policy by demonstrating whether progress in Oklahoman teacher evaluations is ensuing. The outcomes can also represent advancement towards meeting the goals within the Race to the Top application or future innovations in education. In addition, advocates in education can apply the results and use the suggestions to collaborate with administrators and teachers within an evaluation framework to establish an efficient organization.

Review of the Research Method Used

The study was a non-experimental quantitative analysis that used questionnaires to collect data. The instrument employed was a questionnaire that involved a Likert scale of closed ended questions to gather data from teachers. I distributed a questionnaire through district emails after I obtained approval from district superintendents.

My initial target sample included all Oklahoma public schools that participated in the Marzano Casual Teacher Evaluation Model (Marzano) and the Tulsa Teacher and Leader Effectiveness Framework (Tulsa). After superintendent recruitment to gain

teacher access, the new sample size is defined by how many usable questionnaires were returned.

This study correlates to the field of social sciences and aims to quantify teacher perceptions. Since this study is being specifically linked between teachers' perceptions in Oklahoma and the newly implemented teacher evaluation system, I developed a customized questionnaire. I modified an established questionnaire and merged my research questions based on the literature to ensure reliability and validity measures. I followed expert techniques to initiate and implement a pilot study to establish reliability scores after I developed the instrument, (Wagoner & O'Hanlon, 1968).

Methodological Assumptions

I made several assumptions for this study including the following: the methods and procedures selected for this questionnaire are appropriate for the subjects studied; the teachers will honestly and accurately answer all questions to the best of their knowledge; the participant teachers represent a purposeful sample of the targeted population identified for this research study; my analysis of the data will be accurate and represent responses of all data collected; email will be an appropriate method to collect information for this research study; and teachers have access to the Internet and the abilities needed to use and complete the questionnaire. Further, teachers will demonstrate an understanding of the information regarding their districts, school sites, and the evaluation systems to accurately respond to the questionnaire. The findings and conclusions of the study will not be generalizable past the limitations of the study.

Limitations and Delimitations

This research study utilized data returned from an Internet questionnaire. This methodology results in limitations. There is limited knowledge on teachers' perceptions of teacher evaluation. Therefore, the literature review base is restricted. The data for the study is limited to the teachers who responded to the questionnaire. The data analysis was completed through a multiple regression test that assesses correlations in variables; however, correlation does not imply causation (Creswell, 2009). Data is self-reported, therefore "the validity of the information is contingent on the honesty of the respondent" (Mertens, 2010, p. 173).

Delimitations. The research review focuses on teachers' perceptions with limited research from other points of view. This study is a quantitative design. The inclusion of a qualitative aspect may develop a more balanced study of teachers' perceptions. The survey included only a closed ended Likert scale, rather than including a section for comments. The review of literature is bracketed within the Race to the Top initiative and excludes previous policy, such as No Child Left Behind reforms and future ones, namely the Dec. 10, 2015 Every Student Succeeds (ESSA) Act. The review of literature emphasizes the purpose of the teacher evaluation and limited information on the use or the various types of evaluation. The topic of teacher evaluation inevitably overlaps with other topics such as value-added, performance pay, termination, teacher effectiveness, accountability systems, and numerous other fields. However, the intent of this research targets the purpose of teacher evaluation and the need for teachers' perspectives. The study's participants were limited to teachers who

had participated in either the Marzano Causal Teacher Evaluation Model or the Tulsa Teacher and Leader Effectiveness frameworks.

Summary

The American Recovery Reinvestment Act allowed states across America the opportunity to accept the challenge to improve teacher effectiveness and apply for a Race to the Top grant. Oklahoma started aligning the components required to be approved as they applied for the grant. This brings forth the dilemma of whether reforms such as teacher evaluation should be written into policy or district-driven. School districts did not have an option if they wanted to participate or not, their only choice was in selecting a framework (Tulsa or Marzano) for evaluating administrators and teachers. Teacher voice was minimal as new policies and procedures ensued and phases of implementation ignited. With all the changes and trainings, what were teachers' perceptions towards the new policy and how it was implemented? In this study, the primary focus is towards the qualitative portion of the system. Effectiveness of teacher focused programs relies on the buy-in of teachers, so to guarantee success of true implementation, it is also vital to ask, what are teachers' perceptions towards the purpose of teacher evaluation? There are three purposes of teacher evaluation: to improve teacher growth, remove ineffective teachers, and improve professional development. Yet, how do teachers perceive these three facets of teacher evaluation? In addition, do teachers perceive the new frameworks as an improvement towards the three purposes of teacher evaluation? The answers to these questions could foreshadow the success of the Oklahoma TLE system as the Oklahoma TLE Commission continues to collaborate with the State Board to fully implement the system to meet the

requirements of the RTTT grant. Fully implementing the program requires effective teacher training resulting in improved teaching competence and an output of increased student success.

The study is framed within an organizational lens as school districts work with a human resource department to maintain an effective teacher workforce. At the school level, school administrators maintain an effective work force through teacher evaluation. This procedure can cause disequilibrium between administrators and teachers.

Operational Definitions

The definitions used in this study were collected from current literature. The following terms are germane to this study:

Behaviors. Behavior controls are standardized teaching practices.

Elementary School. For the purpose of this study, elementary school represents schools that have the term elementary in their title.

Evaluator. Person (usually an administrator) who has the responsibility to make a fair assessment of the teacher's performance and competence using district standards and the evidence collected during the process (Nolan & Hoover, 2005).

Formative Evaluation. Examines how teachers can improve by identifying the needs for professional development and making the resources available (Kyriakides & Demetriou, 2007).

Growth plan (Also known as Improvement Plans). For the purpose of this study, growth plan and improvement plan is an agreement signed between teacher and

administrator to help improve a teacher who has been identified as ineffective on a teacher evaluation.

Innovations. Innovations are an adoption of an idea or behavior that is new to an organization (Hage, 1999).

Non-tenured teachers. For the purpose of this study, non-tenured teacher is a teacher that has been teaching less than 3 years.

Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE). For the purpose of this study, commission is a group of experts gathered by the State Board of Education to develop the TLE system.

Professional Development. For the purpose of the study, professional development signifies training to promote continued teacher development; the output is to improve student learning (Papay, 2012).

Race to the Top. A grant program that is rewarded to states that meets specific criteria under four educational reforms (USDOE, 2009).

Secondary Schools. For the purpose of this study, secondary schools are defined as schools that are middle schools, junior highs, and high schools.

Summative Evaluation. Examines the scores or results of the evaluation tool and can be used to determine career advancement; implements performance rewards; or establishes sanctions for underperforming teachers (Kyriakides & Demetriou, 2007).

Teacher Effectiveness (Also known as performance appraisals). A teacher whose students achieve acceptable rates (e.g., at least one grade level in an academic year) of student growth.

Teacher Evaluation. The primary tool used to promote and/or improve teaching and determine a teacher's level of quality; an ongoing data gathering process that enables an evaluator to discover, document, and verify a teacher's strengths and weaknesses (Danielson & McGreal, 2000).

Teacher Evaluation Framework. For the purpose of this study the teacher evaluation framework is defined as the preselected teacher evaluations by the TLE Commission. Both the Marzano Casual Teacher Evaluation Model and the Tulsa Teacher and Leader Effectiveness Framework.

Teacher Evaluation System. For the purpose of this study the teacher evaluation system is defined as the TLE system as a whole representing both the quantitative and qualitative measures.

Tenured teacher. For the purpose of this study tenured teacher is a teacher who has been teaching more than 3 years.

Overview of Chapters

Chapter one will present the introduction, historical roots, purpose of teacher evaluation, before Race to the Top, teacher training for implementation, the background of the study, research problem, purpose statement, statement of intent, conceptual framework, organizational theory, schools within organizational framework, teacher evaluation through an organizational lens, research questions, significance of the study, review of the research method used, methodological assumptions, limitations and delimitations, summary, operational definitions, and overview of proposed chapters.

Chapter two provides a comprehensive review of literature related to the concepts, issues, and key scholarly findings for the study. It includes an in-depth

overview from general to specific the topic of teacher evaluation, why do we have teacher evaluation, capturing teacher effectiveness, policies, Oklahoma data collection, continuous changing policy landscape: Oklahoma TLE system, teacher training for implementation, and administrator and teacher perceptions of teacher evaluation systems.

Chapter three includes the study's research questions, sample, research design, pilot study, instrumentation, data collection, data analysis, methodology assumption, limitations and delimitations, and summary.

Chapter four presents the results narratively using tables and graphs.

Chapter five summarizes the findings, draw conclusions, and makes recommendations for practice and future research.

The study concludes with references and appendices presenting the questionnaire used.

CHAPTER II

Literature Review

Introduction

This review of literature provides a background of current research and practices as they have evolved within the field of teacher evaluation. This chapter addresses eight topics including, (a) an overview of the history of teacher evaluation, (b) why do we have teacher evaluation, (c) capturing teacher effectiveness (d) policies, (e) Oklahoma data collection, (f) the continuously changing policy landscape: Oklahoma TLE system, (g) teacher training for implementation, and (h) administrator and teacher perceptions of the teacher evaluation systems.

On July 2nd, 2009a, Secretary of Education, Arne Duncan, addresses the National Educational Association in a speech called *Partners in Reform*:

> We created tenure rules to make sure that a struggling teacher gets a fair opportunity to improve, and that's a good goal. But when an ineffective teacher gets a chance to improve and doesn't- and when the tenure system keeps that teacher in the classroom anyway- then the system is protecting jobs rather than children. That is not a good thing. We need to work together to change that. (U.S. Department of Education, p. 4)

The previous words were the beginning of many conversations that followed the release of the new national reform to improve education. How do we, as an educational organization, fix a broken system known as teacher evaluation?

History of Teacher Evaluation

"Evaluating teachers has been an educational activity since Socrates," declares Danielson and McGreal (2000) when addressing the history of teacher evaluation. As teacher evaluation has advanced, the relationship between teachers and administrators has gone back and forth on the spectrum between inspector on one end and supporter on

the other. How has teacher evaluation transformed? How has the relationship between teachers and administrators changed? Answering these questions provides understanding and validates the perceived dissonance between administers and teachers today. Trained professionals engaged in evaluating teachers did not formally occur until after the Civil War (Danielson & McGreal, 2000; Nolan & Hoover, 2005). Nolan and Hoover recognized that evaluators, or as they were known at the time, inspectors, were often ministers, selectmen, schoolmasters, and other distinguished citizens in the early 18th century. Their supervision emphasized strict control and close inspection of school facilities and continued throughout the 19th century until inspectors become encumbered with multiple task and new reforms led to the creation of the superintendent position (2005). The relationship between administrators and teachers was nonexistent as most teachers were disenfranchised females state Bolin & Panaritis, who describe female teachers of the time as a "bedraggled troop, incompetent and backward in outlook" (as cited in Nolan & Hoover, 2005, p. 22).

In the early 20th century, the development of multiple rating scales designed to rate teacher effectiveness emerged as focus shifted towards efficiency and effectiveness (Danielson & McGreal, 2000; Nolan & Hoover, 2005). Edward Elliot and Clifton Boyce were two of the pioneers of teacher evaluation during this time. In the 1940s, Alexander recognized teacher evaluation as a system used to critique future teacher as a protocol to determine readiness to enter a school of pedagogy (1957). Danielson and McGreal found that in the 1940s and 1950s, emphasis was on presage variables (2000), meaning teacher were evaluated on their teacher traits to determine effectiveness. During the 1950s, the focus shifted again when superintendents wanted to develop an

evaluation system to measure teachers to determine a salary scale aligned to effectiveness (Alexander, 1957). Evaluation was a tool to assist teachers to improve their practice in the 1960's (Anderson, 1969; Danielson and McGreal, 2000; Nolan & Hoover, 2005) and in the 1970's, it was used to determine whether the objectives of education were achieved and to identify effective and ineffective teachers to give contributions to improve education, and to provide motivation and self-improvement (Beller, 1971; Danielson and McGreal, 2000). The 1970's also gave rise to Madeline Hunter's model of clinical supervision, which was a process of teacher evaluation, but districts across the United States changed it to a teacher rating checklist, causing much confusion between teacher supervision and evaluation among teachers (Nolan & Hoover, 2005). Teacher evaluations were defined as a measure of teacher competence in a standardized fashion in the 1980's (Darling- Hammond, Wise, & Pease, 1983). Many alternative models of teacher supervision were developed in the 1980's to try and counteract the impact of Hunter's model, such as developmental supervision and reflective supervision to "espouse collaborative efforts to improve teaching and insisting that evaluation and supervision be viewed as separate activities" (Nolan & Hoover, 2005, pg. 24). These models are still used today.

Relationship Between the Administrator and Teachers

The relationship between administrators and teachers during these historical times was often confusing as administrators' roles varied across a spectrum from inspector to supporter. The role of the supervisor was easily defined as teacher evaluator as the use of rating scales grew in popularity. The supervisor's role involved helping teachers to grow and being collegial. In the late 1950's, as rating scales came

under scrutiny and Sputnik launched, the focus for supervisors became more about being an instrument of change. Collaboration became less important and supervisors monitored implementation and punished resisters with lower evaluation ratings. In the late 1960's, research establish "best practice" to include developing a trusting relationships with teachers. During this time, supervisors found themselves caught between teacher evaluation and teacher improvement, resulting in declining teacher trust. Confusion continued to grow as multiple models were developed to try and measure teacher effectiveness, yet each model's focus was on different components ranging from clinical supervision to reflective supervision. Overall, Nolan and Hoover note, "teacher evaluation as an important function of supervisors, which causes many school systems to be predominately inspectoral in evaluation" (2005, p.25).

From the 1990's to present, prominent scholars still advance the same central notion for the definition of teacher evaluation. Teacher evaluation is currently defined as a "function designed to make comprehensive judgments concerning teacher performance and competence for the purpose of personnel decisions such as tenure and continuing employment" (Nolan & Hoover, 2005, p. 26). Danielson states that teacher evaluations are used to measure teachers' performances and promote professional development (2010). Teacher evaluations are often designed to serve two purposes: to measure teacher competence and foster professional development and growth (Weems and Rogers, 2010). Teacher evaluations assess individual teaching performances in the classroom, the school context, and student outcomes (Looney, 2011). They seek to improve student learning systems systematically and to promote continued teacher development (Papay, 2012). The common element of each definition includes: teacher

evaluation is a tool used to evaluate multiple measures to facilitate student success. The most common measures are: (a) promoting growth of teachers, (b) removal of ineffective teachers, and (c) improvement of professional development.

Why Do We Have Teacher Evaluation?

Teachers question the purpose of teacher evaluation each year as they prepare for principals to enter their classrooms for annual observations. Danielson claims the most fundamental reason why teacher evaluation exist is, "Because public schools are public institutions; they take public money, and the public has a right to expect high-quality teaching" (2007, p.36). As such, teacher evaluation is currently the primary tool being promoted to improve teaching (Darling-Hammond, 2012) and determine a teacher's level of quality. Traditionally, using teacher evaluations was a long-term method for assessing teacher effectiveness, which typically included working within a one page checklist completed by administrators bi-annually or annually (Derrington, 2011; Papay, 2012). Most school districts have a formal procedure for teacher evaluation by order of state law or regulation (Danielson & McGreal, 2000; OECD, 2005).

Researchers identified many studies that parallel how a well-designed teacher evaluation system, aligned with professional learning and development, can contribute to improvements in the quality of teaching and raise student achievement (Harris & Sass, 2009; Looney, 2011; Marzano, 2012; Papay, 2012; Weems and Roger's, 2010). Conversely, there are many researchers whose findings are inconclusive about the significance of teacher influence on student learning (Darling- Hammond, Amrein-Beardsley, Haerel, & Rothstein, 2012; Darling-Hammond, Wise, & Pease, 1983;

OECD, 2005). As stated earlier, the most common measures for teacher evaluation systems are to: (a) Promote growth of teachers, (b) Remove ineffective teachers, and (c) Improve professional development. Despite scholarly disagreement about teacher evaluation efforts on student learning and teaching quality, these three areas of teacher evaluation merit greater attention.

Promote Growth of Teachers

Information gathered in evaluation processes is used to identify teachers' strengths and weaknesses with opportunities of recognizing where they have grown. Promoting the growth of teachers validates that what they do is important and sends a message that their work is significant to the teaching and learning process (Donaldson, 2009; OECD, 2005). Further, many researchers maintain that teacher quality is the most important school-level factor affecting student achievement (Derrington, 2011; Donaldson, 2009; Kupermintz, 2003; Looney, 2011; Menuey, 2005). Therefore, maintaining high standards of teacher quality is vital. Danielson and McGreal concur, "To ensure teaching quality, schools and districts must base the evaluative criteria on recent research on teaching and learning" (2000, p.22). It is important to have an objective tool that represents a teacher's abilities that reflect mastery or weaknesses to measure the qualities. In addition, Weems and Rogers advise that evaluations should be implemented to ensure that teachers are effectively teaching and helping other educators to improve in the areas of weakness (2010). Yet when teachers are evaluated, an administrator usually visits the classroom for an hour, completes a scale of questions, and then has teachers sign to verify that it was completed either satisfactorily or unsatisfactorily (Derrington, 2011; Papay, 2012). This does not allow for collaboration

between teachers and administrators; hence, in this evaluation process, actionable feedback is negligible (Maslow & Kelly, 2012; McGee, 2013; Peterson & Comeaux, 1990). This practice does not meet the expectations of productive teacher evaluations. Instead, a teacher evaluation system should give teachers useful feedback on classroom needs, the opportunity to learn new teaching techniques, and counsel from principals and other teachers on how to make changes in their classrooms (Looney, 2011). To achieve these goals, new trends discussed later are developing in teacher evaluations.

Weems and Roger's (2010) address the demands of maintaining and keeping quality teachers. They state that teacher evaluations are a direct link to teacher growth; therefore, they can be used administratively as incentives and enhancements. Basically, evaluations give teachers the opportunity to receive resources necessary for professional growth in areas of weakness. For incentive and enhancements, evaluations offers teachers the benefit to receive incentive pay for reaching certain masteries and/or the opportunity to become National Board Certified, one of the highest symbols of excellence in teaching.

Removal of Ineffective Teachers

Incompetent teachers have an impact on student learning and can decrease student achievement (Menuey, 2005). Teacher evaluations can be used as a tool to remove ineffective teachers and the new Race to The Top reform requires teacher evaluations to do so. OECD found that two-thirds of countries surveyed report they have evaluations as a tool to remove ineffective teachers and that teachers can be dismissed because of chronic underperformance, but public school teachers are rarely dismissed on grounds of performance (2005). Removing ineffective teachers from the

classroom is not as easy as it sounds. Poor-performing teachers present one of the toughest challenges school principals may ever face," suggests Yariv (2009). For tenured teachers, procedures have to be followed by administrators and the school board must give proper notice of intent by legal precedent as was found in *Roth v. Board* of Regents (1972). These procedures can be cumbersome and time consuming, which may be a possible reason for principals not following through on removing ineffective teachers using evaluations systems. This is not the case if a teacher is non-tenured; they may be non-renewed without reason by an employer.

Another possible explanation for minimal numbers of teacher contract non-renewals is that ineffective teachers are enabled by principals who avoid writing honest performance appraisals (Menuey, 2005; Nixon, Packard, and Douvanis, 2010; OCED, 2005; Waintroob, 1995). Eventually non-tenured teachers become tenured and the procedure of removal becomes more difficult when the principal repeatedly submits positive appraisals for ineffective teachers. Waintroob affirms that tenured teacher's dismissal is typically due to ineffectiveness, involving teachers who have been in the system for years and receive inaccurate, "satisfactory" ratings, making administrators their own worst enemy (1995). Administrators are going to have to lean on their professional ethics and not only use evaluations to identify strengths, but to confront unsatisfactory performance with the possibility of discharge for the concern of the students (Waintroob, 1995).

Jacob (2011) found evidence that principals do consider teacher productivity in determining which teachers to dismiss. The dilemma is that evaluations are so subjective and do not show all the strengths and weaknesses of a teacher. In Jacob's

2004 study in Chicago Public schools, his analysis shows that 38.8% to 46.2% of elementary principals and 28% to 34% of secondary principals — including those in some of the worst performing schools in the district — did not dismiss any teachers despite how easy it was to do under its new policy. Further, in that year half of the dismissed teachers were rehired the following year by another school in the district (Jacob, 2011). These results should raise concern and reconfirm the reason for making evaluation a tool for documentation that is transparent and entails follow-through from year to year.

Dismissing teachers should be the last resort option. When administrators become aware of teacher weaknesses that are not being remediated, teachers and administrators should collaborate on a professional growth plan. Weems and Roger's discuss professional growth plans that are based on professional standards and the individual needs of a teacher. Basically, they give teachers the opportunity to receive professional development and resources to support growth in areas of weakness (2010). Giving teachers the resources and training that they need to improve should always be an option within a growth plan if the goal is to have and maintain effective teachers. After a growth plan has been implemented and a teacher continues to represent themselves with a rating of "unsatisfactory" on an evaluation system, removal of the teacher should always be the next option. The courts will uphold the dismissal of the teacher if the administrator has engaged in the labor intensive process of conducting their evaluations properly, documenting them well, and adhering to deadline with proving incompetence (Range, Duncan, Scherz & Gaines, 2012).

Improvement of Professional Development

"As a professional development tool, evaluation can prove useful in helping build organizational capacity," assert Papay (2012, p. 134). Teacher evaluation systems can improve professional development by identifying common thematic areas where teachers represent weaknesses. The areas, in turn, can be embedded and presented within a district or local school's professional development program to work on improving weaknesses. Papay states more specifically that it can help principals identify areas of instructional strength and weakness and target resources appropriately (2012). If teacher evaluation is used appropriately to promote continued teacher development, the output would be to improve student learning (Papay, 2012).

Yet, using a large-scale international survey on teacher evaluation, the OECD found, that professional development is often fragmented, unrelated to teaching practice, and lacks intensity and follow-up (2005). Often times, professional development days merely fulfill district requirements and provide professional development credits to help teachers maintain an eligible rehiring status. Even though it creates a great opportunity for the majority, teachers who need the professional development or those on improvement plans are often overlooked. The OECD states the disjointedness is mostly due to the fact that professional development can be used for multiple services, e.g., curriculum, evaluation indicators, or individual/district school trainings (2005).

Another factor affecting professional development is the lack of connectedness between teacher evaluation and professional development creating a deficiency of information received from the evaluation tool (Rowan, 1990). McGreal (1983)

identified, “Even though districts state that they have evaluations to improve instruction, they choose systems with high supervisor, low teacher involvement” (p.8-9). In the past, this trend may have been due to a lack of other evaluation options, today it is also due to lack of administrative time (McGreal, 1983; OCED, 2005). The OCED found the evaluation tools are rarely linked to specific needs targeted through professional development. The study reveals administrators do not spend enough time in the classrooms, so many teachers feel they do not receive helpful feedback creating feelings of isolation and invisibility (2005). Donaldson concurs, by showing how a small number of teachers in Bernardino, CA, in 2002-03 stated that they received feedback on their evaluation and that evaluation were useful and effective (2009). This may be because teachers are passive in the traditional evaluation process and administrators are doing all the work (Clipa, 2011; Danielson, 2010; Donaldson, 2009; Nolan and Hoover, 2005). Clipa found most teachers see the evaluation process as dependent (2011). With the traditional evaluation system, Danielson says, “This process violates everything we know about learning- that learning is done by the *learner* through a process of active intellectual engagement” (2010). These results lend themselves to further study towards improvement with teacher evaluation systems and opportunities of growth through professional development systems.

Capturing Teacher Effectiveness

Teacher effectiveness is subjective and can be defined in multiple ways. The U.S. Department of Education defines effective teachers as, “Those teacher whose students achieve acceptable rates (e.g., at least one grade level in an academic year) of student growth (as defined in this notice)” (2009b, p. 12) and further establish how it

will be measured. Researcher Kupermintz defines it as, "Differences in student learning determines- by definition- teacher effectiveness: a teacher whose students achieve larger gains is the "effective teachers," (2003, p. 289). Goe, Bell, and Little from the National Comprehensive Center for Teacher Quality profoundly recognize a five-point definition:

> Effective teachers (a) have high expectations for all students and help students learn, as measure by value-added or other test-based growth measures, or by alternative measures, (b) contribute to positive academic attitudinal, and social outcomes for students such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior, (c) use diverse resources to plan and structure engaging learning opportunities; monitor student progress formatively, adapting instruction as needed; and evaluate learning using multiple sources of evidence, (d) Contribute to the development of classrooms and schools that value diversity and civic-mindedness, and (e) collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of the students with special needs and those at high risk for failure. (2008, p.8)

A teacher who is successful with the before mentioned task is considered an effective teacher. The teacher evaluation frameworks can capture the measures of teacher effectiveness.

Scholars debate the notion that teacher effectiveness predicts student learning. Regardless, teachers should always seek ways to improve their practices, which in turn, should affect student learning. Teacher evaluation tools are a catalyst to identifying areas for improvement. If the objective is to develop teacher evaluation systems that teachers find meaningful and from which teachers can learn, "we must use processes that not only are rigorous, valid, and reliable, but engage teachers in those activities that promote learning" confirms Danielson (2010).

Some researchers elect to not look at the objectives of quality teaching, but the personality traits of a teacher. When examining teacher qualities to determine teacher effectiveness, Harris and Sass attempt to identify the determinants of teacher productivity to predict quality teachers. They contend that teacher productivity is the most important component of a school's effect on student learning. Therefore, one should attempt to evaluate teachers based on a combination of subjective assessment and student outcomes to more accurately gauge teacher performance and predict future teacher value added (2009). Harris & Sass could not find one factor or set of factors in their data set that had a strong positive relationship to a quality teacher. They advocate the best way to determine what makes a good teacher is to locate long-term patterns by using multiple sources rather than a one-day observational snap shot. Darlington-Hammond, Wise, and Pease (1983) found little evidence supporting single teaching performance variables as essential for effective teaching, concluding that there is not one factor that captures teacher effectiveness. There are numerous ways to collect data regarding teacher qualities; one may not be better than another, but all serve to improve a school system.

The Different Types of Evaluation Sources

Data collection for determining the effectiveness of a teacher can be gathered through multiple sources. Some of the most common sources are through classroom observations, interviews, portfolios, checklist, and value-added systems (Darling-Hammond, Amrein-Beardsley, Haerel, & Rothstein, 2012; Derrington, 2011; Donaldson, 2009; Goldrick, 2002; OECD, 2009; Stecher, Garet, Holtzman & Hamilton, 2012; Stufflebeam, 2001). More controversial evaluation sources include peer, student,

and parent surveys; multiple shortened unannounced walk-throughs; Other Academic Measures; and self-evaluations (Danielson & McGreal, 2000; Darling-Hamilton, Wise, & Pease, 1983; Marshall, 2013). Even though all sources can provide benefit toward teachers' insights into his or her practice, they should be used cautiously when used alone. Derrington analyzes teacher evaluations by addressing the principal's role as moving from the traditional task of 60-minute walk- through and checklists to more teacher centered and multi- faceted (2011). Multiple researchers state the need for future evaluations to include more than one source to develop a more holistic approach of a teacher's effectiveness (Derrington, 2011; Goldrick, 2002; Kyriakides & Demetriou, 2007; Looney, 2011). Most of the sources that administrators require or teachers use to determine effectiveness are a part of summative evaluations, but can be used as formative evaluation measures.

Summative Versus Formative Evaluation

The general purpose for evaluation is accountability and improvement, which can be facilitated by summative and formative evaluation. Accountability reflects the need to determine the effectiveness of teachers in order to ensure that services delivered are efficient and effective towards increases in student success. Summative evaluation supports this because it is a final assessment. Formative evaluation supports this because it is continuous and provides regular feedback. Improvement reflects the need for professional growth and development of the individual teacher. (Danielson and McGreal, 2000; Kyriakides & Demetriou, 2007; OECD, 2009; Popham, 1988; Scriven, 1996). Both evaluation results merge together in assessing the effectiveness of a teacher. Even though each evaluation is conducted separately, they are not exclusive of

each other; both summative and formative components are needed to help establish the level of teacher effectiveness (Scriven, 1996). In addition, there are several types of summative and formative evaluations that can align together depending on a district's mission.

A district's focus could determine which evaluation tools a district selects. The OECD recognizes that summative results can be used to determine career advancement; implement performance rewards or establish sanctions for underperforming teachers. Formative results play a key role in ensuring the improvement of teachers by identifying the needs for professional development and making the resources available (2009). Kyriakides and Demetriou's study alleged that most teachers understand the difference between the two and both should be involved in the design and selection of the evaluation system attempting to measure teacher effectiveness (2007). "The problem is that teacher evaluation systems have not accurately measured teacher quality because they have failed to do a good job of discriminating between effective and ineffective teachers," states Marzano (2012). So, what is essential in an effective evaluation framework?

Best Practices

School districts throughout the United States have developed evaluation systems that reflect the best of what is known (Danielson & McGreal, 2000), but what are the best practices and can a teacher evaluation system capture them all? Teachers consistently state the purpose of evaluation should be to promote teachers' reflection on practice, consider the content, and be tailored to different teachers' needs (Peterson & Comeaux, 1990). Yet, teacher perception is rarely accounted for when considering

revisions for teacher evaluation. There are many solutions to improve the measure of teacher effectiveness.

Goldrick (2002) specifies policy-makers can transform teacher evaluation into a more effective tool for improving instructional practice and raising student achievement by aligning evaluation with academic standards for students and professional standards for educators. Goldrick states that policymakers should consider taking action to define teaching quality, focus evaluation policy on improving teaching practices, incorporate student learning into evaluation; create professional accountability, train evaluators, and broaden participation in evaluation design (2002). Historically, evaluations have lacked focus and a clear purpose of what teachers should be expected to be doing, therefore, this lack of clarity implies lack of reliability. In addition, measuring teacher effectiveness based on student achievement also lacks reliability due to complexity of capturing teaching skills. Goldrick (2002) declares that despite the challenge, strengthening teacher evaluation is still worthwhile if investment in educators includes providing greater information, confidence, and the ability to improve teacher's instruction practices to help students achieve their fullest potential.

After observing some "model" districts, Danielson and McGreal identified four best practices that should be recognized when beginning to implement new evaluation framework: (1) new evaluation systems should be directly linked to the mission of the school district, (2) new evaluation and professional development systems should be viewed as a continuing process, (3) They should emphasize student outcomes, and (4) must be a commitment to allocating adequate resources to allow new systems to be successful (2000). Danielson and McGreal feel these practices will create a district

culture where there is collaboration and supports for all involved giving opportunity for districts to break away from old practices and developing new ones (2000).

Is the System Irreparable or in Need of Repair?

Is the teacher evaluation system broken or are evaluation frameworks not being actively used for their intended purpose? The literature base establishes that teacher evaluation is for teacher growth, removal of ineffective teachers, and improvement of professional development, yet most use them as a reflection of student achievement. Maybe it is time for the definition to be modified, broadened, or changed all together. Danielson and McGreal state that the, "Shortcomings of teacher evaluations are that they are (1) outdated, (3) have limited evaluative criteria, (2) lack of shared values and, (4) have assumptions about what constitutes good teaching" (2000, p. 11). There are many reasons why something does not work, nonetheless, Darling-Hammond et al. remind us that there is no quick fix (1983). Since 1983, researchers have been trying to address it and have come to realize that, "The public has come to believe that the key to educational improvement lies in upgrading the quality of teachers rather than in changing school structure or curriculum (Darling-Hammond et.al., 1983)," and it seems even today, that the "dog is still chasing its tail" (McGreal, 1983).

Wallace concurs, stating:

> On a macro level, lack of trust for the teaching profession has created an opening for businesses and policy makers to take control of teacher evaluation. External stakeholders trying to implement a business model assuming that teaching and learning can be broken down into easily measurable units, but teaching and learning are incredibly complex and hard to measure. (2012, p. 45)

Students deserve effective teachers who can develop students into productive citizens in society. The issue is measuring effectiveness with an evaluation framework. Checklist

forms do not constitute a system. An effective teacher evaluation system is far more multifaceted than checklist (Danielson & McGreal, 2000).

Policies

National Policy

American Recovery and Reinvestment Act.

On February 17, 2009, President Obama signed into law the American Recovery and Reinvestment Act (ARRA) to improve education (USDOE, 2009). Within the ARRA, $4.35 billion dollars was provided to a fund called Race to The Top (RTTT) (USDOE, 2009). This fund is a grant program that is rewarded to states that meet specific criteria under four educational reforms. McGuinn says:

> RTTT's program has articulated a promising new approach to federal education policy in the competitive grant program, and it has generated a substantial amount to state policy change in a short period of time, particularly for a program of its relatively small size. Perhaps, most important, it has had a sizable impact on the intensity and character of school reform discourse across the country. (2011, p.141).

In the Race to the Top Program Executive Summary, teacher evaluation is under criteria D, Improving teacher and principal effectiveness based on performance. All of criteria D will be addressed in the literature, but focus will be given to criteria (D) (2). Under this reform (p.9), the criteria states the extent to which the State in collaboration with its participating LEAs, has a high-quality plan and ambitious yet achievable annual targets. Criteria D's targets are to:

> (i) Establish clear approaches to measuring student growth and measure it for each individual student;
>
> (ii) Design and implement rigorous, transparent, and fair evaluation systems for teachers and principals that (a) differentiate effectiveness using multiple rating

categories that take into account data on student growth (as defined in this notice) as a significant factor, and (b) are designed and developed with teacher and principal involvement;

(iii) Conduct annual evaluations of teachers and principals that include timely and constructive feedback; as part of such evaluations, provide teachers and principals with data on student growth for their students, classes, and schools; and

(iv) Use these evaluations, at a minimum, to inform decisions regarding—

(a) Developing teachers and principals, including by providing relevant coaching, induction support, and/or professional development;

(b) Compensating, promoting, and retaining teachers and principals, including by providing opportunities for highly effective teachers and principals (both as defined in this notice) to obtain additional compensation and be given additional responsibilities;

(c) Whether to grant tenure and/or full certification (where applicable) to teachers and principals using rigorous standards and streamlined, transparent, and fair procedures; and

(d) Removing ineffective tenured and untenured teachers and principals after they have had ample opportunities to improve, and ensuring that such decisions are made using rigorous standards and streamlined, transparent, and fair procedures.

These criteria trigger an opportunity for much needed conversations about how existing teacher evaluation, tenure, and dismissal policies are broken and imped efforts

to improve teacher quality and student achievement. It has prompted an unprecedented wave of state reforms with hope of educational improvement (McGuinn, 2011).

Oklahoma State Policy

Oklahoma was already implementing criteria for teacher evaluation prior to the American Recovery and Reinvestment Act. Districts across the state were consistent in their implementation of the Oklahoman Criteria for Effective Teaching and Administrative Performance. This policy was enacted on January 9, 1987 and was agreed upon by the Attorney General Opinion No. 86-146:

> It is, therefore, the official opinion of the Attorney General that all evaluation policies adopted by Oklahoma school districts be based upon minimum criteria developed by the State Board of Education; that in those school districts with previously existing professional negotiation agreements, the negotiated provisions must comply with the State Board of Education minimum criteria; that the provisions of the evaluation procedure are mandatory topics of professional negotiations; and that the criteria negotiated and adopted may exceed the minimum criteria promulgated by the State Board of Education pursuance to 70 O.S. § 6-101.10. (OSDOE, 1999)

As such, each school's administrative board, in collaboration with their negotiation unions, have the responsibility of implementing, maintaining, and ensuring that the state evaluation policy is being abided by teachers and administrators. In addition, the policy is to be reviewed annually to maintain alignment with the state policy as nothing in the act should be construed or modified (OSDOE, 1999). It continues to state that every district policy so adopted of teacher evaluation in Oklahoma shall:

1. Be based upon a set of minimum criteria developed by the State Board of Education;

2. Be prescribed in writing at the time of adoption and at all times when amendments thereto are adopted. The original policy and all amendments to the policy shall be promptly made available to all persons subject to the policy;

3. Provide that all evaluations be made in writing and that evaluation documents and responses thereto be maintained in a personnel file for each evaluated person;

4. Provide that every probationary teacher be evaluated at least two times per school year, once prior to November 15 and once prior to February 10 of each year;

5. Provide that every teacher be evaluated once every year, except as otherwise provided by law;

6. Provide that, except for superintendents of independent and elementary school districts and superintendents of area school districts, who shall be evaluated by the local school board, all certified personnel, including administrators, shall be evaluated by certified administrative personnel designated by the local school board;

7. All personnel designated by the local board to conduct the personnel evaluations shall be required to participate in training conducted by the State Department of Education prior to conducting such evaluations;

8. The State Department of Education shall develop and conduct workshops pursuant to statewide criteria which train such administrative personnel in conducting evaluations;

9. The State Board of Education shall monitor compliance with the provisions of this section by local school districts; and

10. Refusal by a local school district to comply with provisions of this section shall be grounds for withholding State Aid funds until such compliance is met. (70 O.S. § 6-101.10).

Oklahoma Criteria for Effective Teaching Framework.

The Oklahoma Criteria for Effective Teaching is highly succinct in its objectives. The outline of its framework can be seen in the appendix (see Appendix E). Eventually, this framework was developed into a Likert scale, where an administrator

rated teachers on a spectrum between effective and non-effective. Tenure teachers were evaluated once a year while non-tenure teachers were evaluated twice a year.

The Tool used to Evaluate Oklahoma Teachers.

The tool used to evaluate teachers to meet the objectives of the Criteria for Effective Teaching policy was called the Teacher Appraisal System (TAS) (Thomas, 2005). TAS was a one sheet triplicate form specially designed to meet the Criteria for Effective Teaching. TAS used a Likert scale that made it easy for administrators to determine teacher effectiveness. Teachers were evaluated using TAS twice if teachers were non-tenured and once if teachers were tenured. The TAS program was known as non-time consuming and providing immediate results from an administrators' perception (L. Johnson, personal communication, August 13, 2014; Thomas, 2005).

2010 School Laws of Oklahoma Chapter 1- Oklahoma School Code Article V: School Districts and Boards of Education

Under Section 118 (70 O.S. § 6-101.10), Evaluation of Teacher and Administrators section A-1, the requirement of teacher evaluations is established (OLOSDE, 2013). It states:

> Every policy of evaluation adopted by a board of education shall be based upon a set of minimum criteria developed by the State Board of Education, which by no later than the 2013-2014 school year, shall be revised and based upon the Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE) developed by the State Board of Education as provided in Section 6 of this act. (OLOSDE, 2013)

Teacher and Leadership Evaluation System (TLE) Program.

The Oklahoma State Department website contains several links to the new options and changes in Oklahoma Public Schools' teacher evaluation system. Recent political changes have increased the awareness of teacher accountability leading to

reform of teacher evaluation. The Teacher and Leader Effectiveness Evaluation (TLE) System, also known as SB 2033 and Race to the Top, was developed by the State board of Education and adopted (2012). The TLE Program has three components, one qualitative measure and two quantitative measures. The first measure of implementation was the qualitative measure, which involves a teacher evaluation framework. On the Oklahoma State Department website under Oklahoma Teacher Evaluation and Marzano Teacher Evaluation, it maintains how local boards of education had to adopt an evaluation system that contains the new TLE minimum criteria to be implemented no later than the 2013-2014 school year (2012).

At a meeting on December 15, 2011, the state school board voted to allow individual districts to freely choose from Tulsa Public Schools' Evaluation System, Marzano Casual Teacher Evaluation Model, or the Danielson's Framework for Teaching (Edger, 2012). The three choices were selected by the Teacher and Leader Effectiveness Commission after meeting the expectations of the Oklahoma selection criteria framework and gaining the highest results on an online survey given to teachers statewide. The model selection criteria for each framework can be compared in Appendix C. On the Oklahoma State Department website under Oklahoma Teacher Evaluation and Marzano Teacher Evaluation, it articulates that the new evaluation system for teachers is designed to encourage continuous professional growth leading toward improved student achievement for all Oklahoma students (2012).

The new evaluation system includes qualitative and quantitative measures (OSDoE, 2012). The qualitative measure is the preselected evaluation framework a district indicates and the quantitative is the value added measures and other academic

measures. Further, the website expresses that the qualitative portion will weight 50% towards an effectiveness score and the qualitative will weigh the other 50% (35% going towards value added measures and 15% for other academic measures) equaling 100% of a teacher's TLE score (OSDoE, 2012).

Qualitative Measures: The Evaluation Systems Chosen

Districts piloted both the Marzano Causal Teacher Evaluation Model and the Tulsa Teacher and Leader Effectiveness frameworks throughout the 2012-2013 school year. The Danielson Evaluation Framework for Teachers, although approved by the State Board of Education, was not piloted due to the lack of schools choosing the model. Nonetheless, it will remains on the adopted approved frameworks list as school districts may decide to select the Danielson Evaluation Framework in the future. Currently, all districts have moved from the piloting phases of their chosen evaluation tools to full implementation in accordance with SB 426 (McGee, 2013).

Tulsa TLE Framework

Tulsa public schools (a district in the state of Oklahoma) developed an evaluation system called the Teacher and Leader Effectiveness (TLE) Observation and Evaluation System in 2010 (Tulsa Public Schools [TPS], 2013). It is designed to help measure and support teacher effectiveness and incorporates current research and best practices—with authorship and input from Tulsa teachers and administrators. The basic evaluation framework includes observations, conferencing, evaluations, supports, and mentoring. The system is based around 20 criteria that flows into five domains that are given a score during formal evaluation. The five domains are: (a) classroom management (30%); (b) instructional effectiveness (50%); (c) professional growth

(10%); (d) interpersonal skills (5%); (e) leadership (5%). The domains carry different weights concerning how they impact student achievement. They are then averaged into one score to determine a teacher's level of effectiveness (TPS, 2013). Specific details about the system can be observed and manipulated on the district's website, which can be found in the appendix (see Appendix C).

Marzano Framework

As stated on the Marzano website, Marzano has been developing and using the Marzano Casual Teacher Evaluation Model for over five decades. His charge is to connect teacher growth to student achievement. The model contains 60 elements that define a knowledge base for teaching and a framework for the systematic development of expertise. The elements are measured within four domains which are: (a) Classroom strategies and behaviors, (b) Planning and Preparing, (c) Reflecting on Teaching, and (d) Congeniality and Professionalism. These domains lend themselves to an increase in student achievement. Specific details about the Model can be found at on the website. The link can be found in the appendix (see Appendix C).

Quantitative Measures

The Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE) Report indicates that fifty percent of a teacher's total evaluation score will be based on quantitative measures (2011). Thirty-five percent will address student academic growth measures by either value added measures (VAM) or student learning objective/student outcome objective (SLO/SOO) and fifteen percent will be Other Academic Measures (OAM) (OSDOE, 2014). Implementation of Student Growth Measures began in 2013-2014 and full implementation of the quantitative measure is expected by the 2015-2016

school year (McGree, 2013). This did not happen quite as expected and is addressed under the section discussing continuous changing of policy landscape beginning on page 80.

Value Added Measures

Teachers that will participate in value added measures are fourth through eighth grade Language Arts and Math teachers in addition to, Algebra I, Algebra II, Geometry, and English III educators. Only the identified teachers will have thirty-five percentage points based on student academic growth using multiple years of standardized test data (OSDoE, 2012). Darlings-Hammond, Amrein- Beardsley, Haerel, & Rothstein state that, "VAMs are a tool to enable officials to use statistical methods to measure changes in student scores over time, while considering student characteristics and other factors often found to influence achievement" (2013, p. 8). Therefore, VAMs address gains in students' academic achievement as part of the compensation system (Rothman, 2010). VAMs advantage is they reflect student proficiency levels while recognizing each student's growth by accounting for differences in student performance and abilities (REACH, 2012). Even with that, VAMs are a new concept, which some researchers feel have brought benefit (Goe, Bell, & Little 2008), but VAMs should be used with caution as researchers have also documented a number of problems (Darlings-Hammond, Amrein- Beardsley, Haerel, & Rothstein, 2013; Paige, 2012).

Student Learning Objective/ Student Outcome Objective

All other teachers who will not participate in VAM will participate in student learning objectives/ student outcome objectives to represent the thirty-five percentage points under academic growth measures. An SLO/SOO is a measurable, long-term,

academic goal informed by available data that a teacher or teacher team sets at the beginning of the year for all students or for subgroups of students (OSDOE, 2014). A teacher using an SLO has a concentration on academic growth relating to particular standards. A teacher focused on SOO has a concentration on outcomes such as librarians, counselors, and nurses (AIR & OSDE, 2014). As of 2014, SLO/SOO is a new model with little research, but Oklahoma has started the process of "training the trainers" and are in developmental implementation stages (AIR & OSDE, 2014).

A basic flow chart for development of an SLO aims at: (a) identifying core content and standards; (b) gather and analyze student data; (c) determine the focus of the SLO/SOO; (d) select or develop an assessment; and (e) develop a growth target (OSDOE, 2014). For implementation across Oklahoma school districts, data was collected during the 2014-15 school year to generate an SLO/SOO score for teachers based on the percentage of students meeting their targets. These scores will be used in totaling to teachers' overall TLE score for the 2015-16 school year (AIR & OSDE, 2014). The awareness and implementation of this phase was very short. The OSDOE was made aware of the component in March 2014, received funding in July 2014, and requested districts implement the following school year of 2014-2015 (AIR & OSDE, 2014). Districts across Oklahoma became over stimulated to efficiently develop and approve SLO/SOO policies and provide teachers with clear policies, training, and procedures required with the given time frame (OSDOE, 2015; SREB, 2014). This haste created an outcry from practitioners (SREEB, 2014), which is discussed further under Oklahoma Data Collection and Continuous Changing Policy Landscape: Oklahoma TLE System.

Other Academic Measures

Under TLE Quantitative Components in the TLE Handbook under other academic measures on the Oklahoma State Department of Education website, it reveals how the Oklahoma State Board of Education adopted policies recommended to them by the TLE Commission concerning Other Academic Measures (OAMs) (OSDOE, 2014). It further proclaims that OAMs are additional alternative instruments ensuring a robust evaluation by: (1) capturing unique facets of effective teaching, (2) reflecting student academic performance impacted by the teacher, and (3) assessments and/ or programs that are specific to teachers' job assignments (OSDOE, 2014). Examples of OAMs that have been adopted by the State of Oklahoma are State Assessments, VAMs, "Off the Shelf" Assessments commonly used throughout the state, A-F Report Card Components, Surveys, and Student Competitions. Statute 70 O.S. § 6-101.10 and 101.16 required all districts to participate in a no-stakes pilot OAM collection during the 2013-2014 school year. In 2014- 2015, all teachers collected OAM data to be included as 15% of their final evaluation scores in 2015-2016 (OSDOE, 2014). Through each stage, the preliminary data were reported back to the Oklahoma State Department of Education to evaluate better ways to calculate data efficiently and accurately (McGree, 2013).

Oklahoma Data Collection

Three studies have collected TLE data since the inception of the Oklahoma Teacher and Leader Evaluation System. The first one, in 2013 by the TLE commission through surveys (McGree, 2013); second, in 2014 by the bargaining union, Oklahoma Education Association (OEA) through surveys (Littrell, 2014); and third, also in 2014

by the Southern Regional Education Board (SERB) through focus groups (SREB, 2014). Each investigation presented the TLE commission with results to improve and continue implementations, and offered areas for further study.

In February of 2013, McGree stated that the TLE office asked district superintendents to gain feedback from their leaders (2013). The TLE office's primary focus was to determine to what extent the new evaluation frameworks provide tangible feedback to teachers. In turn, they received 327 administrator responses. They found that ninety-three percent (93%) of respondents answered average to a tremendous extent; ninety-one percent (91%) indicated actionable feedback to leaders; nine-one percent (91%) stated the frameworks distinguished between effective teachers from average to a tremendous extent; and that eighty-nine percent (89%) distinguished effectiveness between leaders from average to tremendous amount (McGree, 2013). Then they developed a survey for the teachers.

The TLE office asked teachers to respond to a ten-question survey on a scale from 1-5, regarding the impact of the new evaluation frameworks on instructional practice in May of 2013 (McGree, 2013). Of the 5,500 teachers that responded, sixty-one percent (61%) of the teachers revealed that the new framework adopted by their district had improved professional dialogue somewhat to a great deal; Sixty-nine percent (69%) of educators stated that the new evaluation framework has provided actionable feedback from an average amount to a tremendous extent; eighty-six percent (86%) felt that they were informed of their evaluation framework based on training from their administrators; Fifty-seven percent (57%) of the teachers who responded indicated that their teaching practices had changed somewhat to a great deal due to the

new evaluation tools; And five percent (5%) gave negative comments towards the new evaluation system (McGree, 2013).

During the fall of 2014-2015, the OEA initiated a statewide survey study of all educators i.e. teachers, administrators, counselors, and "other" (Littrell, 2014). The focus of the survey was the TLE system and both its qualitative and quantitative portions. The survey was comprised of 18 Likert scale questions with an opportunity for participants to comment. Data was collected through email and Facebook; yielding 2,411 returns. Only one question concentrated on the teacher evaluation framework. The question assessed the overall feeling of the teacher evaluation system. The highest percent was 48.6% (1148 people) for somewhat positive and the lowest percent was 6.3% (149 people) with very negative. Three concerns were identified: The frameworks are still too subjective and results can vary by administrator; too labor intensive, redundant, and requires too much work for administrators; and both administrators and teachers need more training.

Related to this study, the survey addressed one of the purposes for teacher evaluation. It requested, as a whole, if the process improved instructional practice (Littrell, 2014). The highest percent was 46.3% (1,071 people) with no significant impact and the lowest, 5.7% (133 people) very positive and/or rewarding. The study did not address professional development or the dismissal of ineffective teachers (Littrell, 2014).

In the fall of 2014, the Oklahoma State Department of Education (OSDE) invited the SREB to conduct focus groups with educators across the state (SREB, 2014). There were 26 focus groups held in 10 cities. Participants included 71

administrators and 60 teachers representing 58 school districts. This opportunity was not by random selection. To participate in the focus group, a nomination by one's superintendent had to be submitted to the OSDE (SREB, 2014).

The results showed that selected educators felt the frameworks were an improvement for the qualitative portion, but more training was necessary before full implementation. Most discussions emphasized the quantitative portion of the survey with highly negative trends. The investigators found that as a system, educators wanted more guidance towards the quantitative portion; did not believe the evaluation system is valid, fair, or helpful for improving instruction; and that educators lack buy-in due to the distrust of the OSDE (SREB, 2014). The TLE commission recommended to the State Board of Education to postpone SLO/SOO implementation for 2014-2015 school year from the results of the SREB discussions (OSDE, 2015).

The TLE commission office expects continued growth as it expedites towards the full implementation stage of the TLE evaluation system and listed recommendations as well as received feedback to further promote professional development and improvements from each study (McGree, 2013, Littrell, 2014; SREB, 2014). At each level of implementation, surveys and feedback are expected to continue.

Gaps in the Research

The results of the surveys are interesting and indicate vast improvements occurred in the area of teacher evaluation in the state of Oklahoma, but through the literature I identified areas for further study:

- The TLE Commission's survey instrument contains a scale from one to five, but it does not state the ranges, leaving ambiguity on the levels of importance.

Further, the categories are inconsistently grouped in the findings, which evoke the speculation about the tools validity/reliability and the need to yield higher percentages.

- The OEA survey yielded 2,411 surveys. One of the highest subgroups was "other," which excluded an educator who works in the school. These results imply that external stakeholders could have influenced the study. Further, data analysis was represented as a whole: teachers' perceptions still were not recognized as a vital subgroup.
- The SREB participants were selected by district superintendents and included both teachers and administrators. Therefore, voices of study are not random and likely are not a true representation of how educators feel. In addition, the study focused more towards the quantitative portion than the qualitative portion, which is the concentration of this study.

The literature reinforces the purpose for teacher evaluation, yet for each investigation, the questions do not determine if the new evaluation frameworks are meeting its purpose from a teachers' perceptive. Do teachers have a positive perception towards the new system, which may create buy-in resulting in the new TLE system being a success? This facet is unclear. In addition, the surveys do not address the implementation phase, leaving options for further study.

Oklahoma's data results leaves a gap for further study in determining the effectiveness of the implementation of the new teacher evaluation system; aligning the purpose of teacher evaluation to teacher perception and determining teacher's level of

favorableness. This could lead to successful operative implementation of the new system.

The Continuously Changing Policy Landscape: Oklahoma TLE System

The past decade has demonstrated an increase in teacher accountability through continuous policy activity aimed at improving teacher quality (Gitomer, 2007). Therefore, this study, which focuses on teacher's perceptions' of teacher evaluation and the implementation of the Oklahoma TLE System is bracketed within this timeframe. It begins with the American Recovery and Reinvestment Act and Oklahoma's attainment of the RTTT grant (which is embedded within the act) and ends Fall 2015 of the academic year when I have captured teachers' perceptions through a questionnaire towards what has already been implemented through policy, training, and implementation in schools.

New policies and mandates continue to be established to fulfill the implementation of the TLE system with fidelity and utility. Most of the revamps have concentrated on the quantitative portion of the TLE system, which is not a focus of this study. When following the TLE commission on the OSDOE website under TLE, On July, 1st, 2015, Senate Bill 706 went into effect revamping many facets of the quantitative portion of TLE system (OSDOE, 2015). The one most prominent changes was the halt of the OAMs and SLO/SOO full implementation to address better methods for training and implementation based on the results from the SERB and pushing back full implementation from academic year 2015-2016 to academic year 2016-2017 (OSDE, 2015). Teachers are aware of these changes, as they have been communicated down to the district level. Yet, changes teachers may not be aware of, is if you look at

the OSDOE website under TLE August update through 2015school year, you will see that the TLE commission is planning to present the State Board soon with additional changes that will change the way the TLE system looks as a whole and will create another year of sets backs that will either be accepted or rejected February 1st, 2016 by the State Board (Thompson & Miller, 2015; OSDOE, 2016). All of the changes set this study up for continuous research and longitudinal data of teachers' perceptions of the system.

Teacher Training for Implementation

Richardson and Placier recognize that investigations to measure teachers' perceptions regarding school-level change have been unexplored (as cited by Teyten & Devos, 2009). There is a connection between teachers' perceptions and success or failure with teacher evaluation innovations (Teytens & Devos, 2009). As teachers cultivate self-awareness concerning their jobs, this development shapes their professional attitudes, including their reaction to innovations. As a result, teachers tend to assimilate policies to accommodate their own knowledge (Tuytens & Devos, 2009). An open line of communication among stakeholders is imperative not only to foster awareness of teacher perception, but to also collaborate and align intended teacher evaluation policy with teachers' understanding. Stakeholders taking into account teachers' perceptions when implementing teacher evaluation policy gives a greater possibility of success. Teachers' perceptions are expected to play a significant role for a top-down executed policy (Tuytens & Devos, 2009, p. 925).

True implementation of teacher evaluation starts in the classroom. As a result, it is imperative to address teacher training when designing teacher evaluation (Sartain,

Stoelinga, & Brown, 2011). There are two stages for successful implementation during teacher training: an awareness of the instrument innovation and its importance, and an understanding the complexity the instrument. When training teachers for implementation of a new evaluation innovation, Tuytens and Devos (2009) address the awareness of the instrument by examining need, clarifying function, and practicality. *Need* refers to the necessity of the innovation. It is essential that trainers help teachers to understand the importance of the new reform. *Clarifying function* helps teachers to better understand the expectations of the innovation, so that teachers can apply it into practice. If the policy is unclear, it may cause confusion, leading to an increase of mistrust from teacher towards administrators and fostering teachers' vulnerability (Sartain, Stoelinga, & Brown, 2011). Last, the innovation has to have *practicality*. Is the innovation practical as a part of the organizational structure in its true form or altered to fit the school's structure? From a teachers' perspective, practicality is measured by how much the new innovation will cost them personally to implement whether through time invested or effort required for effectiveness (Tuytens & Devos, 2009).

While Tuytens and Devos (2009) discussed teacher training and the implementation of the instrument by explaining awareness of the instrument innovation and its importance, Sartain, Stoelinga, and Brown (2011) deeply explored teacher training by addressing how the instrument is used and its complexity. Trainers must address content, time-frame, and structure of the innovation for increased teacher buy-in during training. When teachers have the opportunity to manipulate the inner workings of the instrument, implementation of the evaluation framework is an easier process, thus

increasing buy-in. Content involves understanding the standards in which teachers will be evaluated. In addition, if there are artifacts to be collected, it is vital to conduct discussions and offer continued professional development to discuss examples of the artifacts and standards (Sartain, Stoelinga, & Brown, 2011). Time-frame addresses making teachers aware of when they will be evaluated. The time-frame should be different for tenured, non-tenured, and probationary teachers. Structure involves the scope and sequence of the framework. Helping teachers to understand the structure gives teachers a sense of clarity regarding what is expected, increasing the chances of buy-in.

Perceptions

"Perception is a process by which individuals organize and interpret their sensory impressions in order to give meaning to their environment" (Schutte, K. J., 2011, p. 6-0). Perceptions are important because they help develop ownership towards ones actions, behaviors, and attitudes based on what their perception of what reality is, not necessarily reality itself. Therefore, the world that is perceived is the world that is behaviorally important (Schutte, 2011, p. 6-0). As educators, teachers and administrators enter the profession with basic beliefs of the teaching profession gained from their educational training. As they evolve through experiences and professional growth, so do their perceptions (Tarman, 2012). Next, I will address administrators' perceptions of teacher evaluation and teachers' perceptions of teacher evaluation, giving pros, cons, and solutions of teacher evaluation from the administrators' and teachers' perceptions.

Administrators' Perceptions of Teacher Evaluation

Despite percentage variance in how much influence teachers have on student success, research strongly supports the indirect influence from school administrators on student achievement (Derrington, 2011; Jacob, 2011; Range, Duncan, Scherz, and Haines, 2012). One way administrators' effect student success is through teacher evaluation (Marshall, 2005; McGreal, 1983) and verifying the evaluation framework fulfills its purpose: to promote growth, remove ineffective teachers, and improvement of professional development. Without principal support of an evaluation system, the quality of the evaluation system carries little merit; an evaluation system can fail due to principal leadership alone (Giacquinta, 1973; Wormmeester, 2005). Therefore, it is important for principals to set the correct tone for its implementation. This charge sounds simplistic, but teacher evaluation is complex as it is often policy driven from the government, and principals can meet many challenges for implementation at the local level (Tuytens & Devos, 2010).

What are principals' general perception towards teacher evaluation? Unfortunately, as policies evolve, research shows that principals are becoming more overwhelmed and finding it cumbersome due to the time demands (Danielson and Marshall, 2005; McGreal, 2000). Danielson states, "Principals have a hard time finding the time to conduct meaningful observations and engage in professional conversations about practice" (2011, p. 39), therefore they do not use the tool accurately or in its entirety. Marshall (2005) complements that the additional work is making the practice more of a process, only when contractual deadlines ensue, instead of an ongoing process of best practice. In a study from Wyoming about principal perspectives of a

new teacher evaluation system, Range, Duncan, and Holt (2011) found negative and positive results. For the advantages, they discovered that principals were taking on more supervisory behaviors than disciplinary behaviors, meaning the new evaluation framework was allowing principals more opportunity for coaching and mentoring rather than directing, creating better relationships with teachers. In addition, the study showed that the new system required them to know more of the standards in the classrooms. However, three common frustrations were identified by principals. These include: time, the instrument, and teachers' willingness to change. These themes were not new as the author listed how several researchers identified the same results. Principals found it difficult to find the time to manage an evaluation system and the instrument itself. Most instruments or evaluation systems were outdated, lacked proper breadth, did not provide meaningful, constructive feedback to teachers, were from a limited point of view, and were cumbersome. Most principals admitted to results in teacher ratings being inflated and adding walk-throughs to gain a better perspective of teacher performance. Teacher's willingness to change depended on administrators using improvement plans, but ownership was admittedly vital towards making it effective.

Teachers' Perceptions of Teacher Evaluation

In general, teachers accept the purpose of teacher evaluation (McGreal, 1983), it is the process that teachers perceive negatively. Nolan and Hoover found that teachers tended to regard the evaluation process as a time to represents themselves in the best light in a short amount of time (2005). The limited time increases stressors, developing a distance between teachers and administrators (Nolan & Hoover, 2005). As a result, teachers perceived themselves as victims and evaluators as someone trying to catch

them in a negative light (Danielson & McGreal, 2000). With the process of teacher evaluation aside, how do teachers feel about the purpose of teacher evaluation?

Research on teachers' perceptions of teacher evaluation systems is scarce (Tuyten & Devos, 2009; Jiang, Sporte, & Luppescu, 2015). Milanowski and Heneman (2001) found only three reputable studies that were previously conducted before their study. Today, that trend is changing as part of the Race of the Top initiative requires data collecting as a process per state on implementation and manipulation of the evaluation system. There are numerous studies with a concentration on the history and purpose of teacher evaluation; administrator as evaluators and their perception; and the reasons that the evaluation system is broken and ways to improve. There is limited research on teacher perception of evaluation systems, but this is slowly changing as states reach to meet the indicators of the RTTT grant (USDOE, 2015). This lack of knowledge about how teachers perceive evaluation systems and their implementation leaves a broad range of areas to be researched in each state, a gap, because teacher perception can be the vital key to the acceptance and stability of an evaluation system. Evaluation requires a partnerships between administrators and teachers (McGreal, 1983). One can have a great evaluation tool, but in the end, an administrator and a teacher must work in partnership to produce effective teaching and student learning. The results will not only help stakeholders in establishing evaluation procedures that are productive for teachers, but further help administrators use the tool to improve the school's organizational system and student success.

With the current wave of political reforms and increasing demand to use teacher evaluations as a tool to identify effectiveness for the result of establishing professional

development, providing growth and or removal, it is imperative to not only understand the purpose of an evaluation tool and determine if it is achieving its purpose, but also to understand how teachers perceive the evaluation tool. There needs to be awareness of teachers' perceptions for true educational change in classroom practice to improve and ensure success of a new teacher evaluation policy (Tuyten & Devos, 2009; Jiang, Sporte, & Luppescu, 2015). Initially, knowing teachers' perception has the potential to overcome previous flaws and prompt positive teachers' reactions to a system (Milanowski & Heneman, 2001). Moreover, teachers' acceptance of system can help to identify factors that account for difference in teachers' favorableness (Milanowski & Heneman, 2001; Peterson & Comeaux, 1990; Tuytens & Devos, 2009; Jiang, Sporte, & Luppescu, 2015) gaining increased buy-in of a program.

Buy-in of a teacher evaluation system is critical to how much responsiveness a teacher puts into the teacher evaluation process (Peterson & Comeaux, 1990; Jiang, Sporte, & Luppescu, 2015). When Peterson and Comeaux (1990) evaluated four teacher evaluation systems based on different purposes, they found that not only are teachers more supportive toward systems they can understand, but understanding how teachers perceive a system is necessary if evaluation is to be considered fair and valid by teachers. Furthermore, and more crucial, is understanding that the process of teacher evaluation is an emotional process for teachers. Teachers invest their 'selves' in their work, therefore there is a lot of vulnerability involved (Nolan & Hoover, 2005; Tuyten & Devos, 2009), and stakeholders need to take that concern into account when adopting new systems. Allowing stakeholders to select an evaluation system is no longer acceptable; teachers need input into the process (Minnici, 2014).

Teachers need to be active participants in the teacher evaluation process (Danielson, 2010). When studying perceptions of administrators and teachers within a new evaluation system, Kyriakides and Demetriou (2007) found statistical significance exhibiting the need to create a change in power from hierarchical to collaborative for a more successful system. This finding is crucial because it represents the need for teachers' perceptions to be recognized. Further, it informs various stakeholders what is needed for systems to be successful. Tuyten and Devos believe, "teachers' perceptions of a new organizationally focused policy on teacher evaluation will be a determining factor for either the success or failure" (2009, p. 926). In most studies, teachers' perception is given minimal recognition (Clipa, 2011; Danielson, 2010; Donaldson, 2009; McGreal, 1983; Rowan, 1990; Tuyten & Devos, 2009), and when it does, it is within other foci stemming toward increasing teachers as active learners-- as an improvement to the system. The trend started to change in the mid-1980s when researchers focused less on teacher behaviors and more on teachers as active decision makers (Rowan, 1990). In the Kyriakides and Demetriou (2007) and the Tuyten and Devos (2009) studies, teachers were given the primary role, making them active participant in deciding the evaluation system. Moving teachers from a secondary role to a more primary role when making decisions about teacher evaluation has higher potential of system success.

Clipa (2011) pinpoints a different perspective by examining teacher evaluation from the teachers' point of view to investigate the purpose of teacher evaluations and the ideal portrait of the assessors for this assessment process. She took a sample of primary teachers and gave them a survey on some aspects of the evaluation process. Her

survey concluded that most teachers see the evaluation process independent of their teaching but that there were some positive correlations. She found that older teachers, who have more work experience recognize that evaluation has an objective for improvement and that permanent training is necessary. Those with a shorter length of service view it as less necessary. Clipa also concluded that evaluations are perceived as a measure to encourage professional responsibility to a significantly higher extent by those who have more work experience from those with less work experience (2011).

Summary

Teacher evaluation is a process that has been conducted for as long as there have been teachers. The intent of the evaluation process has changed over time driven by research and governmental policies changes, but the main themes remain: promoting growth, removing ineffective teachers, and improving professional development. The ultimate goal is to enhance teacher practice and improve student achievement simultaneously.

To meet new policy and research standards, a new teacher evaluation system is being implemented in schools across Oklahoma. The problem is identifying teachers' perceptions towards teacher training for the new implementation and the purpose of teacher evaluation. Being aware of teachers' viewpoints can help district leaders be more effective with implementation procedures creating stronger buy-in and sustainability of the new teacher evaluation system. Further, sharing the level of favorableness teachers have towards teacher evaluation systems with external stakeholders can create stronger partnerships between those that are writing the policies and those implementing and engaged with the system.

CHAPTER III

Methodology

Purpose

This study seeks to quantify (1) the level of favorableness with Oklahoma teachers and their perceptions towards teacher training for implementation of the new teacher evaluation system and (2) teachers' perceptions towards teacher evaluation and its purpose. It is intended to add to the growing field of teacher evaluation by contributing to the knowledge base of necessary awareness in understanding teachers' perceptions when implementing teacher evaluation systems. The study is quantitative and uses a questionnaire for instrumentation. Chapter three will provide a review of the research questions, sample, confidentiality, research design, pilot study, instrumentation, data collection, data analysis, methodology assumptions, limitations, and summary.

Research Questions Re-stated

The primary purpose for this study is to quantify teachers' perceptions towards the new qualitative measure of the Teacher and Leader Effectiveness Evaluation framework. The following questions organize the information related to the identified areas of interest:

1. What are teachers' perceptions regarding teacher training for implementation?
2. What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth?
3. What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers?

4. What are teachers' perceptions about the purpose of the evaluation in regard to receiving professional development?

Sample

The research design was a purposeful sample of teachers in the state of Oklahoma. I used a single-stage sampling procedure due to available access to teachers' districts (Creswell, 2009). The study was limited to Oklahoma teachers participating in either the Marzano Causal Teacher Evaluation Model or Tulsa Teacher and Leader Effectiveness Frameworks.

The state of Oklahoma has 43,840 teachers and 516 school districts (C. Hassell, personal communication, October 25, 2013; K. Isenhour, personal communication, April 14, 2015). 483 school districts selected to use the Tulsa framework and 50 school districts selected to use the Marzano framework. This does not include private schools, but it does include some charter schools that are a part of a public school district. The list of school districts used for this study, was retrieved from the OSDE (2015b). The extracted list provided the names of each district's superintendent, principals of each school, and the email addresses for each superintendent and principal. Data collection was in two phases, superintendent recruitment and the teacher recruitment. During the superintendent recruitment phase, the focus was teacher's email retrieval through superintendent's approval. Permission was required from each school district's superintendent for their teachers to participate in the research study as the questionnaire was being distributed through district electronic mail. The focus of the teacher recruitment phase was to provide all identified teachers with the opportunities to participate in the research study within a cross-sectional timeframe.

Superintendent Data Collection

A recruitment letter was addressed to every Oklahoma district superintendent who used Marzano or Tulsa evaluation frameworks to secure permission for teacher distribution. Embedded in the letter was required procedures and a survey. First, superintendents had to print off the recruitment letter and sign it to represent permission. Then they had to copy the form onto their school's letterhead to confirm consent. Finally, within the survey, they had to upload the signed letterhead as well as district's teacher roster into an online survey software program called Qualtrics. Recruitment for superintendent consent lasted six weeks. From the list, I was able to contact 532 superintendents. The goals was for a 20% return rate (106 superintendents), but of the 234 superintendents who opened the email, 37 superintendents agreed to participate, and only 26 actually followed through reaching 5% of superintendents across the state of Oklahoma. At least eight superintendents who at first agreed to participate, eventually dropped out due to lack of time to follow through with the IRB requirements to provide consent.

Teacher Data Collection

Phase two instigated after receiving consent from superintendents to send the questionnaire to their districts' teachers. I merged all the teachers' email addresses into one panel in Excel and uploaded it into Qualtrics as superintendent responses were received. Next, I used Qualtrics to distribute the questionnaire. Each teacher received a recruitment letter to provide consent before gaining access to the questionnaire. I received permission to distribute questionnaires to 4,856 teachers representing 11% of the total teacher population. From the 4,856 (11 %) teacher emails sent, 466 opened the

email, 41 refused consent, 425 agreed, but only 385 completed the teacher questionnaire in its entirety, which is nearly 8% of the total teacher recruitment. 388 will represent the teacher sample for this study.

I gathered detailed information on teachers, districts, and TLE models that are used in this study from the Oklahoma State Department of Education's legal department at and TLE's executive director.

Confidentiality

Confidentiality of the sample was maintained throughout the study. First, participants were not asked to provide any identifying information, e.g., name, date of birth, employee identification, in the demographic section of the questionnaire. Second, I did not identify the specific schools and districts from which data was collected. Districts are only recognized by either non-rural or rural locations. Third, I informed teachers of these safeguards and gain informed consent at the start of the questionnaire to confirm confidentiality. Ensuring teacher's confidentiality increases the chances of receiving honest responses and lowers any potential validity threats. All data collected were viewed by either the researcher or committee member and transferred into spreadsheets with descriptors removed. All variables were dummy coded.

Research Design

This study is quasi-experimental and uses a cross-sectional research design. The design falls under quantitative methodology, using multiple regression data to locate probable causation between naturally occurring phenomena (Creswell, 2009). The design includes a pilot study in addition to the original research design.

Questionnaire Development Process

I constructed a customized questionnaire to gather data from teachers in the study. Established questions were modified from a questionnaire called The Teacher Evaluation Profile (TEP) used by Tuytens & Devos (2009) originally developed by Stiggins and Duke (1988). The questionnaire was created based on the literature review. Each question selected and developed for the questionnaire instrument was aligned to help answer the research questions. Fowler's (2009) recommendations for designing a survey instrument were followed: (a) drafting tentative questions; (b) conducting a critical review to detect common flaws; (c) conducting individual cognitive interviews; (d) putting questions into the survey instrument, and (e) pre-testing the instrument using proposed data collection procedures. To establish validity and reliability measures, I developed an analysis plan to ascertain the quality of the questionnaire instrument. The process included (a) asking experts in the field to review the questionnaire (Dr. Mary Derrington, University of Tennessee; Dr. Kent Seidel, University of Colorado-Denver; and my committee), (b) making additional revisions, (c) conducting a pilot study, (d) making revisions, and (e) distributing the official questionnaire. During the pilot study, each respondent was asked to answer each question individually. I facilitated discussion with all respondents to determine whether the directions were clear, if the questions were easy to read, whether there was consensus among the respondents in their understanding of each question, whether the answer choices were easy to understand and whether the respondents had any insights into the practical aspects of the questionnaire instrument after respondents completed

the pilot study (Fowler, 1995). Finally, I analyzed results from the pilot study to determine if the results were systematic and replicable. Necessary revisions were made to the questionnaire instrument based upon the data analysis from the pilot study and discussions with participants from the pilot study and committee members. The revised instrument was reviewed by the aforementioned expert panel and approved by my research chair.

The Teacher Evaluation Profile (TEP)

The Teacher Evaluation Profile (TEP) questionnaire was developed by Stiggins and Duke in 1988 (Stiggins & Nickel, 1989) and revised by Doherty (2009). For this study, The TEP questionnaire was the foundation, but each question used was modified to fit my specific needs. Permission was gained from Dr. Daniel Duke to use the TEP questionnaire (see Appendix B). The instrument has been used in multiple studies, usually with revisions as the understanding of teacher evaluation advances (Doherty in 2009; Sheppard, 2013). The original questionnaire consisted of 44 key attributes of teacher evaluation experiences and ranged on a nine-point Likert scale to quantify perceptions (Stiggins & Nickel, 1989). The internal consistency reliability of the questionnaires was .93, representing a highly cohesive set of questions. The validity of the TEP was established during its development and confirmed through three research studies discussed in Stiggins' and Nickel's research (1989).

When Doherty modified the questionnaire to 58 items adding some minor modification to measure perceptions in the overall quality of the TEP process, the internal consistency reliability remained the same (Doherty, 2009). Therefore, the reported internal consistency reliability coefficient of .93 is in line with Cronbach's

(1951) guidelines indicating that reliability coefficients above 0.6 are desirable and values about 0.8 are required for a developed scale. The TEP instrument examined perceptions relating to attributes for effective growth-oriented or formative teacher evaluation. Doherty explains, "The estimate of internal consistency of the total instrument suggest that the scales of each attribute are both internally consistent and highly correlated" (2009, p. 51). With Doherty's modified instrument, the Likert scale was also modified to decrease the range to 1 through 5 (see Appendix D). My modified questionnaire, Oklahoma TLE: Implementation and Purpose of Teacher Evaluation, can be found in the appendix (see Appendix A). Table 3.1 represents the modified TEP items used and how it aligns to this study.

Table 3.1

Teacher Evaluation Profile for Teachers aligned to my questionnaire constructs

Item Number	TEP Item description	My Questionnaire Constructs
Item 30	Were standards communicated to you? (between not at all to in great detail)	Implementation
Item 31	Were the standards clear to you (between Vague to very clear)	Implementation
Item 32	Were the standards endorsed by you as appropriate for our teaching assignment (between not endorsed to highly endorsed)	Implementation
Item 33	Were the standards… (between the same for all teachers to tailored for your unique needs)	Implementation

Item 45	Amount of information received (between none to great deal)	Purpose of teacher evaluation: Improving teacher growth
Item 46	Frequency of formal feedback (between infrequent to frequent)	Purpose of teacher evaluation: Improving teacher growth
Item 47	Frequency of informal feedback (between infrequent to frequent)	Purpose of teacher evaluation: Improving teacher growth
Item 48	Depth of information provided (between shallow to in depth)	Purpose of teacher evaluation: Improving teacher growth
Item 49	Quality of the ideas and suggestions contained in the feedback (between low and high)	Purpose of teacher evaluation: Improving teacher growth
Item 50	Specificity of information provided (between general to specific)	Purpose of teacher evaluation: Improving teacher growth
Item 51	Nature of information provided (between judgmental to descriptive)	Purpose of teacher evaluation: Improving teacher growth
Item 52	Timing of feedback (between delayed to immediate)	Purpose of teacher evaluation: Improving teacher growth
Item 53	Feedback focused on the TAP standards (between ignored the TAP standards to reflected the TAP standards)	Purpose of teacher evaluation: Improving teacher growth
Item 55	Time allotted during the school year for professional development aligned with standards (between none to great deal)	Purpose of teacher evaluation: Improving professional development

Item 56	Availability of training programs and models of good practices (between none to great deal)	Purpose of teacher evaluation: Improving professional development
Item 57	Clarity of policy statements regarding the purpose of evaluation (between vague and very clear)	Purpose of teacher evaluation: Improving professional development. Continuing to train

Researcher Questions

The survey instrument was divided into four sections to align respondents' answers to the research questions. The questionnaires sections included: (1) demographics, (2) teachers' perceptions towards teacher training for implementation of the evaluation framework, (3) teachers' perceptions towards the three purposes of teacher evaluation (professional growth, removal of ineffective teachers, and professional development), and (4) policy. Policy is not a research question for this study, yet it was added to the questionnaire to fill a gap between policy that initiated the problem statement and full implementation of the evaluation framework and eventually the evaluation system.

In section one, categorical questions were developed for the demographic section of this study. While questions were created in such a way that does not reveal the identity of the participants, only one question was needed to be included in the questionnaire that asked for teachers' location of their district. For the purpose of this study, location represents the population of an area. The Economic Research Service [ERS] (2000) recognizes that there are multiple types of population's i.e. small towns, suburban, or country, but for this study population has been categorized into rural and

non-rural. In the state of Oklahoma, the definition of a rural population can be changed depending on its context (ERS, 2000). It can be defined fiscally, geographically, by population, or based on educational needs, to name a few. Even though the rural definition can be vague, the non-rural definition has stayed constant defined as locations with a population of 50,000 or more (ERS, 2000). To represent the Oklahoma population, I had 21 rural and 5 non-rural superintendents respond to my recruitment letter.

Section two, implementation, is divided into three parts to measure teachers' perceptions of received training for implementing their teacher evaluation framework. Part 2A informs on training from the state, district, and building levels; part 2B measures clarity and understanding of the standards that make up the content; and part 2C recognizes awareness of the scope and sequence to understand the structure of the evaluation framework. I created the questions under 2A and 2C based on Sartain's, Stoelinga's, and Brown's (2011) research in the literature review. They discuss the structure of evaluation frameworks with teachers during teacher training and implementation.

In section three, the purpose of teacher evaluation, there are multiple constructs. In 3A, improving teaching growth, I added item 15 to directly align to research question 2. In 3B, removal of ineffective teachers, I added items 16 through 18 to help answer research question 3 and reflect the work of experts in the literature such as Menuey (2005), Donaldson (2009), Weems and Rogers (2010), and Waintroob (1995). In 3C, improve teacher development, I added items 22 through 25 to address research question 4 and reflect the work of the experts in that literature such as Darlinton-Hammond,

Wise, and Peas (1983), McGreal (1983), Danielson (2010), Papay (2012), and Nolan and Hoover (2005).

In section four, Policy and teacher evaluation, I created the questions to tie together the research discussed around policy and the new Oklahoma TLE system. The charge is to be able to discuss various teacher perceptions towards the TLE system, framework, and how policy should be driven.

I ensured content validity (Creswell, 2009) by aligning research questions to the literature and experts in the field. Additional researchers in the field are listed in Table 3.2.

Table 3.2

Content Validity

Research Questions	Item Numbers	Content Validity
What are teachers' perceptions regarding teacher training for implementation?	1-5	Tuytens and Devos (2009), Wagoner and O'Halen (1968), Stiggins and Duke (1988), Doherty (2009), Dainelson and McGreal (2000), and Sertain, L, Stoelinga, S.R., & Brown, E.R., 2011.
What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth	6-15	Tuytens and Devos (2009), Wagoner and O'Halen (1968), Stiggins and Duke (1988), Doherty (2009),, Looney (2011), Weems and Rogers (2010), Danielson and McGreal, Peterson and Comeaux (1990), Papay (2012), and Nolan

		and Hoover, Derrington (2011)
What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers	16-18	Tuytens and Devos (2009), Wagoner and O'Halen (1968), Menuey (2005), Yariv (2009), Donaldson (2009), Weems and Rogers (2010), and Waintroob (1995)
What are teachers' perceptions about the purpose of the evaluation in regards to professional development	19-25	Tuytens and Devos (2009), Wagoner and O'Halen (1968), Stiggins and Duke (1988), Doherty (2009), Papay (2012), McGreal (1983), Danielson (2010), Nolan and Hoover (2005), Darlington- Hammond, Wise, and Peas (1983)

Pilot Study

I submitted my plan of action to the Institutional Review Board (IRB) before beginning the pilot study. The IRB submission was approved with contingency until the official questionnaire was submitted. The questionnaire was developed to align the literature review with the research questions to establish content validity and to confirm that the research questions could be answered on a scale of favorableness.

I developed the questionnaire within the parameters of Fisicaro's (2010) (as cited in Ary, 2010, p.63) and Fowler's (2010) criteria for survey development and effective instrumentation. Criteria for survey development include the following: should be short, simple and direct; questions should be understood by all respondents, and technical terms should be avoided. Further asking a pilot group of respondents similar to the main study group to evaluate the meaning of questions is recommended,

as is avoiding questions that lead to ambiguous answers, avoiding bias in the question wording, avoiding questions that assume traits that might not be present in the sample, avoiding leading questions; avoiding psychologically threatening questions, avoiding double-barreled questions that ask two questions in one, and making answer choices where all possible responses to a question for closed answers are provided, and questionnaire should be kept brief as possible. Respondents are more likely to answer completely and honestly if the survey takes a minimum of time to complete, and care should be taken to ensure that respondents are appropriately knowledgeable to answer the questions. The criteria Fowler listed for the instrument itself are: the questionnaire should be self-explanatory; the items should mainly involve closed answers; only a few forms of questions should be used; the instrument should be visually uncluttered; and cues for respondents to inform them of the next steps in the survey should be provided (Fisicaro, 2010). I worked within these parameters while developing my survey to ensure an effective instrument.

After I developed the questionnaire, I had to recruit pilot study participants. I initially recruited twenty teachers to volunteer to take the questionnaire through social media although only thirteen teachers attended. The procedures for the pilot study was to take the questionnaire in the same situation as the original. Meaning, I gave no discussion beforehand or explanation of the questions during the time they were taking the questionnaire. Afterwards, I gave opportunity for general comments and then discussion for each question. I made several modifications for improvement from the pilot study discussion: I dropped two questions due to confusion (and to increase reliability scores), I reworded two questions, flipped the scale under improving teacher

growth for consistency to align to the other subscales, changed three "I" questions to give teachers something to agree to, and added a back button as an option. The only suggestion I did not adhere to was embedding a dialog box to the end of each construct. After the pilot study meeting, I was ready to analyze the data.

I conducted Reliability Analysis and Exploratory Factor Analysis (EFA) on the pilot study. The questionnaire yielded 4 demographic questions and 33 survey questions within 5 constructs. After discussion with the pilot study participants and using the item deletion and recoding for reliability purposes, the reliability scores for the constructs are: implementation $\alpha = .789$, Teacher growth $\alpha = .777$, Removal of ineffective teachers $\alpha = .674$, Professional development $\alpha = .674$, and Policy $\alpha = .537$. The policy construct scored lowest at $\alpha = .537$. There are five questions within the construct and four questions were recoded for reliability purposes and they were recoded again before conducting the analysis. Removing a question was an option for a higher reliability score, but I decided against it as each question had relevance. The questions for the official questionnaire can be reviewed in Appendix A.

Once I gathered reliability scores, I was ready to begin the research study. I submitted the official questionnaire as a hard copy and as an online survey to the IRB. The IRB gave approval to move forward with the official study.

Instrumentation

The study used a non-experimental cross-sectional questionnaire design that employed a Likert scale from 0 to 5 of closed ended questions to gather data from teachers. A letter of approval was sent to district superintendents to gain consent for distribution of the questionnaire following the approval of the questionnaire by the

Institutional Review Board (IRB) at the University of Oklahoma. The Web-based questionnaire was distributed to teachers through district emails. The questions were entered into a web based software program to collect the required data. Questionnaire results provided quantitative descriptions (Creswell, 2009) of the teachers' level of favorableness towards teachers' perceptions regarding teacher training for implementation of the qualitative portion of the evaluation systems and how teachers' perceive the purpose of teacher evaluation. The questionnaire was completely anonymous even though the questionnaire was implemented through a software program and distributed through school email.

Validity and Reliability

Validity.

The content validity was established by a panel of experts who are scholars in the field. The experts were asked to provide feedback based upon their knowledge and experience with the subject matter. I then asked 13 practitioners to review the questionnaire instrument, engage in a discussion, and offer feedback. Modifications were made based upon a consensus by the dissertation chair, panel experts, and the researcher.

Reliability.

Alpha scores revealed how consistent teachers were in their answers. Scores on a 5 point scale for 33 items were averaged to generate the scores. I established reliability of the questionnaire twice. Once during the pilot study and again after the official study, which are listed in Table 3. There are multiple reliability scores due to the questionnaire having multiple constructs.

Table 3.3 illustrates that some Cronbach's alpha reliability scores were calculated at less than the desirable level of .70. It represents that teachers were not consistent in the way they answered the questions that were grouped within a specific construct. During the pilot study, under the construct *removal of ineffective teachers,* originally there were 4 questions with a .314 reliability score. The factor analysis showed that all questions were one component and was loading above.7 except for question number two, which asked if annual employment should be dependent upon evaluation. This question loaded at -.534. Removing this question and doing the reliability analysis again gave a new reliability score of .674. Under the *Professional Development* construct, I originally had 8 items receiving an alpha of .758, due to discussion with teachers during the pilot study, I felt the need to remove the question professional development was the intended role of evaluation as several participants had different meanings of what the term *role* meant and removing it made the reliability score decline from .718 to .674. For *policy* under pilot study, there are two set of questions that involve answering in opposite sides of the scale. When I take out the questions that teachers tend to score opposite, the score increases, but I felt that all questions in this section were important to this study and future studies.

Removal of ineffective teachers and *policy* continued to establish a reliability score below .70 in the official study. Not only does it represents two topics being discussed within teacher evaluation currently, but likewise teacher inconsistency in their answers.

Table 3.3

Reliability Scores

Constructs	Pilot Study	Item Number	Official Study	Item Number
Implementation	α = .789	8	α = .886	8
Teacher Growth	α = .777	10	α = .937	10
Removal of ineffective teachers	α = .674	3	α = .532	3
Professional development	α = .674	7	α = .851	7
Policy	α = .537	5	α = .623	5

Data Collection

During the official study, data collection was cross-sectional, collected during a six week time period. The data was obtained using a self-administered questionnaire. The questionnaire was distributed to school districts state-wide whose superintendents consented to participation and used the Marzano or Tulsa evaluation frameworks making the research design quasi-experimental, meaning the participants were not randomly selected and assigned (Creswell, 2009).

I requested permission from the IRB to distribute the questionnaire before data collection. I secured permission from each district superintendent to distribute a questionnaire to her teachers, from the Oklahoma State Departments website. The letter

to the superintendent for approval is in the appendix (see Appendix F). Each superintendent gave me access to teachers' email addresses as the questionnaire were distributed through teacher emails. Teachers received a letter of invitation as the initial correspondence where there was a link to the survey (see Appendix G). The letter described the study with its purpose and significance. Questionnaires were emailed multiple times to receive the highest number of participants.

Questionnaires were selected as the tool for data collection because they are inexpensive, easy to administer, and take only a short amount of time to complete. By offering a questionnaire to teachers, this tool may have increased more responses due to the lack of free time in a teacher's day. Questionnaires increase the teacher's level of honesty due to them taking minimal time to complete (Creswell, 2009).

Perceptions will serve as the dependent variables in this study. The frameworks, teacher experiences, locations, and school levels will denote the independent variables (Creswell, 2009).

The questionnaire was intended to measure the different aspects of favorableness of teachers' perceptions: (1) favorableness towards the teachers' perceptions regarding teacher training for implementation of the new frameworks, and (2) favorableness towards the three purposes of teacher evaluation (professional growth, removal of ineffective teachers, and alignment to professional development). In this study, I was looking to see how much of the variance among dependent variables is explained by perceptions.

Once the questionnaires were returned, data was uploaded from Qualtrics into the software program called Statistical Package for Social Sciences (SPSS). SPSS is a

statistical program used to manipulate the data outputting multiple analyses of the results (Salkind, 2011).

Data Analysis

Data analysis informs a reader which form of statistical analysis will be used (Creswell, 2009). For this study, descriptive statistics documented the response to each question in the survey. It describes the frequencies of the teacher respondents. Mean scores were identified to determine the level of favorableness towards the four research questions. Means were also used to compare each of the independent variables with the dependent variables. The study used a linear multiple regression test to predict outcomes and to determine statistical significance and the direction of the relationships between variables. All of the variables were measured against each other to determine the strength and predictability of the variables to determine the known total variance and if there is a statistical significance between any two variables.

Using all the data collected, I was able to answer the research questions to discuss teachers' perceptions towards the Oklahoma teacher evaluation frameworks, its implementation and policy. From the results of the SPSS output, it identified the level of favorableness to the multiple constructs, helped to fill in a gap in the literature and stimulate future research. Additionally, it provides insight to various stakeholders in the roles of implementing the Oklahoma teacher evaluation system. Not only do the results show how teachers perceived the implementation of the TLE framework and purposes, but demonstrated the importance of being aware of teachers' perceptions and the role teachers plays in establishing effective implementation and stability of a school program, such as the teacher evaluation framework and system.

Validity

Internal validity threats are issues of validity with results based upon flaws within the research design and external validity threats are incorrect conclusions made from the data. I identified three internal and one external threats in this study. For each threat recognized, I was transparent and tried to minimize their influence to the best of my ability.

Internal Validity

The most serious validity threat involved ensuring the survey instrument was developed efficiently to answer the research questions. I followed an analysis plan to safeguard the quality of the questionnaire instrument. The steps are listed under research design in chapter three.

History was another internal validity threat. In attempt to control for this threat the survey was distributed at the beginning of the school year before teachers went through their first evaluation process to try avoiding temporary influence of teachers' results.

In addition, this study is bracketed within a specific policy framework called Race to the Top. The focus in only on the quantitative portion of the system. Therefore, it is vital to understand new policies were introduced to amend the TLE system from the qualitative portion, I avoided distractions and maintain consistency with the scope of this study.

The last internal validity threat was diffusion of treatment, which deals with participants' intercommunicating. As teachers converse they can likely influence each other's perceptions. This act may influence how they respond to the survey, but there is

not much I can do to minimalize this concern as it is a current topic in academia and is a required tool teachers are communicating and engaging with on various personal and political levels.

External Validity

Population validity is the only external validity threat identified. Population validity evaluates whether a sample population signifies the entire population and if the sampling method was adequate (Shuttleworth, 2009). Teachers receiving access to the questionnaire required approval from district superintendents and interest from teachers. The sample size represents 8% (388 teachers responded to the questionnaire) of 11% (4,856 teachers) of the Oklahoma teacher population who were allowed by their superintendents to receive the questionnaire, the findings may not generalize across all Oklahoma regions or other states.

Methodological Assumptions

I made several assumptions for this study including the following: the methods and procedures selected for this questionnaire are appropriate for the subjects studied; the teachers will honestly and accurately answer all questions to the best of their knowledge; the participant teachers represent a purposeful sample of the targeted population identified for this research study; my analysis of the data will be accurate and represent responses of all data collected; email will be an appropriate method to collect data for this research study; teachers have access to the Internet and the abilities needed to use and complete the questionnaire. Further, I assumed teachers would demonstrate an understanding of the information regarding their districts, school sites,

and the evaluation systems to accurately respond to the questionnaire. The findings and conclusions of the study will not be generalizable past the limitations of the study.

Limitations and Delimitations

This research study utilized data returned from an Internet questionnaire. This methodology results in limitations. There is limited knowledge on teachers' perceptions of teacher evaluation. Therefore, the literature review base is restricted. The data for the study is limited to the teachers who responded to the questionnaire. The data analysis was completed through a multiple regression test that assesses correlations in variables; however, correlation does not imply causation (Creswell, 2009). Data is self-reported, therefore "the validity of the information is contingent on the honesty of the respondent" (Mertens, 2010, p. 173).

Delimitations

Review of research is focused on teachers' perceptions with limited research from other points of view. This study is a quantitative design. The inclusion of a qualitative aspect may develop a more balanced study of teachers' perceptions. The survey included only a closed ended Likert scale, rather than including a section for comments. The review of literature is bracketed within the Race to the Top initiative and excludes previous policy such as No Child Left Behind reforms and future ones such as the Every Student Succeeds Act (ESSA) that was signed December 10, 2015. The review of literature emphasizes the purpose of the teacher evaluation and limited information on the use or the various types of evaluation. The topic of teacher evaluation inevitably overlaps with other topics such as value-added, performance pay, termination, teacher effectiveness, accountability systems, and numerous other fields.

However, the intent of this research is to focus on the purpose of teacher evaluation and the need for teachers' perspectives. The study participants were limited to teachers who had participated in either the Marzano Causal Teacher Evaluation Model or the Tulsa Teacher and Leader Effectiveness frameworks.

Summary

This chapter summarized the quantitative research design that was used in this study to examine teachers' perceptions of Oklahoma's new teacher evaluation system. The study was based upon teachers' perceptions regarding teacher training for implementation, teachers' perceptions about the purpose of the evaluation in regards to improving professional growth, teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers, and teachers' perceptions about the purpose of the evaluation in regards to aligning professional development to the teacher evaluation. The research methods, data collection techniques, and data analysis were thoroughly explained, along with content validity and pilot testing procedures.

CHAPTER IV

Results

Overview

The findings presented in Chapter 4 may provide insight to policymakers and educational leaders who are involved with continuous improvement developing the Oklahoma TLE System. The data collected was examined to quantify the level of favorableness teachers hold toward the three purposes of teacher evaluation, the training for implementation in Oklahoma public schools, and the policy that initiated this study.

This chapter encompasses the findings from the data collected through Qualtrics. This chapter is organized into seven sections: (1) review of research methods, design, and instrumentation, (2) review of research questions, (3) data collection, (4) data analysis, (5) descriptive statistics, (6) inferential analyses, and (7) summary. Review of research methods, design, and instrumentation reestablishes the methods and design. The review of research questions reminds readers of the research questions for the study. Data collection is a review of how data was collected, stored, and kept confidential. Data analysis identifies the reliability and validity measures and discusses why I chose multiple regression. Descriptive statistics categorizes the frequencies of respondents. Frequencies were then used to gain mean scores. Mean scores were identified to determine the level of favorableness towards the four research questions and policy and to compare each of the independent variables against the dependent variables. Five multiple regression were completed to ascertain inferential analysis: I performed one multiple regression for each dependent variable (questionnaire construct). The dependent variable was run against the independent

variables to determine variance and predictability of the variables to establish if there was statistical significance between any two variables. Using all the data collected, I answered the research questions to predict teachers' perceptions towards the Oklahoma teacher evaluation framework, its implementation, and policy to determine if the study has statistical significance. The summary concludes the chapter.

Review of Research Methods, Design, and Instrumentation

This quantitative multiple regression study used data collected from teachers across Oklahoma through a questionnaire instrument. The study was cross-sectional being completed in a six-week timeframe. The design included a pilot study in addition to an original research design to determine reliability and validity. The independent variables included location, framework, experience, and grade level. The dependent variables included implementation, improvement of teacher growth, removal of ineffective teachers, improvement of professional development, and policy. I used data to make predictions between naturally occurring phenomena, resulting in the findings below.

Research Questions Restated

The primary purpose for this study was to quantify teachers' perceptions towards the new qualitative measure of the Teacher and Leader Effectiveness Evaluation framework. The following questions organize the information related to the identified areas of interest:

1. What are teachers' perceptions regarding teacher training for implementation?
2. What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth?

3. What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers?
4. What are teachers' perceptions about the purpose of the evaluation in regard to receiving professional development?

Data Collection

I received IRB approval prior to the distribution of the research questionnaire. All data collection was gathered electronically through Qualtrics. To begin, an informational email was sent to all Oklahoma superintendents (see Appendix F) to explain the purpose of the survey and provide contact information for any questions. In response to the initial email, superintendents gave consent to distribute questionnaires to their teachers. This study was limited to the number of superintendents who consented to their districts participating in the study. Superintendents then chose to either forward their teachers email roster for me to send teachers the questionnaire or they personally forwarded the questionnaire to their teachers. In the fall of 2015, I received permission to distribute questionnaires to 4,856, teachers representing 11% of the Oklahoma teacher population. Then the official survey was distributed using Qualtrics to collect data from teachers. The survey was available for six weeks. The survey consisted of 37 Likert scale questions separated into demographics and five constructs: implementation, professional growth, removal of ineffective teachers, professional development and policy (see Appendix A). From the 4,856 (11 %) teacher emails sent, 466 opened the email, 41 refused consent, 425 agreed, but only 385 completed the teacher questionnaire in its entirety, which is nearly 8% of the total teacher recruitment. 388 represents the teacher sample for this study.

Data Analysis

Data analysis procedures were initiated by transferring data from Qualtrics to SPSS v.23 software. Reliability scores were established and compared to the pilot study's reliability scores. I dummy coded all of the demographics or independent variables from the questionnaire to protect confidentiality. I then calculated the descriptive statistics to determine the rating of each item on the survey (frequencies, means, and standard deviations). I used linear multiple regression for my inferential analysis. I chose multiple regression to identify the variance and statistical significance between variables. More specifically, if dependent variables were influenced by independent variables, in what direction and magnitude did one independent variable show more favorability than another? Five linear multiple regressions were completed in this study where one regression aligned with each of the instrument constructs.

Official Questionnaire Reliability Analysis

This section will discuss the reliability scores of the official survey instrument. The study has multiple reliability scores due to the questionnaire having multiple constructs. I tested for internal reliability by calculating Cronbach alpha reliability scores. The scores were computed to ascertain the degree of internal consistency and reliability among the key dependent variables. Reliability scores revealed how consistent teachers were in their answers. Scores were averaged on a 5-point scale for 33 questionnaire items. Reliability coefficients were calculated twice: once during the pilot study and again after the official study. The constructs' reliability scores for the questionnaire: *Implementation* is $\alpha = .886$; *Teacher Growth* is $\alpha = .937$; *Removal of*

Ineffective Teachers is α = .674; *Professional Development* is α =.851; and *Policy* is α = .623.

Alpha scores between .70 and .79 are considered acceptable; scores of .80 or more are desirable (George & Mallery, 2003). The construct's *Removal of Ineffective Teachers* and *Policy* Cronbach's alpha reliability scores were calculated at less than the desirable level of .70. It infers that teachers were inconsistent in answering the questions that were grouped within the construct. It also represents two controversial topics currently within teacher evaluation.

Descriptive Statistics

Section one of the questionnaire ask teachers to respond to demographic items. For this study, the sample included teachers who were evaluated using the Marzano or Tulsa frameworks. The sample was divided by teachers who lived in either rural or non-rural locations; who were tenured or non-tenured; and who taught at the elementary or secondary school level. From the analysis, three hundred and eighty-five teachers participated in the teacher evaluation study. Of these, 178 (46.2%) teachers used the Marzano framework and 207 (53.8%) used the Tulsa framework. The majority of participants were tenured teachers with 307 (79.9%). The teachers were mostly located in non-rural 224 (58.2%) locations. Teacher experience was mostly at the secondary level with 199 (51.7%). There was not a big disparity between descriptive statistics except for the tenured and non-tenured teachers. All frequencies and percentages are provided in Table 4.1.

Table 4.1

Frequencies and Percent of Questionnaire Demographics (N= 385)

Characteristics	n	%
Location		
Non-Rural	161	41.8
Rural	224	58.2
Frameworks		
Marzano	178	46.2
Tulsa	207	53.8
Experience[a]		
Tenure	307	79.9
Non-Tenured	77	20.0
Grade Level[a]		
Elementary	185	48.2
Secondary	199	51.7

[a]One participant reported neither experience nor grade level taught, so N = 384.

Means

Means scores align the average teacher responses to the Likert scale. The alignment answers the research questions by identifying the level of favorableness of teachers' perceptions. The survey instrument was developed in construct form to accommodate the literature review and answer the research questions. Each construct asked questions ranging from 1 (strongly disagree/low favorableness) to 5 (strongly agree/ high favorableness) on a Likert scale with choices as Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree and Not Applicable (which gave zero points). The questionnaire included five constructs and this section is arranged in the order of the questionnaire: Implementation, improving teacher growth, removal of ineffective teachers, improve professional development, and policy. Below depicts the statistics

and frequencies of each item in the questionnaire excluding the demographics, see Table 4.2. The means scores indicate a range between a two and three.

Table 4.2

Questionnaire Statistics and Frequencies

Implementation	$\bar{X}$	s.d	SD (%)	D (%)	N (%)	A (%)	SA (%)
… STATE LEVEL…	2.20	1.065	87(22.6)	147(38.2)	76(19.7)	41(10.6)	7(1.8)
… DISTRICT…	2.90	1.223	44(11.4)	109 (28.3)	69 (17.9)	112(29.1)	30 (7.8)
… BUILDING LEVEL…	3.20	1.243	41(10.6)	71(18.4)	79(20.5)	122(31.7)	54(14.0)
… Standards communicated….	3.14	1.156	27(7.0)	88 (22.9)	69(17.9)	141(36.6)	30(7.8)
… Very clear to me.	3.24	1.142	21(5.5)	88 (22.9)	62(16.1)	148(38.4)	37(9.6)
… Standards are appropriate …	2.81	1.245	61(15.8)	101 (26.2)	64(16.6)	103 (26.8)	27 (7.0)
… Training for understanding…	2.84	1.066	29(7.5)	132 (34.3)	73(19.0)	105 (27.3)	14 (3.6)
… High level of understanding …	2.93	1.135	33 (8.6)	104 (27.0)	95(24.7)	90 (23.4)	30(7.8)
Improving teacher growth	$\bar{X}$	s.d	SD (%)	D (%)	N (%)	A (%)	SA (%)
Amount of information…	2.89	1.189	29(7.5)	126(32.7)	65(16.9)	82(21.3)	35(9.1)
I received formal feedback…	2,79	1.233	41(10.6)	131(34.0)	51(13.2)	79(20.5)	34(8..8)
I received informal feedback…	2.74	1.245	42(20.9)	140(36.4)	49(12.7)	65(1.9)	38 (9.9)
Depth of feedback…	2.99	1.283	38(9.9)	108(28.1)	58(15.1)	82(21.3)	51(13.2)
The quality of the ideas … helpful.	3.02	1.305	39(10.1)	100(26.0)	57(14.8)	86(22.3)	53(13.8)
Information provided was specific …	2.89	1.304	46(11.9)	116(30.1)	56(14.5)	66(17.1)	52(13.5)
The feedback provided was objective …	2.79	1.233	39(10.1)	133(34.5)	65(16.9)	57(14.8)	43(11.2)
Timing of the feedback was useful…	2.77	1.247	43(11.2)	129(33.5)	61(15.8)	63 (16.4)	40(10.4)
Feedback focused on the standards.	2.42	1.109	56(14.5)	162(42.1)	54(14.0)	39(10.1)	22(5.7)

The evaluation… improved growth.	3.48	1.416	17(4.4)	53(13.8)	68(17.7)	78(20.3)	106(27.5)
Removal of ineffective teachers	$\bar{X}$	s.d	SD (%)	D (%)	N (%)	A (%)	SA (%)
… Fire teachers who don't function	2.91	1.246	43(11.2)	101(26.2)	68 (17.7)	87 (22.6)	37(9.6)
… Will remove ineffective teachers	2.73	1.250	64(16.6)	91(23.6)	75 (19.5)	79 (20.5)	27(7.0)
… Could remove EFFECTIVE teachers	3.59	1.254	19 (4.9)	48(12.5)	54 (14.0)	123 (31.9)	89(4.9)
Improve professional development	$\bar{X}$	s.d	SD (%)	D (%)	N (%)	A (%)	SA (%)
Time allotted during the school year…	2.87	1.211	44(11.4)	102(26.5)	47(12.2)	116(30.1)	19(4.9)
A useful variety of PD …	2.75	1.173	49(12.7)	110(28.6)	59(15.3)	95(24.7)	16(4.2)
The policy statements … were clear.	2.98	1.115	31 (8.10	95(24.7)	71(18.4)	114(29.6)	19(4.9)
… Increase professional development…	2.64	1.119	45(11.7)	132(34.3)	60(15.6)	78 (20.3)	14(3.6)
… Build strengths/support weaknesses.	2.55	1.095	58(15.1)	113(29.4)	82(21.3)	65(16.9)	10(2.6)
… Individualized PD…	1.92	1.117	108(28.1	12(31.4)	39(10.1)	38(9.9)	3(.8)
… Framework improved PD	2.03	1.063	97(25.2)	118(30.6)	69(17.9)	26(6.8)	5 (1.3)
Policy	$\bar{X}$	s.d	SD (%)	D (%)	N (%)	A (%)	SA (%)
… SYSTEM eventually effective…	2.48	1.171	82(21.3)	86(22.3)	76(19.7)	68(17.7)	9(2.3)
… Reform should be district driven.	3.70	1.067	12(3.1)	35(9.1)	58(15.1)	144(37.4)	72(18.7)
… Reform should be policy driven.	2.65	1.262	70(18.2)	92(23.9)	61(15.8)	74(19.2)	24(6.2)
… FRAMEWORK was improvement.	2.35	1.107	81(21.0)	96(24.9)	84(21.8)	52(13.5)	5(1.3)
The … SYSTEM was an improvement.	2.23	1.107	84(22.1)	99(25.7)	77(20.0)	53(13.8)	4(1.0)

After examining mean scores for each question in the questionnaire, mean scores were analyzed by construct. I used SPSS v.23, to identify the total items, points possible, and means of the questionnaire by construct (see Table 4.3). To determine the

level of favorableness a range was developed using respondents' total numbers and the total mean scores. The *Implementation* construct has eight survey items with a total of 40 possible points. *Improving Professional Growth* construct has ten survey items with a possible 50 possible points. *Removal of Ineffective Teachers* construct has three survey items with a possible 15 possible points. *Improving professional development* construct has seven survey items with a possible 35 possible points. The *Policy* construct has five survey items with a possible 25 possible points.

I calculated mean item responses after I identified range and mean scores, see Table 4.3. A means item response aligned favorableness of each construct to the Likert scale. For implementation, the mean score was a 20.11 of 40 possible points (range of 8) representing 2.513 responses towards the research question: what are teachers' perceptions regarding teacher training for implementation? Professional growth's mean score was a 26 of 50 possible points (range of 10) representing 2.606 responses towards the research question: what are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth? Removal of ineffective teachers mean score was a 9.22 of 15 possible points (range of 3) representing 3.073 responses towards the research question: what are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers? Professional development mean score was a 17.78 of 35 possible points (range of 7) representing 2.540 responses towards the research question: what are teachers' perceptions about the purpose of the evaluation in regard to aligning professional development to the teacher evaluation? Policy was 13.50 out of 25 possible points (range of 5) representing 2.700 responses towards teachers' perceptions of the policy initiative.

Table 4.3

Total Possible Point and Means Report

	Implementation	Improving Teacher Growth	Removal of Ineffective Teachers	Improve Professional Development	Policy
Total Items	8	10	3	7	5
Total Possible	40	50	15	35	25
Mean	20.11	26.06	9.22	17.78	13.50
Total Number	351	324	338	327	322
Standard Deviation	6.030	9.093	2.695	2.695	3.616
Mean Item Response	2.513	2.606	3.073	2.540	2.700

Comparison Means

The construct mean scores fell between a 2.5 and 3.1 leaning towards non-favorableness in multiple constructs. Because the scores did not confirm strong indication of favorableness nor non-favorableness towards the research questions, I compared independent variables against each other, see Table 4.4. Analyzing frameworks reveal that teachers who were evaluated using the Marzano frameworks scored higher favorability in the area of teacher growth. Teachers who were evaluated using the Tulsa framework scored higher favorability in all other areas. For experience, tenured teachers exhibited higher favorability than non-tenured teachers in the areas of implementation, teacher growth, and removal of ineffective teachers. Non- tenured teachers scored higher favorability with professional development and policy. For

teacher location, rural teachers scored higher favorability in the areas of implementation, removal of ineffective teachers, professional development and policy. Non-rural teachers' scored higher favorability in the area of teacher growth. For grade level, secondary teachers scored higher favorability in every construct. Table 4.4 represents which independent variable had a higher level of favorability within each dependent construct.

Table 4.4

Comparison Means

	Framework						Location					
	Marzano			Tulsa			Rural			Non-Rural		
	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N
Implementation	18.57	6.218	160	21.40	5.562	191	20.98	5.597	147	19.49	6.262	204
Growth	28.93	8.972	147	23.67	8.504	177	24.47	8.767	138	27.23	9.175	186
Removal	8.97	2.955	154	9.43	2.444	184	9.45	2.648	143	9.06	2.723	195
PD	16.61	5.451	150	18.77	5.793	177	18.39	5.897	138	17.33	5.581	189
Policy	13.15	3.653	150	13.81	3.564	172	13.87	3.611	135	13.24	3.605	187
	Experience						Grade Level					
	Non-Tenured			Tenured			Elementary			Secondary		
	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N
Implementation	20.03	6.888	74	20.13	5.803	276	19.74	6.111	169	20.45	5.960	181
Growth	24.16	8.780	70	26.57	9.141	253	25.94	9.005	153	26.14	9.219	170
Removal	8.86	3.063	71	9.32	2.590	266	8.93	2.794	162	9.50	2.584	175
PD	18.65	6.453	68	17.56	5.528	258	17.46	5.627	160	18.10	5.847	166
Policy	14.82	3.992	67	13.15	3.438	254	13.49	3.533	157	13.50	3.710	164

Inferential Analyses: Multiple Regression

I used multiple regression to study multiple independent variables, called predictor variables, in relation to the single dependent variable, called an outcome variable (Fields, 2009). A series of linear multiple regressions were calculated to predict whether location, framework, experience, and grade level (independent variables) had a relationship with teachers' perceptions towards the new implementation of teacher evaluation frameworks, the three purposes of teacher evaluation (improvement of teacher growth, removal of ineffective teachers, and improvement of professional development), and policy (dependent variables). Organization of this section followed the order of the questionnaire. To begin I coded the categorical independent variables as numerical 0 or 1, see Table 4.5

Table 4.5

Categorical Independent Variables

Independent Variable	0	1
Location	Non-rural	Rural
Framework	Marzano	Tulsa
Experience	Tenured	Non-Tenured
Grade Level	Secondary	Elementary

Implementation

A multiple regression was run to determine if grade level, teacher experience, location, and framework were predictors of teachers' perceptions towards teacher training for implementation of the Oklahoma TLE frameworks. Results of the regression analysis indicated there was a statistically significant model, $R^2 = .060$, F

(4,345) = 5.481, p. = < .001, see Table 4.6. The model results indicate that 6% of the variance in the implementation scores is attributable to the four independent variables.

Table 4.7 includes the regression coefficients for implementation. Framework was the only independent variable that was significantly related to implementation. Teachers who were part of the Tulsa framework scored 3.3 points higher than those using the Marzano framework. None of the other independent variables were statistically significant.

Table 4.6

Simultaneous Multiple Regression Model Summary for Implementation

Model	R	R^2	R^2adj	Std Error of the Estimate	F *chg*	*df*	*Sig*
1	.244	.060	.049	5.83	5.481	4, 345	.000

a. Predictors: (Constant), Level, Experience, Location, Framework
b. Dependent Variable: Implementation

Table 4.7

Simultaneous Multiple Regression Results for Regression Equation

	Unstandardized Coefficients		Standardized Coefficients		
model	B	SE	β	t	Sig.
1 (Constant)	18.589	.641		29.020	.000
Framework	3.340	.852	.276	3.921	.000
Experience	.622	.792	.042	.786	.433
Location	-647	.852	-.053	-.760	.448
Level	-.338	.635	-.028	-.532	.595

Improving Teacher Growth

I ran a second multiple regression to determine if grade level, teacher experience, location, and framework were predictors of teachers' perceptions towards teacher evaluation and improving teacher growth. Results of the regression analysis indicated there was a statistically significant model, R^2= .118, F (4,318) = 10.613, p. = < .001, see Table 4.8. The model results indicated that 12% of the variance in the improving teacher growth scores is attributable to the four independent variables.

Table 4.9 includes the regression coefficients for improving teacher growth. Framework and experience were the only independent variables that were significantly related to improving teacher growth. Teachers who were part of the Marzano framework scored -6.7 points higher than those using the Tulsa framework. Tenured teachers scored -.3.9 points higher than non-tenured teachers. None of the other independent variables were statistically significant.

Table 4.8

Simultaneous Multiple Regression Model Summary for Improving Teacher Growth

Model	R	R^2	R^2*adj*	Std Error of the Estimate	F *chg*	*df*	*Sig*
1	.343	.118	.107	8.606	10.613	4, 318	.000

a. Predictors: (Constant), Level, Experience, Location, Framework
b. Dependent Variable: Teacher Growth

Table 4.9

Simultaneous Multiple Regression Results for Regression Equation

model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	SE	β		
1 (Constant)	30.603	.970		31.539	.000
Framework	-6.706	1.295	-.367	-5.178	.000
Experience	-3.892	1.196	-.17	-3.255	.001
Location	.950	1.293	.052	.735	.463
Level	-.945	.968	-.052	-.976	.330

Removal of Ineffective Teachers

I ran a third multiple regression to determine if grade level, teacher experience, location, and framework were predictors of teachers' perceptions towards the teacher evaluation and the removal of ineffective teachers. Results of the regression analysis indicated there was not a statistically significant model, R^2= .020, F (4,332) = 1.692, p. = > .050, see Table 4.10. In addition, there were no independent variables that were significantly related to removal of ineffective teachers see Table 4.11.

Table 4.10

Simultaneous Multiple Regression Model Summary for Removal of Ineffective Teachers

Model	R	R^2	R^2*adj*	Std Error of the Estimate	F *chg*	*df*	*Sig*
1	.141	.020	.008	2.688	1.692	4, 332	.152

a. Predictors: (Constant), Level, Experience, Location, Framework
b. Dependent Variable: Removal of Ineffective Teachers

Table 4.11

Simultaneous Multiple Regression Results for Regression Equation

model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	SE	β		
1 (Constant)	9.369	.297		31.581	.000
Framework	.238	.397	.044	.599	.550
Experience	-.365	.368	-.055	-.993	.321
Location	.138	.397	.025	.347	.729
Level	-.530	.296	-.098	-1.791	.074

Professional Development

I ran a fourth multiple regression to determine if grade level, teacher experience, location, and framework were predictors of teachers' perceptions towards the teacher evaluation and improvement of professional development. Results of the regression analysis indicated there was a statistically significant model, R^2= .050, F (4,321) = 4.251, p. = < .001, see Table 4.12. The model results indicate that 5% of the variance in professional development scores is attributable to the four independent variables.

Table 4.13 includes the regression coefficients for professional development. Framework and experience were the only independent variables that were significantly related to professional development. Teachers who were apart of the Marzano framework scored 2.7 points higher than those using the Tulsa framework. Tenured teachers scored 1.6 points higher than those who were non-tenured teachers. None of the other independent variables were statistically significant.

Table 4.12

Simultaneous Multiple Regression Model Summary for Professional Development

Model	R	R^2	R^2*adj*	Std Error of the Estimate	F *chg*	*df*	*Sig*
1	.224	.050	.038	5.628	4.251	4, 321	.002

a. Predictors: (Constant), Level, Experience, Location, Framework
b. Dependent Variable: Professional Development

Table 4.13

Simultaneous Multiple Regression Results for Regression Equation

	Unstandardized Coefficients		Standardized Coefficients		
model	B	SE	β	t	Sig.
1 (Constant)	16.352	.633		25.822	.000
Framework	2.704	.844	.235	3.204	.001
Experience	1.625	.786	.115	2.068	.039
Location	-.518	.845	-.045	-.613	.540
Level	-.315	.630	-.027	-.500	.618

Policy

I ran a fifth multiple regression to determine if grade level, teacher experience, location, and framework were predictors of teachers' perceptions towards policy. Results of the regression analysis indicated there was a statistically significant model, R^2= .056, F (4,316) = 4.683, p. = < .001, see Table 4.14. The model results indicate that 5.6% of the variance in the policy scores is attributable to the four independent variables.

Table 4.15 includes the regression coefficients for policy. Teacher experience was the only independent variable that was significantly related to policy. Teachers

who were non-tenured scored 1.9 points higher than teachers who were tenured. None of the other independent variables were statistically significant.

Table 4.14

Simultaneous Multiple Regression Model Summary

Model	R	R^2	R^2adj	Std Error of the Estimate	F *chg*	*df*	*Sig*
1	.237	.056	.044	3.538	4.683	4, 316	.001

a. Predictors: (Constant), Level, Experience, Location, Framework
b. Dependent Variable: Policy

Table 4.15

Simultaneous Multiple Regression Results for Regression Equation

	Unstandardized Coefficients		Standardized Coefficients		
model	B	SE	β	t	Sig.
1 (Constant)	12.473	.399		31.277	.000
Framework	.638	.537	.088	1.189	.235
Experience	1.939	.496	.218	3.906	.000
Location	.535	.539	.073	.993	.321
Level	.104	.400	.014	.260	.795

Summary

Chapter four discussed data that was collected through a computerized survey called Qualtrics and exported into SPSS v.23 a statistical software. The study was initiated with the approval of the IRB. The chapter reviewed the research design and data analyzed to measure mean scores and identify statistical significance by running five multiple regressions to predict relationships among independent and dependent variables. The data quantified the level of favorableness teachers hold toward the three

purposes of teacher evaluation, the training for implementation in Oklahoma public schools, and the policy that initiated the study.

CHAPTER V

Discussion

In academic year 2013-2014, the state of Oklahoma began the implementation phase of a new teacher evaluation system. The goal was to align with the facets in the Race to the Top Grant to increase the possibility of receiving the funds. The Oklahoma TLE commission was charged with implementation of a more rigorous evaluation system to increase teachers' accountability and foster student success. The benefits of an effective teacher evaluation system are numerous and well documented in the research, yet the study is lacking when discussing the teachers' perspectives or framed within quantitative analysis (Milanowski & Heneman, 2001; Tuytens & Devos, 2009; Wormmeester, 2005; Jiang, Sporte, & Luppescu, 2015). This empirical study is intended to partially fill the literature gap within the current scholarship.

Summary of the Purpose

This study examined Oklahoma teachers' perceptions regarding the newest teacher evaluation system. There were two purposes for this study: (1) to quantify perceptions held by Oklahoma teachers to determine levels of favorableness towards the new teacher evaluation framework (the qualitative portion of the TLE); and (2) to quantify perceptions held by Oklahoma teachers to determine levels of favorableness towards teacher training of the TLE implementation. Do teachers believe that the new evaluation system effectively met the purpose of evaluations including improving teacher quality, removing ineffective teachers, and professional development? In addition, were teachers' perceptions favorable or unfavorable toward the

implementation of the new program? The concomitant intent was to identify educators' perceptions and recognize variance among the identified perceptions.

Summary of the Literature

The teacher evaluation process has been conducted for many years. The intent of the evaluation process has transformed over time, driven by research and governmental policies changes, but the main themes remain: promoting growth, removing ineffective teachers, and improving professional development. The ultimate goal of evaluation is to enhance teacher practice and improve student achievement simultaneously.

In order to fulfill new policy requirements and research standards, a new teacher evaluation system was implemented in schools across Oklahoma. The problem for this study was identifying teachers' perceptions towards teacher training for implementation and the three purposes of teacher evaluation.

The study of the Oklahoma TLE System was bracketed within a specific timeframe. It began with the American Recovery and Reinvestment Act and Oklahoma's attempts to receive the RTTT grant (which is embedded within the act) and ended in the Fall 2015 academic year once I had collected teachers' perceptions of favorableness through a questionnaire.

Research Questions

The primary purpose for this study was to quantify teachers' perceptions towards the new qualitative measure of the Teacher and Leader Effectiveness Evaluation framework. The following questions organized the information related to the identified areas of interest:

1. What are teachers' perceptions regarding teacher training for implementation?
2. What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth?
3. What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers?
4. What are teachers' perceptions about the purpose of the evaluation in regard to receiving professional development?

Summary of the Methodology

This quantitative multiple regression study used data collected from teachers across Oklahoma through an on-line questionnaire instrument. The questionnaire included 23 questions ranked using a Likert scale. The study was completed in a six-week timeframe, making it cross-sectional. The design included a pilot study in addition to an original research design to determine reliability and validity. The independent variables were location, framework, experience, and grade level. The dependent variables were implementation, improvement of teacher growth, removal of ineffective teachers, improvement of professional development, and policy. I identified the means, compared means, and analyzed the data for statistical significance and relationships using five multiple regressions.

Summary of Findings

This research analysis proposes the succeeding findings to further an understanding of the teacher evaluation process and teachers' perceptions towards its implementation and the purpose:

Descriptive Statistics

From the analysis, three hundred and eighty-five teachers participated in the teacher evaluation study. Of these, 178 (46.2%) teachers used the Marzano framework and 207 (53.8%) used the Tulsa framework. The majority of participants were tenured teachers with 307 (79.9%). The teachers were mostly located in non-rural 224 (58.2%) locations. Teacher experience mainly encompassed the secondary level with 199 (51.7%). There was not a large disparity between descriptive statistics except between the tenured and non-tenured teachers.

Means Scores

The mean scores answered the research questions by identifying the level of favorableness of teachers' perceptions. Each construct asked questions ranging from 1 (strongly disagree/low favorableness) to 5 (strongly agree/ high favorableness) on a Likert scale with the choices: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree and Not Applicable (which gave zero points). When examining individual questions and questions within a construct, the mean scores' results indicated a range between a two and a three on the Likert scale; when assessing individual questions and questions within each construct, responses depicts a level of low favorableness of teachers' perceptions toward the purpose of teacher evaluation and its implementation. The construct, removal of ineffective teachers, was the only construct that scored neutral on the Likert scale.

Comparison Means

Recognizing teachers mostly showed low favorableness towards teacher evaluation and its implementation, the independent variables were compared against

each other to determine which variables had the lowest favorability towards the research questions. Analyzing frameworks revealed that teachers who were evaluated using TLE scaled lower favorability towards teacher growth and Marzano teachers scaled lower favorability with implementation, removal of ineffective teachers, and professional development. Analyzing experience showed non-tenured teachers exhibited lower favorability than tenured teachers in the areas of implementation, teacher growth, and removal of ineffective teachers. Analyzing experience revealed tenure teachers had lower favorability with professional development and policy. Analyzing teacher location revealed non-rural teachers scored lower favorability in the areas of implementation, removal of ineffective teachers, professional development, and policy; rural teachers scored lower favorability in the area of teacher growth. Analyzing grade level unveiled elementary teachers scored lower favorability in every construct.

Multiple Regressions

A series of linear multiple regressions were calculated to determine statistical significance and which independent variables had a relationship with teachers' perceptions of the dependent variables. There was 6% variance with implementation to denote statistical significance with Marzano framework having the highest relationship. There was 12% variance with improving teacher growth and it showed statistical significances with teachers who used the Marzano framework and tenured teachers. Removal of ineffective teachers did not show statistical significance or significant variables. There was 5% variance with professional development and it showed statistical significances with Marzano and tenured teachers having the highest

relationship. There was 5.6% variance with policy and it showed statistical significance with non-tenured teachers having the highest relationship.

Research Questions Answered

What are teachers' perceptions regarding teacher training for implementation?

For individual questions, the means scores demonstrated that a majority of teachers deemed implementation of the framework as the responsibility of districts and schools, not the state. Teachers felt they were trained appropriately and that the standards were appropriate for what they were teaching, but educators did not have a high level of understanding towards the structure of the given framework.

Teachers' perceptions regarding the construct scaled with low favorableness. The construct represented a 2.5 mean item response. Independent variables that scored the lowest favorableness were non-tenured elementary teachers who used the Marzano framework and lived in a non-rural locations. There was statistical significance with framework having the strongest relationship.

What are teachers' perceptions about the purpose of the evaluation in regards to improving professional growth?

For individual questions, the means scores showed a majority of teachers represented favorably towards the evaluation frameworks improving teacher growth but they represented low favorableness towards the evaluator's guidance being the origin of teacher growth. All questions representing what actions the evaluator used to help promote growth was scored negatively. Yet, as a framework, improving teacher growth was scored positively

Teachers' perceptions of the construct professional growth scaled with low favorableness. The construct represented a 2.6 mean item response. The independent variables that scored the lowest favorableness were non-tenured elementary teachers who used the TLE framework and lived in rural locations. There was statistical significance with framework and experience having the strongest relationship.

What are teachers' perceptions about the purpose of the evaluation in regards to removing ineffective teachers?

For individual questions, the means scores disclosed that teachers felt the new frameworks were not implemented to fire teachers nor did they favor firing ineffective teachers who received low evaluation scores. Teachers did perceive that the frameworks could remove effective teachers. Therefore, teachers did not feel that the new frameworks would achieve the purpose to remove ineffective teachers.

Teachers' perception of the construct removing ineffective teachers scaled neutral. The construct represented a 3.1 mean item response. The independent variables that scored the lowest favorableness were non-tenured elementary teachers who used the Marzano framework and teachers who lived in non-rural locations. There was no statistical significance or strong relationships. Which displayed lack of reliability in teacher responses towards this construct.

What are teachers' perceptions about the purpose of the evaluation in regard to receiving professional development?

For individual questions, the means scores showed that the majority of teachers felt they received adequate professional development on the new frameworks and system. Teachers perceived negatively in regards to receiving professional

development aligned to their teacher evaluation results to enhance individual growth. Teachers did not recognize that districts were using the new frameworks to gather evaluation results and use them to build upon teachers' strengths and improve teachers' needs. Therefore, teachers disagreed that the new teacher evaluation frameworks would meet the purpose to improve professional development.

Teachers' perceptions of the construct to improve professional development scaled with low favorableness. The construct represented a 2.5 mean item response. The independent variables that scored the lowest favorableness were tenured elementary teachers who used the Marzano framework and lived in non-rural locations. There was statistical significance with framework and experience having the strongest relationships.

What are teachers' perceptions about the purpose of the evaluation in regard to policy?

In addition to the four research questions, I analyzed another construct to enclose the study, called policy. For individual questions, the means scores showed that the majority of teachers preferred reforms as district driven, not state driven mandated through policies. Teachers identified both the frameworks and the system as improvements from past evaluations, yet teachers did not discern that the fully implemented Oklahoma TLE system would improve teacher competency.

Teachers' perceptions of the construct represented a 2.7 mean item response. Teachers' perceptions regarding policy scaled with low favorableness. The independent variables that scored the lowest favorableness were tenured elementary teachers who

used the Marzano framework and lived in non-rural locations. There was statistical significance with experience having the strongest relationship.

Connections to the Literature

Implementations

From the literature, there are two stages for successful implementation during teacher training/implementation: awareness of the instrument innovation and its importance, and understanding the complexity of the instrument (Tuytens and Devos, 2009). Analysis showed that communication is lacking with understanding the complexity of the instrument. Sartain, L., Stoelinga, S.R., and Brown, E. R. (2011) advised that during training, trainers must address the structure of the innovation for increased teacher buy-in. When teachers have the opportunity to manipulate the inner workings of the instrument, implementation is easier and buy-in increases. Without both steps of implementation, this void causes a lack of clarity leading to teacher mistrust towards administrators and fostering teacher vulnerability (Sartain, L., Stoelinga, S.R., and Brown, E. R., 2011). Hence, the low favorableness towards implementation correlated to the findings.

Professional Growth

Administrators promoting teachers' growth validate that what teachers do is important (Donaldson, 2009; OECD, 2005). The literature and analysis results confirmed that teachers accept the purpose of teacher evaluation and the notion that teachers can grow from evaluation frameworks. The breakdown involved how teachers perceived their alleged professional growth gains. Results showed teachers view negatively administrators' intentions to improve teachers' growth. Yet, evaluation

entails a partnership between administrators and teachers (McGreal, 1983). The relationship between the administrators and teachers has been vulnerable through the years as administrators' roles have continued to fluctuate between teacher evaluator and teacher improvement, creating confusion and mistrust. To gain higher favorableness, evaluators must continue to transition from hierarchical partnerships with teachers to more collaborative partnerships, giving teachers an active role in the learning process (Kyriakides and Demetriou, 2007; Tuyten and Devos, 2009). Principals and teachers working together facilitates effective teaching and student learning creating an output of teacher growth.

Removal of Ineffective Teachers

Teacher evaluation can be used as a tool to remove ineffective teachers and the new Race to The Top reform requires the tool to be used to fulfill the requirement. Yet research shows that teachers are rarely dismissed on the grounds of performance (OECD 2005). Likewise this study recognized that teachers continue to perceive that the new evaluation system will not remove ineffective teachers. In addition, teachers did not believe the framework was implemented to fire teachers. Conversely, teachers did perceive that the new evaluation system would remove effective teachers. This construct represented low reliability results, neutral on the Likert scale, and no statistical significance. Further research needs to be generated in this area of study.

Professional Development

Teacher evaluation systems can improve professional development by identifying common thematic areas where teachers represent weaknesses and then implement training to improve the identified areas. The study analysis aligned with the

literature finding that teachers' perceived professional development unrelated to teaching practices (OECD, 2005). The study also confirmed the literature that professional development lacks connectedness between teacher evaluation and professional development, creating a deficiency of information received from the evaluation tool (Rowan, 1990).

Policy

Based on low favorableness that the evaluation framework will improve teacher competency, I still ask, is the evaluation tool broken or in need of repair? For all constructs, Oklahoma teachers validate what researchers have identified in previous studies concerning ongoing mistrust between teachers and administrators to develop teacher growth; that teacher evaluation is not removing ineffective teachers; that professional development is not individualized to benefit the teacher; and that teachers view the evaluation frameworks unfavorably towards improving teacher competency. So, what is the new evaluation system measuring? What is the intended purpose of teacher evaluation today? Maybe it is time for the definition to be modified, broadened, or changed all together. Perhaps it is time to change the school structure or curriculum and stop trying to change the teacher (McGreal, 1983).

Recommendations

To gain higher favorability of implementation, districts should be proactive with communicating to teachers how the TLE framework is progressing. As a result, meetings should be held frequently to review and comprehend the standards and framework, and the evaluators should be working with teachers to help them gain a higher understanding of the system and its results. This communication cannot be

shared once a year at the beginning of a school year; instead, it must be continuous. Teachers need to understand that the framework or system is a part of the district's culture, not something that is just required until the next policy is implemented. Increased professional development will foster the process of true implementation.

To gain higher favorability of the evaluation system improving teacher growth, evaluators need to have a more collaborative approach to the process. Teachers need to feel that the evaluation meetings are for educators' benefits, not solely a requirement set by policy. Administrators can accomplish this collaborative approach with teachers by helping educators to identify teachers' strengths and weaknesses and then provide objective feedback that is frequent (informally and formally), informative, and goal-oriented.

To gain higher favorability towards the evaluation system removing ineffective teachers, teachers must see direct outcomes. Teachers are aware of which teachers are ineffective in their buildings and continue to hinder students. School administrators have the responsibility of protecting students from ineffective teachers and should have the authority to make changes when necessary.

Further, administrators need to have authentic conversations about why teachers are concerned that the new frameworks will possibly remove effective teachers. It requires clear and constant communication through each phase of implementation

To gain higher favorability towards the evaluation system and professional development, professional development will need to go through systemic change from an organizational level. The research aligns with the teachers' perspectives of this study

that districts perform the required professional development to meet the majority of teachers' needs, but rarely focuses on individualized teacher's needs.

To gain higher favorability towards the evaluation system and policy, everything discussed above will need to progress. This study further demonstrates that teachers do not have buy-in of the Oklahoma TLE framework or system. Lack of buy-in creates difficulty from policy makers to administrators when attempting to fully implement the system with utility.

Suggestions for Further Research

The findings and conclusions in this study lead to the following recommendations for further research and study: (a) Continuing to work on the questionnaire to gain higher reliability scores; (b) Gaining permission from more superintendents to involve a larger number of teachers; (c) A more in-depth study utilizing qualitative data obtained through interviews. Qualitative data will clear up some of the complexities under the constructs removing ineffective teachers and the additional construct, policy; (d) An addition study can be done to determine if Oklahoma's teacher shortage is effecting how principal's score teachers in high need areas; (e) An addition study can be done to determine if teachers who have chosen to leave in the last two years is due to the constraints of the evaluation framework; and (f) Additional study could be done to see how teachers' perceptions have changed as reforms have been made and the districts continue to work towards full implementation.

Summary

Oklahoma teachers' perceptions towards the implementation of the new evaluation framework are favored negatively from the analysis. The majority of focus

pertained to non-tenured elementary teachers who work in non-rural locations. These unfavorable viewpoints towards the framework or system represent a lack of teacher buy-in and difficulty in understanding the true purposes of teacher evaluation: to increase teacher growth, remove ineffective teachers, and receive professional development. Further research needs to be executed to deepen the knowledge of teachers' perceptions towards teacher evaluation. To have open conversations towards implementing a highly effective evaluation tool that gives teachers the desired effects of improving growth, removing ineffective teachers, and getting professional development.

REFERENCES

AIR & OSDE (2014). Oklahoma student learning objectives/ Student outcome objectives frequently asked questions (SLOs/SOOs FAQ) 2014–15. *American Institute for Research.* Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/SLO%20SOO%20FAQ.pdf

Alexander, R.V. (1957). Teacher evaluation program. Ladue public schools. *Journal of Teacher Education, 8*(2), 148-152. Doi: 10.1177/002248715700800209

Alexandrov, D. C. (1989). Teacher evaluation in an era of school reform. New York, NY: United States Department of Education. (ERIC Document Reproduction Service No. ED318081). University of Oklahoma Library retrieved April 15, 2014, from the ERIC database.

Alkin, M.C. & Christie, C.A. (2004). An Evaluation Tree. In Evaluation Roots (Chapter two). Retrieved from http://www.sagepub.com/upm-data/5074_Alkin_Chapter_2.pdf **DOI:** http://dx.doi.org/10.4135/9781412984157

Anderson, J.M. (1969). Teacher evaluation. *The Clearing House, 43*(9), 522.

Beller, E.K. (1971). Teacher evaluation: Why, what, and how! *Peabody Journal of Education, 48*(2), 125-139.

Bolman, L. G. & Deal, T. E. (1997). *Reframing organizations*. San Francisco, CA: Jossey-Bass Inc.

Clark, D. (1993). Teacher evaluation: A review of literature with implication for educators. Received from ERIC database. (ED359174)

Clipa, O. (2011). Teacher perceptions on teacher evaluation: the purpose and the assessors within the assessment process. *Procedia- Social and Behavioral Sciences, 29,* 158-163. Doi: 10.1016/j.sbspro.2011.11.220

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Cronbach, LJ. (1951). Coefficient alpha and internal structure of tests. Psychometrics, 297-334

Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership, December/January,* 35-39.

Danielson, C. & McGreal, T. L. (2000). *Teacher Evaluation to enhance professional practice.* Alexandria, VA: Association for Supervision and Curriculum Development and Educational Testing Service.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6), 8-15.

Darling-Hammond, L., Wise, A. E., and Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, Fall, 53*(3),285-328. Doi: 10.3102/00346543053003285

Derrington, M. L. (2011). Changes in teacher evaluation: Implications for the principal's work. *The Delta Kappa Gamma Bulletin, Spring*, 51-54.

Derrington, M. L., & Campbell, J. (2013). The changing conditions of instructional leadership: Principal perceptions of teacher evaluation accountability mandates. In B. G. Barnett, A.R. Shoho, & A. J. Bowers (Eds.), *International research on school leadership: Vol. 4. School and district leadership in an era of accountability* (pp. 231-251). Charlotte, NC: Information Age Publishing Inc.

Doherty, J. F. (2009). *Perceptions of teacher and administrators in a Massachusetts suburban school district regarding the implementation of a standards-based teacher evaluation system* (Order No. 3431766). Available from ProQuest Dissertations & Theses Full Text. (814798727). Retrieved from http://search.proquest.com/docview/814798727?accountid=12964

Donaldson, M. (2009). *So long Lake Wobegon? Using teacher evaluation to improve teacher Quality.* Washington, DC: Center for American Progress. Retrieved from http:/www.americanprogress.org/issues/2009/06/pdf/teacher_evaluation.pdf

Fisicaro, R. J. (2010). *Teacher evaluation: Assessing principals' perceptions in the state of New Jersey.* (Order No. 3391110, Liberty University). *ProQuest Dissertations and Theses,* 141. Retrieved from http://search.proquest.com/docview/305244976?accountid=12964. (305244976).

Fowler, F. (1995). Improving survey questions: Design and evaluation. Thousand Oaks, CA: Sage

Fowler, F. (2009). Survey research methods (4th ed.). Thousand Oaks, CA: Sage.

Fullen, M. (1998). The meaning of organizational change: A quarter of a century of learning. In A. Liberman (Ed.), *The roots of educational changes* (pp. 202-216). New York, NY: Springer.

George, D., & Mallery, P. (2003). SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update, (4th ed.) Boston: Allyn & Bacon.

Glisson, J. (2014). Teacher and leader effectiveness: Next steps for 2014. Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/TLE%20Next%20Steps%20for%202014%20FINAL.pdf.

Goldrick, L. (2002). Improving teacher evaluation to improve teaching quality. Issue Brief. *National Governors' Association, Washington DC. Center of Best practices.* Retrieved from http://www.nga.org/cda/files/1202IMPROVINGTEACHEVAL.pdf.

Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. *National Comprehensive Center for Teacher Quality.* Retrieved from http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf

Giacquinta, J.B. (1973). The process of organizational change in schools. *Review of Research in Education, 1,* 178-206. Doi: 10.3102/0091732X001001178.

Gross, N., Giacquinta, J. B., & Bernstein, M. (1971). *Implementing Organizational Innovations: A Sociological Analysis of Planned Educational Change*. New York, NY: Basic Books, Inc.

Harris, D. N. & Sass, T. R. (2009). What makes for good teacher and who can tell? Working Paper 30. Washington, D.C.: National Center for the Analysis of Longitudinal Data I Education Research, Urban Institute.

Jacob, B.A. (2011). Do principals fire the worst teachers? *Educational Evaluation and Policy Analysis, 33*(4), 403-434. Doi: 10.3102/0162373711414704.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of Tennessee value added assessment system. *Educational Evaluation and Policy Analysis,25*(3), 287-298.

Kyriakides, L. & Demetriou, D. (2007). Introducing a teacher evaluation system based on teacher effectiveness research: An investigation of stakeholders' perception. *Journal of Personnel Evaluation in Education*, *20*(1-2), 43-64.

Littrell, K. (2014). OEA TLE evaluation system survey fall 2014: Results and analysis [PowerPoint slides]. Retrieved from OEA.

Looney, J. (2011). Developing High-Quality Teachers: teacher evaluation for improvement. *European Journal of Education, 46*(4), 440-455.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan,8,* 727-735.

Marshall, K. (2012). Fine-tuning teacher evaluation. *Educational Leadership, 70*(3), 50-53.

Maslow, V. J. & Kelly, C. J. (2012). Does evaluation advance teaching practice? The effects of performance evaluation on teaching quality and system change in large diverse high schools. *Journal of School Leadership, 22*(3), 66-32.

Marzano, R. J. (2012). The two purpose of teacher evaluation. *Educational Leadership, 70*(3), 14-19.

McGee, L. (2013). TLE Report to the State Board of education. Retrieved from http://ok.gov/sde/sites/ok.gov.sde/files/documents/files/TLE_Report_Oklahoma_State board Education7_15_2013.pdf

McGreal, T. L. (1983). *Successful teacher evaluation.* Alexandria, VA: Association for Supervision and Curriculum Development.

McGuinn, P. (2011). Stimulating reform: Race to the top, competitive grants and the Obama education agenda. *Educational Policy, 26 (1)*, 136-159.

Menuey, B. P. (2005) Teachers' perceptions of professional incompetence and barriers to the dismissal process. *Journal of Personnel Evaluation in Education, 18,* 309-325.

Mertens, D. M. (2010). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Milanowski, A. T. & Heneman, H. G. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: A pilot study. *Journal of Personnel Evaluation in Education, 15*(3), 193-212.

Minnici, A. (2014). Mind shift in teacher evaluation: Where we stand- and where we need to go. *American Educator, 38*(1), 22-36.

Moe, T. M. (1995). The politics of structural choice: Toward a theory of public bureaucracy. In O. E. Williamson (Ed.), Organizational theory (pp.116-153). Oxford, NY: Oxford University Press.

National Council on Teacher Quality. (2011). State of states: Trends and early lessons on teacher evaluation and effectiveness policies. Retrieved from http://www.nctq.org/p/publications/docs/nctq_stateOfTheStates.pdf.

Nixon, A., Packard, A. L., & Douvanis, G. (2010). Non-renewal of probationary teachers: Negative retention. *Education, 131*(1), 43-52.

Nolan, J. & Hoover, L. A. (2005). *Teacher supervision and evaluation: Theory into practice.* Hoboken, NJ: John Wiley & Sons, INC.

OECD. (2005) *Teachers Matter: Attracting, Developing and Retaining Effective Teachers,* OECD, Paris. Doi: 10.1787/9789264018044-en.

OECD. (2009) *Teacher Evaluation: A Conceptual Framework and Examples of Country Practices,* OECD, Mexico.

Oklahoma State Department of Education. (1999). Oklahoma criteria for effective teaching and administrative performance: Activities workbook. New York, NY: United States Department of Education. (ERIC Document Reproduction Service No.ED439103). Retrieved from http://files.eric.ed.gov/fulltext/ED439103.pdf.

Oklahoma State Department of Education. (2012). Oklahoma Law. Retrieved from http://www.ok.gov/sde/school-laws-oklahoma-16

Oklahoma State Department of Education. (2012). Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE) Report. Retrieved from http://ok.gov/sde/sites/ok.gov.sde/files/TLE-ReportLeg.pdf

Oklahoma State Department of Education. (2012).Oklahoma Teacher evaluation. Retrieved from http://www.marzanoevaluation.com/evaluation/Oklahoma_Teacher_Evaluation/

Oklahoma State Department of Education. (2013). Teacher and Leader Effectiveness: Other Academic Measures. Retrieved from http://ok.gov/sde/tle-quantitative-components#oam

Oklahoma State Department of Education. (2014). Teacher and Leader Effectiveness: Other Academic Measures. Retrieved from http://ok.gov/sde/tle-quantitative-components#oam

Oklahoma State Department of Education (2014, July 15). Exploring student learning objectives (SLOs) & student outcome objectives (SOOs) [PowerPoint Slides]. Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/Stepping%20Up-Vision%202020%20Presentation.pdf

Oklahoma State Department of Education (2015). Current TLE information (March 3, 2015). Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/Current%20TLE%20Information030315.pdf.

Paige, M. (2012). Using VAM in high-stakes employment decisions. *Kappan Magazine, 94 (3),* 29-32.

Papay, J. P. (2012). Refocusing the debate: assessing the purpose and tools of teacher evaluation. *Harvard Educational Review, 82*(1), 123-141.

Peterson, P. L. & Comeaux, M. A. (1990). Evaluating the system: Teachers' perspectives on teacher evaluation. *Educational Evaluation and Policy analysis, 12*(1), 3-24.

Popham, W. J. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education, 1,* 269-273.

Range, B. G., Duncan, H. E., & Holt, C. R. (2011) Supervision and evaluation: The Wyoming perspective. *Springer, (23),* 243-265. Doi: 10.1007/s11092-011-9123-5

Range, B. G., Duncan, H. E., Scherz, S. D., & Haines, C. A. (2012). School leaders' perception about incompetent teachers: Implications for supervision and evaluation. *NASSP Bulletin, 96 (4),* 302-322. Doi; 0.1177/0192636512459554.

Roth v. Board of Regents, 408 U.S. 564 (1972).

Rothman, R. (1010). Value added and the teacher incentive fund. *School Administrator, 67* (22), 22.

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education, 16,* 353-389.

Salkind, N. J. (2011). *Statistics for people who (Think They) hate statistics* (4th ed.). Thousand Oaks, CA: Sage Publications.

Sartain, L., Stoelinga, S. R., and Brown, E. R. (2011). Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation. Retrieved from https://ccsr.uchicago.edu/sites/default/files/publications/Teacher%20Eval%20Report%20FINAL.pdf

Schutte, K. J. (2011). Perception and individual decision making. In Robbins & Judge Global (Eds.), Organizational Behavior (pp. 6-0 – 6-24). New York, NY, Pearson Education. Retrieved from http://www.academia.edu/6526846/Perception_and_Individual_Decision_Making

Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice, 17*(2), 151-164.

Sheppard, J. D. (2013). *Perceptions of teachers and administrators regarding the Teacher Evaluation Process.* Unpublished doctoral dissertation. Georgia Sourthern University. Retrieved from http://digitalcommons.georgiasouthern.edu/cgi/viewcontent.cgi?article=1856&context=etd.

Shuttleworth, M. (2009). Population Validity. Retrieved Aug 04, 2016 from https://explorable.com/population-validity

SREB. (2014). Oklahoma teacher and leader evaluation system: Focus group findings. Retrieved from http://bloximages.newyork1.vip.townnews.com/tulsaworld.com/content/tncms/assets/v3/editorial/d/9c/d9c9f060-7e70-5ecc-9a0f-7a5559e0ee1e/54cb184b87911.pdf.pdf.

Stiggins, R. J. & Duke, D. (1988). The case for commitment to teacher growth: Research on teacher evaluation. Albany, NY: State University of New York Press

Stiggins, R. & Nickel, P. (1989) The Teacher Evaluation Profile: A technical analysis. *Journal of Personnel Evaluation in Education*. L(S).

Stufflebeam, D. L. (2001). Evaluation checklist: Practical tools for guiding and judging evaluations. *American Journal of Evaluation, 22*(1), 71-79.

The Center on great teachers and Leaders, N.D. (N.D.). Databases on state teacher and principal valuation policies. *American Institute for Research.* Retrieved from http://resource.tqsource.org/stateevaldb/StateRoles.aspx.

Thomas, T. (2005). Better teacher evaluation. Retrieved from http://www.orosok.org/aOROS2006/jan05newsletter.html.

Tuyten, M. & Devos, G. (2009). Teachers' perception of the new teacher evaluation policy: A validity study of the policy characteristics scale. *Teaching and Teacher Education, 25* (6), 924-930.

U.S. Department of Education. (2009a). *Partners in Reform.* Retrieved from http://www.ed.gov/ news/speeches/partners-reform

U.S. Department of Education. (2009b). *Race to the Top Program: Executive Summary.*Retrieved from http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

U.S. Department of Education. (2014). *Race to the Top Funds*. Retrieved from https://www2.ed.gov/programs/racetothetop/phase1-applications/index.html?exp=0.

Wagoner R. L. & O'Hanlon, J. P. (1968). Teacher attitudes towards evaluation. New York, NY: United States Department of Education. (ERIC Document Reproduction Service No. ED013236). Retrieved from http://files.eric.ed.gov/fulltext/ED013236.pdf.

Waintroob, A. R. (1995). Don't fool with incompetent teachers. *Educational Digest, 61*(2), 36-40.

Wallace, J. D. (2012). Teacher evaluation: A conversation among educators. *Phi Delta Kappan, 94*(30), 44-46.

Weems, D. M. & Rogers, C. B. H. (2010). Are US teachers making the grade? A proposed framework for teacher evaluation and professional growth. *Management in Education, 24*(1), 19-24. Doi: 10.1177/0892020609354959

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Retrieved from http://www.widgeteffect.org.

White House: Office of the Press Secretary, "Race to the Top Fact Sheet". Retrieved from https://www.whitehouse.gov/the-press-office/fact-sheet-race-top

Wormmeester, R. B. (2005). High school teachers' and administrators' perceptions of the teacher evaluation process in California's public schools (Doctoral dissertation). Retrieved from ProQuest. 3180408.

Yariv, E. (2009). Principals' informal methods for appraising poor-performing teachers. *Educational Assessment, Evaluation, and Accountability, 21*, 283-298. Doi: 10.1007/s11092-009-9081-3.

Zepeda, S. J. & Ponticelli, J. A. (1998). At cross-purposes: What do teachers need, want, and get from supervision? *Journal of Curriculum and Supervision, 14*(1), 68-87.

APPENDICES

APPENDIX A

Questionnaire

Oklahoma TLE: Implementation and Purpose of Teacher Evaluation

Section 1: Demographic Information

I. The location of your district is mostly:
   a. Non rural (Population more than 50,000)
   b. Rural (Population less than 50,000)

II. Which framework is your district implementing?
   a. Tulsa (TLE)
   b. Marzano

III. With total teaching experience, are you a tenured teacher or non-tenured teacher?
   a. Non-tenure (< 3 years)
   b. Tenure (≥ 3 years)

IV. Your current teaching assignment grade level (select the answer that best describes)
   a. Elementary (PreK- 6)
   b. Secondary (7-12)

Section 2: Part A. Teacher Training for Implementation of the teacher evaluation framework: General

Teacher Evaluation Framework. For the purpose of this study, the teacher evaluation framework is defined as the preselected teacher evaluations by the TLE

Commission which includes both the Marzano Casual Teacher Evaluation Model and the Tulsa Teacher and Leader Effectiveness Frameworks.

Please rank your perceptions towards the statement:

1. The overall quality of teacher training provided by the STATE LEVEL for implementing the teacher evaluation framework was high.

   Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

2. The overall quality of teacher training provided by the DISTRICT LEVEL for implementing the teacher evaluation framework was high.

   Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

3. The overall quality of teacher training provided by the BUILDING LEVEL for implementing the teacher evaluation framework was high.

   Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

Section 2: Part B. Teacher Training for Implementation of the teacher evaluation framework: Content.

"Standards" are the content criteria used in your evaluation framework to evaluate your teaching.

Directions: Please rank your perceptions towards procedures related to content in the items below:

1. During training, the standards were communicated thoroughly and clearly.

   Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

2. The standards are very clear to me.

Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

3. The standards are appropriate for my teaching assignment.

Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

Section 2: Part C. Teacher Training for Implementation of the teacher evaluation framework: Structure.

"Structure" means the outline and organization of the evaluation framework to give teachers an understanding of evaluation expectations.

Directions: Please rank your perceptions towards procedures related to structure in the items below:

4. The level of training for understanding the structure of the framework was adequate to comprehend the scope and sequence.

Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

5. I have a high level of understanding of the structure of the framework.

Strongly Disagree 1 2 3 4 5 Agree Strongly N/A

Section Three. Part A. Purpose of teacher evaluation: Improving teacher growth.

Directions: Please rank your perceptions towards administrator's level of providing feedback.

6. Amount of information received from evaluation meetings were helpful.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

7. I received formal feedback from my evaluator frequently.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

8. I received informal feedback from my evaluator frequently.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

9. Depth of feedback towards my teaching was thorough.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

10. The quality of the ideas and suggestions in the feedback was very helpful.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

11. Information provided was specific in relation to my teaching.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

12. The feedback provided was objective in relation to my teaching.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

13. The timing of the feedback was useful to me.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

14. Feedback focused on the evaluation standards.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

15. The new teacher evaluation framework has improved my teacher growth.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

Section Three. Part B. Purpose of teacher evaluation: Removal of ineffective teachers.

Directions: Please rank your perceptions towards the removal of ineffective teachers.

16. Teacher evaluation is a means to fire teachers who don't function well.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

17. The new teacher evaluation framework will help to remove ineffective teachers.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

18. The new evaluation framework could remove EFFECTIVE teachers.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

Section Three. Part C. Purpose of teacher evaluation: Improve professional development.

Directions: Please rank your perceptions towards resources available for evaluation.

19. Time has been allotted during the school year for professional development aligned with evaluation standards.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

20. A useful variety of professional development programs and models of good practices have been offered.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

21. The policy statements about the purpose of evaluation were clear.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

22. There has been an increase in professional development to assist the implementation of the teacher evaluation framework.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

23. Teacher evaluation has been aligned to build on teachers' strengths and support their weaknesses.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

24. I was able to participate in individualized professional development that addressed weaknesses identified by my evaluation.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

25. The new teacher evaluation framework improved professional development.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

Section Four. Part A. Policy and the Teacher Evaluation System

Teacher Evaluation System. For the purpose of this study, the teacher evaluation system is defined as the TLE system as a whole representing both the quantitative and qualitative measures.

Teacher Evaluation Framework. For the purpose of this study the teacher evaluation framework is defined as the preselected teacher evaluations by the TLE Commission. Both the Marzano Casual Teacher Evaluation Model and the Tulsa Teacher and Leader Effectiveness Framework.

Directions: Please rank your perceptions towards the Oklahoma TLE System.

26. Teacher evaluation reform should be district driven.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

31. Teacher evaluation reform should be policy driven.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

32. The policy adoption of the teacher evaluation FRAMEWORK was an improvement.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

33. The policy adoption of the teacher evaluation SYSTEM was an improvement.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

34. The teacher evaluation SYSTEM will eventually be an effective means towards improving the competence of a teacher.

Strongly Disagree 1 2 3 4 5 Agree Strongly

N/A

APPENDIX B

Permission to use Teacher Evaluation Profile (TEP)

From: Clute, Sharla
Sent: Tuesday, April 28, 2015 10:00 AM
To: beatty

Thank you for requesting to use material from the SUNY Press book The Case for Commitment to Teacher Growth: Research on Teacher Evaluation edited by Richard J. Stiggins and Daniel L. Duke. It is our policy to not require permission for use of our material in an unpublished thesis.

If the thesis is later published in any format using this material you will need to seek permission. Please feel free to review our guidelines for requesting reprint permission that is available on our website: http://www.sunypress.edu/l-43-reprint-permissions.aspx

The material may be photocopied by your dissertation committee for internal display/review purposes only.

We appreciate the standard source citation such as the following:

"Reproduced by permission from The Case for Commitment to Teacher Growth: Research on Teacher Evaluation edited by Richard J. Stiggins and Daniel L. Duke, the State University of New York Press ©1988, State University of New York. All rights reserved."

Best wishes with your thesis defense.

Sincerely,

Sharla Clute

SUNY Press/Rights and Permissions/ 22 Corporate Woods Blvd., 3rd Floor Albany, NY 12211

From: Beatty
Sent: Wednesday, April 22, 2015 8:35 PM
To: SUNY Press Web Site
Subject: Permission to use Questionnaire
Hello:

I am currently a doctoral student at the University of Oklahoma and I am working on my dissertation with a focus on Teacher Evaluation. In the book, The Case for Commitment to Teacher Growth: Research on Teacher Evaluation (Richards & Stiggins, 1988), there is a questionnaire in the appendix called Teacher Evaluation Profile (TEP). I am requesting permission to use the questionnaire.

Thanks,

Ivana Beatty

From: Daniel L. Duke
Sent: Thursday, April 23, 2015 6:53 AM
To: beattylms@aol.com

Dear Ivana:

Thank you for your request to use the TEP Questionnaire. You have my permission to use the instrument. I'll be interested to find out what you discover.

All the best,
Daniel L. Duke
Professor
University of Virginia

On Thu, 23 Apr 2015 00:49:08 +0000
Hello:

I am writing to request permission to use the questionnaire you developed called the Teacher Evaluation Profile (TEP). Currently, I am a doctoral student at the University of Oklahoma, Norman Campus, in the Department of Educational Leadership and Policy Studies Executive, Ed.D. I am in the process of writing my dissertation, which focuses on Oklahoma Teachers' Perceptions towards the New Teacher Evaluation System and its Implementation.

Thanks for your help,

Ivana Beatty
The Case for Commitment to Teacher Growth: Research on Teacher Evaluation (with Richard Stiggins), (Albany, NY: State University of New York Press, 1988).

APPENDIX C

Model Section Criteria for Tulsa Teacher and Leader Effectiveness Framework and the Marzano Causal Teacher Evaluation Model

http://www.tulsaschools.org/4_About_District/_documents/TLE/Handbook_TLE_Observation_and_Evaluation_System_8-7.pdf.

Teacher and Leader Effectiveness Commission Teacher Evaluation Framework/Model Selection Criteria

OKLAHOMA STATE DEPARTMENT of EDUCATION JANET BARRESI

The review following the criteria set forth by the Teacher and Leader Effectiveness Commission will be conducted by the Oklahoma State Department of Education and submitted to the Commission for consideration.

Selection	Description
Does not meet the criteria	From the materials provided, does not appear to meet the criteria
Meets the criteria	From the materials provided, appears to meet the criteria
Exceeds the criteria	From the materials provided, appears to exceed the criteria

Teacher Framework: Tulsa Public Schools Teacher Leader Effectiveness Observation and Evaluation System

TULSA PUBLIC SCHOOLS - TEACHER

Required by Statute	Teacher Evaluation Framework/Model Selection Criteria for the Qualitative Measures	Does not meet the criteria	Meets the criteria	Exceeds the criteria	Reviewers' Notes
✓	Includes a Five-Tier Rating System (qualitative measures) 1) Superior 2) Highly Effective 3) Effective 4) Needs Improvement 5) Ineffective		✓		
✓	Annual evaluations that provide feedback to improve student learning and outcomes		✓		
✓	Comprehensive remediation plans and instructional coaching for all teachers rated as Needs Improvement or Ineffective		✓		
✓	Be evidence-based (e.g. research and field proven)		✓		The evidence base and field experience is one district
✓	Include observable and measurable characteristics of personnel and classroom practices (rubrics and evidences for each rubric)		✓		
✓	Be correlated to student performance success (validation studies and causal links studies for instructional strategies/behaviors)	✓			There is some encouraging evidence emerging in the district but not yet a research study
✓	Be based on research-based national best practices and methodology (contemporary research and practices of expertise development and strategies/behaviors for raising student achievement)		✓		

Teacher and Leader Effectiveness Commission Teacher Evaluation Framework/Model Selection Criteria

TULSA - TEACHER CON'T

Required by Statute	Teacher Evaluation Framework/Model Selection Criteria for the Qualitative Measures	Does not meet the criteria	Meets the criteria	Exceeds the criteria	Reviewers' Notes
✓	Must contain minimally: a) Organizational and classroom management skills b) Ability to provide effective instruction c) Focus on continuous improvement and professional growth d) Interpersonal skills e) Leadership skills		✓		
	Accounts for years of service since teacher expertise develops over time	✓			
	Granular enough with "thin slices" of instruction to support deliberate practice because teachers develop expertise through engaging in focused practice with focused feedback	✓			Framework is the most broad of all with 20 elements
	Identifies the instructional context or lesson type or segment for when it is instructionally appropriate to see certain research-based strategies	✓			
	Includes scales or rubrics to identify the level of implementation for the strategies		✓		
	Reflects the elements for a research-based common language of instruction that accurately reflects the complexity of teaching		✓		
	Clearly defines and articulates teacher and student evidences for each rubric	✓			
	Validation studies (Specific studies on the model/framework to verify its ability to identify levels of teaching performance correlated to student achievement results)	✓			This question was raised during the presentation and the framework has not yet been validated
	Research studies verifying the specific classroom practices in the rubrics have a "causal link" to raising student achievement	✓			No research studies were submitted
	Weights/emphasizes classroom instruction		✓		
	Depth of supports provided for the framework (qualified capacity to support statewide efforts)	✓			Given only one district is currently implementing, there would likely be capacity issues trying to scale it statewide

http://www.marzanoevaluation.com/

Teacher and Leader Effectiveness Commission Teacher Evaluation Framework/Model Selection Criteria

The review following the criteria set forth by the Teacher and Leader Effectiveness Commission will be conducted by the Oklahoma State Department of Education and submitted to the Commission for consideration.

Selection	Description
Does not meet the criteria	From the materials provided, does not appear to meet the criteria
Meets the criteria	From the materials provided, appears to meet the criteria
Exceeds the criteria	From the materials provided, appears to exceed the criteria

Framework: Marzano Causal Teacher Evaluation Model

MARZANO CAUSAL - TEACHER

Required by Statute	Teacher Evaluation Framework/Model Selection Criteria for the Qualitative Measures	Does not meet the criteria	Meets the criteria	Exceeds the criteria	Reviewers' Notes
✓	Includes a Five-Tier Rating System (qualitative measures) 1) Superior 2) Highly Effective 3) Effective 4) Needs Improvement 5) Ineffective		✓		Presentation included calculation to generate Oklahoma's five-tier rating system and is currently in use within Oklahoma City as presented by Dr. Brian Staples
✓	Annual evaluations that provide feedback to improve student learning and outcomes			✓	In addition to the causal link research, the model also contains reflection questions, video examples, teacher and student evidences, etc. to provide teachers with annual evaluations that support their growth and development to raise student learning and outcomes.
✓	Comprehensive remediation plans and instructional coaching for all teachers rated as Needs Improvement or Ineffective			✓	All teachers are required to develop Professional Growth Plans and engage in deliberate practice in order to document improvements in their teaching. Processes include supports and tools for instructional coaches to engage with struggling teachers and supervision models for more support and observational feedback for struggling teachers.
✓	Be evidence-based (e.g. research and field proven)			✓	The Art and Science of Teaching, upon which the evaluation model was developed is widely used. The evaluation model is also being used in large scale including a pilot in Oklahoma City and the state of Florida.

Teacher and Leader Effectiveness Commission Teacher Evaluation Framework/Model Selection Criteria

MARZANO CAUSAL - TEACHER CON'T

Required by Statute	Teacher Evaluation Framework/Model Selection Criteria for the Qualitative Measures	Does not meet the criteria	Meets the criteria	Exceeds the criteria	Reviewers' Notes
✓	Include observable and measurable characteristics of personnel and classroom practices (rubrics and evidences for each rubric)			✓	The model contains rubrics, teacher and student evidence for each rubric, and coaching supports for each rubric.
✓	Be correlated to student performance success (validation studies and causal links studies for instructional strategies/behaviors)			✓	The model has substantial research documentation for raising student achievement through the 41 classroom strategies and behaviors in Domain 1.
✓	Be based on research-based national best practices and methodology (contemporary research and practices of expertise development and strategies/behaviors for raising student achievement)			✓	The model draws upon 35 years of research for what works for raising student achievement. The model also cites contemporary research for the development of expertise and incorporates national best practices for accounting for years of service and growth over time.
✓	Must contain minimally: a) Organizational and classroom management skills b) Ability to provide effective instruction c) Focus on continuous improvement and professional growth d) Interpersonal skills e) Leadership skills			✓	The model exceeds the minimum requirements. Domains 3-4 incorporate interpersonal skills and leadership skills.
	Accounts for years of service since teacher expertise develops over time			✓	Model includes for categories of teachers accounting for years of service with recommendations for 0-3, 3-9, and 10 plus years of service.
	Granular enough with "thin slices" of instruction to support deliberate practice because teachers develop expertise through engaging in focused practice with focused feedback			✓	Classroom strategies and behaviors (Domain 1) includes 41 elements that are granular enough to support deliberate practice.
	Identifies the instructional context or lesson type or segment for when it is instructionally appropriate to see certain research-based strategies			✓	41 elements in Domain 1: Classroom Strategies and Behaviors are classified into lesson types or segments for both teachers and observers to identify when it is appropriate to see certain strategies.
	Includes scales or rubrics to identify the level of implementation for the strategies		✓		All elements in the framework include 5 point scales/rubrics identifying levels of implementation of the strategies.
	Reflects the elements for a research-based common language of instruction that accurately reflects the complexity of teaching			✓	The complexity of teaching is represented in the model and the model reflects a substantial research base drawn from 35 years of research and meta-analysis.

Teacher and Leader Effectiveness Commission Teacher Evaluation Framework/Model Selection Criteria

MARZANO CAUSAL - TEACHER CON'T

Required by Statute	Teacher Evaluation Framework/Model Selection Criteria for the Qualitative Measures	Does not meet the criteria	Meets the criteria	Exceeds the criteria	Reviewers' Notes
	Clearly defines and articulates teacher and student evidences for each rubric			✓	Every scale/rubric includes examples of teacher and student evidence.
	Validation studies (Specific studies on the model/framework to verify its ability to identify levels of teaching performance correlated to student achievement results)		✓		Only model where validation studies conducted within Oklahoma were cited for the model
	Research studies verifying the specific classroom practices in the rubrics have a "causal link" to raising student achievement			✓	Over 300 individual experimental and control studies have been completed identifying the causal link for use of strategies cited in the model to increases in student learning.
	Weights/emphasizes classroom instruction			✓	Of the 60 total elements in the Marzano model, 41 or 68% represent classroom instruction.
	Depth of supports provided for the framework (qualified capacity to support statewide efforts)			✓	Given the statewide implementations currently underway with the model, there is both capacity to support Oklahoma districts and a depth of supports from trainings or certification for evaluators for accuracy for observations.

APPENDIX D

Teacher Evaluation Profile (TEP)

Section 1: Demographic Information

1. What is your gender?	1. Female 2. Male
2. Including the current year, how many years of teaching experience?	1. 1 year 2. 2 to 5 years 3. 6 to 10 years 4. 11 to 15 years 5. 16 or more years
3. What is your degree level?	1. Bachelors Degree 2. Masters Degree 3. Doctorate Degree
4. Did your school meet AYP in 2008?	1. Yes 2. No 3. Don't Know
5. Your current teaching assignment	1. Language Arts Literacy 2. Mathematics

Section 2: Overall Rating

Please reflect on your most recent experience with the evaluation process in your school district. Consider the entire evaluation process including goal-setting, meetings with evaluator, planning for evaluation, formal and informal observations, or other procedures and feedback.

1. Rate the overall quality of the evaluation:

Very poor quality 1 2 3 4 5 Very high quality

2. Rate the overall impact of the evaluation on your professional practices. (Note: A rating of 5 would reflect a strong impact leading to profound changes in your teaching practices, attitudes about teaching, and/or understanding of the teaching profession. A rating of 1 would reflect no impact at all and no changes in your practices, attitudes, and/or understanding.)

No impact 1 2 3 4 5 Strong impact

3. Rate the overall impact of the evaluation process on your professional growth as an educator. (Note: A rating of 5 would reflect a strong impact on your professional growth. A rating of 1 would reflect no impact at all on your professional growth.

No impact 1 2 3 4 5 Strong impact

Section 3: Rating Attributes of Evaluation

A. Describe yourself in relation to the following attributes:

4. The strength of your professional expectations of yourself

I demand little	1	2	3	4	5	I demand a great deal

5. Orientation to risk taking

I avoid risks	1	2	3	4	5	I take risks

6. Orientation to change

I am relatively slow to change	1	2	3	4	5	I am relatively flexible

7. Orientation to experimentation in your classroom

I don't experiment	1	2	3	4	5	I experiment frequently

8. Openness to criticism

I am relatively closed	1	2	3	4	5	I am relatively open

9. Knowledge of technical aspects of teaching

I know a little	1	2	3	4	5	I know a great deal

10. Knowledge of curriculum content

I know a little	1	2	3	4	5	I know a great deal

11. Experience with teacher evaluation prior to most recent experience

Waste of time	1	2	3	4	5	Very helpful

B. Describe your perceptions of the person who most recently evaluated your performance:

12. Credibility as a source of feedback

Not credible	1	2	3	4	5	Very credible

13. Working relationship with you

Adversary	1	2	3	4	5	Helper

14. Level of trust

Not trustworthy	1	2	3	4	5	Trustworthy

15. Interpersonal manner

Threatening	1	2	3	4	5	Not threatening

16. Temperament

Impatient	1	2	3	4	5	Patient

17. Flexibility

Rigid	1	2	3	4	5	Flexible

18. Knowledge of technical aspects of teaching

Not knowledgeable	1	2	3	4	5	Very knowledgeable

19. Capacity to model or demonstrate needed improvements

Low	1	2	3	4	5	High

20. Familiarity with your particular teaching assignment

Unfamiliar	1	2	3	4	5	Very familiar

21. Usefulness of suggestions for improvement

Useless	1	2	3	4	5	Very useful

22. Persuasiveness of rationale for suggestion

Not persuasive	1	2	3	4	5	Very persuasive

C. Describe the attributes of the procedures used during your most recent evaluation:

Standards are the criteria used to evaluate your teaching. Describe the procedures related to standards in the items below:

23. Were standards communicated to you?

Not at all	1	2	3	4	5	In great detail

24. Were the standards clear to you?

Vague	1	2	3	4	5	Very clear

25. Were the standards endorsed by you as appropriate for your teaching assignment?

Not endorsed	1	2	3	4	5	Highly endorsed

26. Were the standards…

The same for all teachers?	1	2	3	4	5	Tailored for your unique needs?

To what extent were the following sources of performance information considered as part of the evaluation?

27. Observation of your classroom performance

Not considered	1	2	3	4	5	Used extensively

28. Meetings with evaluator

Not considered	1	2	3	4	5	Used extensively

29. Examination of artifacts (lesson plans, materials, home/school communication)

Not considered	1	2	3	4	5	Used extensively

30. Examination of student performance

Not considered	1	2	3	4	5	Used extensively

31. Self evaluations

Not considered	1	2	3	4	5	Used extensively

Describe the extent of the observations of your classroom, based on your most recent evaluation experience in your school district. (Note: In these items, <u>formal</u> refers to observations that were pre-announced and/or were accompanied by a pre- or post-conference with the evaluator; <u>informal</u> refers to unannounced drop-in visits.)

32. Number of <u>formal</u> observations per year

1. 0 Observations
2. 1 Observation
3. 2 Observations
4. 3 Observations
5. 4 Observations

33. Approximate frequency of informal observations per year

1. None
2. Less than 1 per month
3. Once per month
4. Once per week
5. Daily

34. Average length of formal observations

Brief (few minutes)	1	2	3	4	5	Extended (40 minutes or more)

35. Average length of informal observations

Brief (few minutes)	1	2	3	4	5	Extended (40 minutes or more)

D. Please describe the attributes of the feedback you received during your last evaluation experience:

36. Amount of information received

None	1	2	3	4	5	Great deal

37. Frequency of formal feedback

Infrequent	1	2	3	4	5	Frequent

38. Frequency of informal feedback

Infrequent	1	2	3	4	5	Frequent

39. Depth of information provided

Shallow	1	2	3	4	5	In-depth

40. Quality of the ideas and suggestions contained in the feedback

Low	1	2	3	4	5	High

41. Specificity of information provided

General	1	2	3	4	5	Specific

42. Nature of information provided

Judgmental	1	2	3	4	5	Descriptive

43. Timing of feedback

Delayed	1	2	3	4	5	Immediate

44. Feedback focused on

Ignored the standards	1	2	3	4	5	Reflected the teaching standards

E. Please describe these attributes of the evaluation context:

Resources available for evaluation:

45. Amount of time spent on the evaluation process, including your time and that of all other participants.

None	1	2	3	4	5	Great deal

46. Time allotted during the semester for professional development

None	1	2	3	4	5	Great deal

47. Availability of training programs and models of good practices

None	1	2	3	4	5	Great deal

District Values and policies in evaluation:

48. Clarity of policy statements regarding purpose of evaluation

Vague	1	2	3	4	5	Very clear

49. Intended role of evaluation

Teacher accountability	1	2	3	4	5	Teacher growth

APPENDIX E

Oklahoma Criteria for Effective Teachers Performance Evaluation

OKLAHOMA CRITERIA
for
EFFECTIVE TEACHING
and
ADMINISTRATIVE PERFORMANCE



Sandy Garrett
State Superintendent of Public Instruction
Oklahoma State Department of Education

Revised

CRITERIA FOR EFFECTIVE TEACHING PERFORMANCE

I. Practice

A. Teacher Management Indicators

1. Preparation
2. Routine
3. Discipline
4. Learning Environment

B. Teacher Instructional Indicators

1. Establishes Objectives
2. Stresses Sequence
3. Relates Objectives
4. Involves All Learners
5. Explains Content
6. Explains Directions
7. Models
8. Monitors
9. Adjusts Based on Monitoring
10. Guides Practice
11. Provides for Independent Practice
12. Establishes Closure

II. Products

A. Teacher Product Indicators

1. Lesson Plans
2. Student Files
3. Grading Patterns

B. Student Achievement Indicators

When the term objectives is used it refers to the mandated Oklahoma academic content standards, *Priority Academic Student Skills (PASS). PASS* may be found on the State Department of Education Web site <http://www.sde.state.ok.us>.

APPENDIX F

Superintendent letter of Approval

RE: Permission to Conduct Research Study

Dear Superintendent:

I am writing to request permission to distribute a questionnaire within your district. Currently a doctoral candidate at the University of Oklahoma, Norman Campus in the Department of Educational Leadership and Policy Studies Executive, Ed.D program, I am engaged in the process of writing my dissertation, which is titled, *Oklahoma Teachers' Perceptions Towards the Qualitative Portion of the New Teacher Evaluation System and its Implementation*. The study will be offered to all teachers in Oklahoma. It will utilize a questionnaire, allowing teachers to share their perceptions regarding the new teacher evaluation system. If approval for your district is granted, teachers will receive an email discussing the research purpose and an attached link to the questionnaire. Once teacher consent is conceded, the survey should take approximately fifteen minutes. The survey results will be pooled for the dissertation project; individual results of this study will remain absolutely confidential and anonymous. Should this study be published, only pooled results will be documented. No costs will be incurred by your district or the individual participants.

Your approval is greatly appreciated. If you choose not to participate, please click on the link below and select no. If you allow your district to participate, there are a few required procedures for teachers to receive the questionnaire. The link below is to a survey program called Qualtrics. Before you click on the link, I will need you to have all required documentation ready for upload on your computer. First, print this permission letter and sign below representing your approval and copy the form on your schools letterhead to confirm consent. Click on the link and select yes, you will upload the letterhead as well as your

district's teacher roster so Qualtrics can disseminate the questionnaire to the teachers. If you have questions or concerns, I am available to correspond with you at Ivana.A.Beatty@ou.edu.

Sincerely,

Ivana A. Beatty

Approved by:

_____ _____ _____

Print your name and title here Signature Date

Survey: https://ousurvey.qualtrics.com/SE/?SID=SV_6nbHstkraR6Wp01

Default Question Block

Options:

Yes, my district will participate in the study
No, my district will not participate in the study.

Please upload thepermssion letter with your signed information on school letterhead

Choose File No file chosen

Please upload your teacher roaster.

Choose File No file chosen

Powered by Qualtrics

The University of Oklahoma is an equal opportunity institution.

APPENDIX G

Teacher Recruitment letter

Hello:

My name is Ivana Beatty and I am a doctorate student at the University of Oklahoma. I am conducting a statewide research study on teachers' perceptions towards the new Oklahoma teacher evaluation framework and the effectiveness of its implementation. I am emailing to ask if you would take about 15 minutes to complete a questionnaire for this research project. Participation is completely voluntary and your answers will be anonymous.

If you have any questions, please do not hesitate to contact me (ivana.a.beatty@ou.edu) or my advisor, Dr. Hollie Mackey (hmackey@ou.edu). You can also contact the University of Oklahoma – Norman Campus Institutional Review Board (OU-NC IRB) at 405-325-8110 or irb@ou.edu (IRB number is 5498) if you have questions about your rights as a research participant and wish to talk to someone other than the researcher.

Please click on the link below to participate.

Thank you for your time,

Ivana Beatty

Doctorate Student

University of Oklahoma

APPENDIX H

Doherty's Teacher Evaluation Profile (TEP) for Teachers

Doherty's Teacher Evaluation Profile for Teachers

Item Number	Item description
Section 2: Overall rating Please reflect on your most recent experience with the evaluation process in your school. Consider the entire evaluation process including goal setting, self- assessment, meetings with your evaluator, planning for evaluation, formal and informal observations, or other procedures and feedback.	
Item 1	Rate the overall quality of the evaluation process (between very poor quality and very high quality)
Item 2	Rate the overall impact of the evaluation process on your professional practices (Note: a rating of 5 would reflect a strong impact leading to profound changes in your teaching practices, attitudes about teaching, and/or understanding of the teaching profession. A rating of 1 would reflect no impact at all and not changes in your practices, attitudes, and/or understandings.) (between no impact and strong impact)
Item 3	Rate the overall impact of the impact of the evaluation process your professional growth as an educator. (Note: a rating of 5 would reflect a strong impact in your professional growth. A rating of 1 would reflect no impact at all in your professional growth.) (between no impact and strong impact)
Next, please rate your perception of the impact of the teacher evaluation process on the school, district and the state goals. Use the scales provided to indicate impact, from 1 meaning no impact to 5 meaning strong impact.	
Item 4	Rate the positive impact on school *student learning*: A strong impact rating (5) would indicate that the evaluation system improves the quality of student learning.
Item 5	Rate the positive impact on *student achievement*: A strong impact rating (5) would indicate that the evaluation system improves student performance on standardized test.

Item 6	Rate the positive impact on *school improvement goals*: A strong impact rating (5) would indicate that the evaluation system helps the faculty achieve school improvement goals.
Item 7	Rate the positive impact on *school climate and culture*: A strong impact rating (5) would indicate that the evaluation system supports and helps foster a positive school culture and climate that supports learning.
Item 8	Rate the positive impact on *quality of teachers:* A strong impact rating (5) would indicate that the evaluation system improves teaching quality.
Item 9	Rate the positive impact on your *goals* that you develop each year. A strong impact rating (5) would indicate that the evaluation system supports and links to the development of your goals.

Section 3: Rating Attributes of Evaluation

Please use the scales provided below (1 through 5) to describe yourself and the nature of your most recent teacher evaluation experience. Do this by:

- Considering the attribute to be described
- Studying the scale to be used to describe it
- Selecting the number that represents the point you select on each continuum
- Marking the answer sheet accordingly

Part A- Describe yourself in relation to the following attributes:

Item 10	Your overall performance on the Teacher Assessment Process (TAP) (between does not meet standard to exceeding the standard)
Item 11	The strength of your professional expectation of your yourself (between I demand little to I demand a great deal)
Item 12	Orientation to risk taking (between I avoid risk to I take risk)
Item 13	Orientation to change (between I am relatively slow to change to I am relatively flexible)
Item 14	Orientation to experimentation in your classroom (between I don't experiment to I experiment frequently)
Item 15	Open to criticism (between I am relatively closed to I am relatively open)
Item 16	Knowledge of technical aspects of teaching (between I know a little to I know a great deal)
Item 17	Knowledge of curriculum content for what you teach (between I know a little to I know a great deal)

Item 18	Experience with teacher evaluation prior to most recent experience (between waste of time and very helpful)
Part B- Describe your perceptions of the person who most recently evaluated your performance:	
Item 19	Credibility as a source of feedback (between not credible and very credible)
Item 20	Working relationship with you (between adversary and helper)
Item 21	Level of trust (between not trustworthy to trustworthy)
Item 22	Interpersonal manner (between threatening to not threatening)
Item 23	Temperament (between impatient to patient)
Item 24	Flexibility (between rigid to flexibility)
Item 25	Knowledge of technical aspects of teaching (between not knowledgeable to very knowledgeable)
Item 26	Capacity to model or demonstrate needed improvements (between low and high)
Item 27	Familiarity with your particular teaching assignment (between unfamiliar to very familiar)
Item 28	Usefulness of suggestions for improvement (between useless to very useful)
Item 29	Persuasiveness of rational for suggestions (between not persuasive to strong impact)
Part C- Describe the attributes of the procedures used during your most recent evaluation: Standards are the criteria used in the TAP process to evaluate your teaching. Describe the procedures related standards in the items below:	
Item 30	Were standards communicated to you? (between not at all to in great detail)
Item 31	Were the standards clear to you (between Vague to very clear)
Item 32	Were the standards endorsed by you as appropriate for our teaching assignment (between not endorsed to highly endorsed)

Item 33	Were the standards… (between the same for all teachers to tailored for your unique needs)
To what extent were the following sources of performance information considered as part of the evaluation?	
Item 34	Observation of your classroom performance (between not considered to used extensively)
Item 35	Meetings with evaluator (between not considered to used extensively)
Item 36	Examination of artifacts (lessons plans, materials, home/school communication, etc.) (between not considered to extensively)
Item 37	Examination of student performance (between not considered to used extensively)
Item 38	Student evaluations (between not considered to used extensively)
Item 39	Peer evaluations (between not considered to used extensively)
Item 40	Self- evaluations (between not considered to used extensively)
Describe the extent of the observations of your classroom, based on your most recent evaluation experience. (Note: in these items, formal refers to observations that were pre-announced and/or were accompanied by a pre- or post- conference with the evaluator; informal refers to unannounced drop-in visits.)	
Item 41	Number of formal observations per year (between 0 to 4 or more observations)
Item 42	Approximate frequency of informal observations (most recent experience) (choices are none, less than 1 per month, once per month, once per week, and daily)
Item 43	Average length of FORMAL observation (most recent experience) (between brief- a few minutes to extend- 40 minutes or more)
Item 44	Average length of INFORMAL observation (most recent experience) (between brief- a few minutes to extend- 40 minutes or more)
Part D- Please describe the attributes of the feedback you received during your last evaluation experience:	
Item 45	Amount of information received (between none to great deal)
Item 46	Frequency of formal feedback (between infrequent to frequent)
Item 47	Frequency of informal feedback (between infrequent to frequent)

Item 48	Depth of information provided (between shallow to in depth)
Item 49	Quality of the ideas and suggestions contained in the feedback (between low and high)
Item 50	Specificity of information provided (between general to specific)
Item 51	Nature of information provided (between judgmental to descriptive)
Item 52	Timing of feedback (between delayed to immediate)
Item 53	Feedback focused on the TAP standards (between ignored the TAP standards to reflected the TAP standards)
Part E- Please describe these attributes of the evaluation context: Resources available for evaluation	
Item 54	Amount of time spend on the evaluation process, including your time and that of all other participants (between none to great deal)
Item 55	Time allotted during the school year for professional development aligned with standards (between none to great deal)
Item 56	Availability of training programs and models of good practices (between none to great deal)
Item 57	Clarity of policy statements regarding the purpose of evaluation (between vague and very clear)
Item 58	Intended role of evaluation (between teacher accountability and teacher growth)