UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE PERCEIVED QUALITY AND IMPACT OF THE TULSA MODEL EVALUATION: A SURVEY OF TEACHERS AND PRINCIPALS IN OKLAHOMA PUBLIC SCHOOLS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF EDUCATION

By

MELONIE R. HAU
Norman, Oklahoma
2020

THE PERCEIVED QUALITY AND IMPACT OF THE TULSA MODEL EVALUATION: A SURVEY OF TEACHERS AND PRINCIPALS IN OKLAHOMA PUBLIC SCHOOLS

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

BY THE COMMITTEE CONSISTING OF

Dr. John R. Jones, Chair

Dr. Curt Adams

Dr. Jeffrey Maiden

Dr. Neil Houser

ACKNOWLEDGEMENTS

A study gathering the perceptions of hundreds of teachers and principals across the state of Oklahoma would not be possible without the help of multiple talented colleagues.

Dr. Beulah Adigun provided critical assistance in analyses of data, and I am grateful for her insight into tackling a large sample via an alternative ANOVA to produce valid results.

I am thankful to the tight knit Oklahoma educator community for their participation in the survey. Our state has endured many accountability reform challenges in the last ten years. However, we continue to find ourselves banding together to make our voices heard. We understand the importance of collaboration, using a combination of intelligence, common sense, and that famous Oklahoma wit to build better lives for children.

Dr. John Jones and members of my committee have assisted me and given me confidence throughout this journey. I greatly respect their dedication to educating future school leaders and their passion for improving education policy. Dr. Jones, especially, provided much needed encouragement just when I doubted I could complete this project.

And to my wonderful family, I am grateful for your patience and resilience as I took time away from family gatherings and weekend visits. I love you dearly, Byron, Trent, and Jesse.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Oklahoma began implementing its Teacher Leader Effectiveness Law in 2011, the state's response to competing for federal Race to the Top grant money and the nation-wide cry from education reformers for tougher teacher evaluation mandates in public schools. This quantitative analysis of variance study explores the effectiveness of the law's implementation in Oklahoma school districts using the Tulsa Model Evaluation process in the fall of 2019. A guiding theoretical framework for this study is based in the research of Stiggins and Nickel (1989) and Stiggins and Duke (1988) and their findings of the usefulness of the Teacher Evaluation Profile to measure conditions for teacher growth using standards-based evaluation instruments. A cross-sectional survey distributed to principals and teachers gathered perceptions of the quality and impact of the Tulsa Model Evaluation process based on six key areas: knowledge and understanding of the evaluation system; perceptions regarding usefulness; qualifications and evaluator leadership; quality of data and feedback from principals; impact on professional practices; and personal impressions of the evaluation system on professional conversations. The study uses a Welch alternative ANOVA to determine the differences in perceived quality and impact of the Tulsa Model evaluation instrument between principals and teachers in 54 Oklahoma schools. The following research questions were addressed:

1. Do principals and teachers differ in their perceptions of knowledge and understanding of the Tulsa Model Evaluation process and if so, to what degree?

2. Do principals and teachers differ in their perceptions of the usefulness of the Tulsa Model Evaluation process and if so, to what degree?
3. Do principals and teachers differ in their perceptions of the qualification of the evaluator and if so, to what degree?
4. Do principals and teachers differ in their perceptions of the quality of data and feedback in the Tulsa Model Evaluation process and if so, to what degree?
5. Do principals and teachers differ in their perceptions of the impact on professional practices for the Tulsa Model Evaluation process and if so, to what degree?
6. Do principals and teachers differ in their perceptions of the context of professional conversations during the Tulsa Model Evaluation process and if so, to what degree?

Results determined there are significant differences between principal and teacher responses to survey questions with the greatest degree of difference in three key areas: 1. knowledge and understanding of the evaluation; 2. usefulness of the process; and 3. quality of data and feedback.

Keywords: standards-based teacher evaluation, teacher evaluation profile, Race to the Top, Teacher Leader Effectiveness Law, Tulsa Model Evaluation

CHAPTER 1

INTRODUCTION

Background of the Study

During the time of sweeping education reform across the country, states like Oklahoma scrambled to adopt models for teacher and leader evaluations in competition for federal dollars promised by Race to the Top grants (Viteritti, 2012; U. S. Department of Education, 2009). Although the federal government has moved passed the days of Race to the Top and the associated waivers which released states from the mandates of No Child Left Behind, many states are left with codified legislative mandates passed during the education reform fervor. The intention of these waivers was to release states from the accountability mandates of the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), also known as No Child Left Behind (NCLB). However, the true policy intent of waivers was to drive a specific education reform agenda, which included reforming teacher and leader evaluation (Viteritti, 2012; Kane, Kerr, & Pianta, 2014). The impact of incentivizing frameworks for teacher and leader evaluations on a federal level was a rush by state officials to adopt policy without regard to research-based practices.

The Teacher Leader Effectiveness Law was implemented in Oklahoma beginning in 2011 as a way for the state to compete for Race to the Top grants and it included Value Added Measures to rank teachers according to student achievement gains. State mandated implementation did not call for educating principals and teachers how to use the Tulsa Model Evaluation instrument within the context of a systems or a learning organization

approach to encourage teacher conditions for growth. As such, principals often use the evaluation process as part of a managerial task list aligned with an outdated organizational culture in which principals, due to time and stress, will choose to maintain climate over enforcing ineffective policy mandates (Hallinger, Heck, & Murphy, 2014). Principals would rather retain teachers rather than alienate them with evaluation instruments that feel burdensome. Without proper implementation, the mandate demoralizes the profession and wastes time and resources and forces us to consider its usefulness.

When reflecting on the history of the past twenty years of school reform, it seems natural to ask questions about the success of these standards-based evaluation frameworks, especially the ones implemented in schools through state and federal mandates. Is it possible that mandated structures improve student learning? Under what conditions do forced structures produce genuine motivation for change? And after implementation of these reforms, especially those focused on teacher effectiveness, do we believe teachers can grow and improve the quality of their practice? Do these mandated standards-based evaluation systems create conditions for teacher growth, which in turn leads to school wide improvement and higher student achievement? If so, how do mandates on teacher evaluation assist that process and what role does it play in the larger cultural context of engagement, professional growth, and school improvement? How does the understanding of the purpose the evaluation process between principal and teacher impact professional outcomes? What can we learn about the motivation of teachers, principals, and schools to change practices to improve student learning through the

Oklahoma Teacher Leader Effectiveness mandate, especially as it relates to the Tulsa Model Evaluation process?

Using the foundation of the Teacher Evaluation Profile research by Stiggins and Nickel (1989) and Stiggins and Duke (1988), the purpose of this study was to measure the quality and impact of the Tulsa Model Evaluation process as perceived by teachers and principal evaluators in Oklahoma. Cross-sectional survey data from schools across the state were collected and analyzed in two phases utilizing demographic statistics and a Welch ANOVA alternative. The variables for the study included perceived teacher and principal understanding of the quality and impact of the Tulsa Model Evaluation process and its procedures surrounding the evaluation experience, attributes of the evaluator, quality of feedback, and the evaluation context.

The overarching question for this study was do teachers and principals differ in their perceptions of the quality and impact of the Tulsa Model Evaluation process based on Teacher Evaluation Profile conditions and previous empirical research establishing conditions for teacher professional growth?

Problem of Practice:

The Unsatisfactory Results of State Mandated School Reform

Today, districts find themselves living with competing education reform movements: implementing standards-based accountability reforms while attempting the collaborative learning cultures championed by the effective schools movement. The state mandated tools used to blend the effective schools movement and the accountability reform movement have not been studied fully to determine their effectiveness on the

teacher's or the school leader's ability to improve student learning. In Oklahoma, one such state mandate is the Teacher Leader Effectiveness Law.

After President Obama announced Race to the Top grants in 2009, states had one year to propose reforms. Race to the Top grants required states to overhaul, or implement in cases where none existed, teacher evaluation systems. Oklahoma did compete for Race to the Top funds, and although the grant request produced in 2010 was not accepted by the United States Department of Education, the reforms generated through the application process became part of Oklahoma's waiver to release the state from the mandates of NCLB (Oklahoma State Department of Education, 2014). One mandate specifically targeted by the waiver request was teacher evaluation. The Oklahoma Legislature codified teacher evaluation as part of law in 2010 (Oklahoma State Courts Network, 2019) and The Oklahoma Teacher Leader Effectiveness Law came into effect in 2011 (Oklahoma State Court Network, 2015). The mandate required each school to choose a teacher evaluation framework. Most districts chose the Tulsa Model Evaluation process based on Danielson's *Framework for Teaching* (The Danielson Group, 2013; The Oklahoma State Department of Education, 2014).

The Oklahoma State Department of Education developed a theory of action to explain implementation to educators in 2011. The theory of action relied on the assumption that teacher effectiveness is the most critical factor in improving student achievement. It also held the state could develop a system of evaluation that would

1. assess a teacher's "strengths and weakness,"
2. "provide high-quality professional development,"
3. support districts in providing "professional learning opportunities," and "seek

ongoing feedback to improve the system and professional development opportunities" (Oklahoma State Department of Education, 2019).

Although Oklahoma teacher evaluation was part of law and best practice before the state's attempt at a Race to the Top grant, the question in the minds of education reformers remain: Do teacher evaluations create conditions for teacher professional growth? Nine years after the Teacher Leader Effectiveness Law began to be implemented, we have no evidence the Tulsa Model Evaluation process produces the outcomes promised for the large number of schools using the model across the state of Oklahoma.

Teacher evaluation as a form of current education reform is focused on accountability. Results have been less than convincing that mandated teacher evaluation efforts make a difference in school improvement or that they increase the proficiency of the profession (Dynarski, 2016; Borko, Liston, & Whitcomb, 2009; Weisberg, et al., 2009). Because reforms have been strongly focused on mechanisms associated with organizational management rather than creating the systems and learning organizations to support them, results have not delivered on NCLB or Race to the Top intentions (Hallinger & Heck, 2014). The problems with state mandated standards-based teacher evaluation tools are many including

1. Lack of understanding of the purpose of the evaluation or ineffective principal training and implementation. Principals do not receive adequate professional development for building collaborative learning cultures in which feedback is a form of coaching. Principals do not see how an evaluation tool fits into a whole system of school improvement or part of a learning organization. In

addition, principals may lack the skill to properly understand what should be recorded as effective practice (Bryk & Schneider, 2002; Connally & Tooley, 2016; Darling-Hammond, et al., 2012; Fullan, 2009; Hill & Grossman, 2013; Hallinger & Heck, 2014; Hallinger, Heck, & Murphy, 2014; Honig, et al., 2010).

2. Lack of teacher input or feedback. Teachers should be treated as learners and leaders in the system (Darling-Hammond et al., 2012). Without attention to teachers as adult learners who need feedback, mental models, and reflection, effective strategies will not bridge the knowing/doing gap. When teachers do not apply what they have learned about effective practice, student achievement will suffer (Hattie, 2015a; Stronge & Tucker, 1999; Taylor & Tyler, 2012; James & McCormick, 2009).
3. Lack of connection to professional development and professional community. Teacher evaluation instruments work best as a tool within schools as learning organizations. When principals and teachers use evaluation to provide focused conversations about specific strategies, student achievement results are likely to increase (Reeves, 2010; Robinson, Lloyd, & Rowe, 2008; Taylor & Tyler, 2012; Goe, Biggers, & Croft, 2012).
4. Limited range and authenticity in evaluation scores. Authenticity can be defined as validity or legitimacy. An evaluation of scores from several evaluation tools ranks most teachers a 3 on a five-point scale. Most systems use a Likert scale rubric based on standards of effective teaching practice. Principals tend to choose mid-range scores resulting in overall ratings of

> "effective" or "highly effective" either because they believe the rating scale to be too blunt an instrument to measure effective practice or they believe teachers should not be rated poorly if proper supports are not given for improvement (Borko, Liston, & Whitcomb, 2009; Connally & Tooley, 2016; Dynarski, 2016; Taylor & Tyler, 2012; Weisberg, et al., 2009). The lack of validity is especially problematic for schools attempting to use the evaluation process as a vehicle to measure effective practice as it is linked to school improvement. Although a regulation can become internalized, leading to sustained change, if the regulation is seen to have value (Gagne & Deci, 2005), one way to add value is to provide an environment for proper feedback. If evaluation ratings consistently yield a 3 rating, the value or usefulness of the tool comes into question.

There is evidence that a particular evaluation instrument, used in isolation, has little effect on student learning or achievement (Darling-Hammond et al., 2012). However, a body of evidence exists suggesting feedback provided to teachers within a context of a learning community or supportive environment can have a positive impact on student learning (Hattie, 2003; Tucker & Stronge, 2005; Taylor & Tyler, 2012). A work environment which helps its employees achieve feelings of competence, autonomy, and relatedness will be more successful in implementing a regulation (Gagne and Deci, 2005). Feedback and a principal's ability to provide a larger, more purposeful context for the teacher evaluation process can create feelings of competence and relatedness needed to turn the evaluation from a mandate to more internalized motivation to improve teacher growth.

A growing segment of inquiry connects principal leadership to teacher evaluation and feedback, which ultimately impacts student learning and achievement results. A study conducted in Chicago Public Schools creates a pathway from principal leadership to student outcomes on the ACT Education Planning and Assessment System (Sebastian & Allensworth, 2012). Through multilevel equation modeling and analysis of teacher survey results, researchers created a link from principal actions to student achievement. The results found that although a principal's direct impact on student learning was small, the overall school climate and learning environment created by the principal made the biggest difference for student achievement.

In addition to the Chicago study, researchers in Cincinnati Public Schools found qualitative data collected via classroom observation tools may also have positive influence on student achievement (Sebastian & Allensworth, 2012). Researchers in this study limited their data to the established Cincinnati Teacher Evaluation System instrument, which is based on Charlotte Danielson's framework of effective teaching (Taylor & Tyler, 2012). In a comparative analysis of teacher evaluation scores and student achievement scores over time, the results suggest the act or the process of evaluation had a positive impact on student achievement, especially in math. Taylor and Tyler (2012) also uncovered evidence suggesting critical feedback to teachers improved student performance. The Cincinnati study makes a case that an evaluation instrument, if used to provide specific feedback to teachers, can make a difference for student learning.

Oklahoma law requires the qualitative component of the Teacher Leader Effectiveness standards-based evaluation processes to be based in evidence. The framers of the policy acknowledged the work of Charlotte Danielson and Robert Marzano in the

adoption of the evaluation instruments (Oklahoma State Courts Network, 2015). The intention of the qualitative evaluation process is to provide teachers with specific feedback to improve practice, and therefore, improve student learning (Hallinger & Heck, 2014). Research suggests teachers need feedback on effective strategies to grow as professionals (Darling-Hammond, et al., 2012; Dufour & Marzano, 2011; Murphy, Hallinger, & Heck, 2013; Hattie, 2002). The law supports and codifies the regular and timely practice for meaningful feedback for teachers and leaders, and as such, initiates radical change in school cultures which attempt to support deeper, authentic student learning. The qualitative nature of the Tulsa Model Evaluation process provides a strong mental model for educators because it connects a focus on instruction to the specific skills needed to develop an effective teacher or leader.

However, after five years of full implementation in Oklahoma from 2014 to 2019, debate persists about the quality and impact of standards-based quantitative teacher evaluation tools on professional growth, student achievement, or school improvement. The lack of agreement on what constitutes successful implementation of teacher and leader evaluation instruments in Oklahoma schools aligns with studies nationwide. In fact, within most local school contexts, the lack of correlation between the use of evaluation processes and student achievement is persistent. In addition, the failure of mandated policy to transform principal and teacher evaluation conversations into conditions for professional growth is evident in teacher and principal perceptions. No statewide studies have been conducted on the quality and impact of the Tulsa Model Evaluation process to determine if schools should continue using accountability driven reforms developed during the era of Race to the Top.

Previous dissertation studies measuring the perceived effect of standards-based evaluation tools have been completed in Iowa, Wisconsin, and New Jersey. A study by Huckstadt (2011) collected quantitative survey data from administrators and teachers in Iowa. Findings in this study indicate both administrators and teachers achieved an understanding of the evaluation tool, but they ranked the quality of feedback and the impact on effective teaching low. Overall, administrators ranked the impact of the evaluation tool higher than teachers. Teachers saw less connection between evaluation feedback conversations and their teaching practices. The summative evaluation seemed to be the most impactful event in the evaluation process for administrators and teachers. Among the recommendations from Huckstadt's (2011) study was a suggestion for training administrators and teachers in coaching conversation skills and in self-evaluation skills to build agency.

The Schumacher (2004) study used a mixed methods approach to collect teacher and administrators perceptions of the implementation of a standards-based teacher evaluation tool in one school district in Wisconsin. The results from this study found teachers generated mixed responses when comparing survey data to opinions given during personal interviews. Teachers understood the evaluation tool and believed indicators used in a standards-based evaluation improved teacher practice. When interviewed, however, teachers did not believe the evaluation process improved their practice. The results of this study hinted there is a void between the intent of the evaluation tool and its implementation to improve professional growth.

A study conducted by Towe (2012) in urban schools in New Jersey utilized a mixed methods approach to determine the impact of a standards-based evaluation model

on teacher effectiveness. Results determined perceptions between principals and teachers were similar indicating the evaluation process had little effect on improving teacher practice; however, the summative evaluation did have some impact on targeting professional development for teachers. Findings suggested the best use of the evaluation process is through a collaborative, coaching conversation in which a common language of effective practice is established. Recommendations from this study included measuring specific training needs of teachers and administrators to drive a stronger "growth-oriented" teacher evaluation process.

The theory of action used in recent studies promotes whole system approaches in implementing accountability movement reforms (Bryk, et al., 2010; Honig, 2008; Stronge & Tucker, 1999; Tucker & Stronge, 2005). In contrast, results of mandated accountability have been iffy at best (Dufour & Fullan, 2012; Taylor & Tyler, 2012; Darling-Hammond, 2013; Connally & Tooley, 2016; Weisberg, et al., 2009).

Critics of accountability driven teacher evaluation bemoan the data which indicate everyone receives a three on a five-point rating scale (Weisberg, et al., 2009) and teacher performance on evaluations have not clearly been connected to student achievement gains (Hallinger & Heck, 2014). Because reforms have been strongly focused on mechanisms associated with organizational management rather than learning organization functions, results have not delivered on NCLB or Race to the Top intentions.

Guiding Theoretical Framework

The Oklahoma Teacher Leader Effectiveness mandates were implemented based on a theory of action not grounded in educational learning theory; however, an assumption was that under the right conditions, these mandates would produce expected

results (Oklahoma State Department of Education, 2019). Documents used during the initial implementation phases of the Teacher Leader Effectiveness Law in Oklahoma promoted teacher evaluation as the key to effective teaching (Oklahoma State Department of Education, 2019). The law itself was codified based on the assumption that states must implement standard teacher evaluation instruments to ensure every student had access to the most effective teacher. At the time, the promulgated argument asserted a standardized process for teacher evaluation would improve student achievement in the state (Oklahoma Courts Network, 2015). As a critical component of the law, the Tulsa Model Evaluation and its proper use was assumed to positively correlate to teacher effectiveness, and by extension, improved student learning.

Because the standards-based evaluation processes were initially implemented with little consideration to the context and dynamics of real-world school environments, Oklahoma educators did speak out against the mandate. In response, the Oklahoma State Department of Education invited the Southern Regional Education Board to conduct a focus group to gather educator feedback (Oklahoma State Department of Education, 2020). Key findings from the Southern Regional Education Board report in 2014 indicate from the beginning of its implementation, teachers lacked understanding and trust in the standards-based evaluation mandate. Educators did not believe the mandate was designed to provide valuable or valid feedback about professional practice or would allow for ongoing feedback from schools about the effectiveness of the Teacher Leader Effectiveness Law.

To understand the environment in which the Tulsa Model Evaluation should be implemented one must consider the school as a learning organization and a system in which conditions are constantly changing.

Learning Organizations and Systems Thinking Theory

Evans, Thornton, and Usinger (2012) establish Peter Senge's learning organization approach as a theoretical framework for school improvement. The principal evaluator and the teacher, within school contexts, establish structures and conversations to adapt to changing conditions in a school environment. For actors within the system to be successful, elements of personal mastery, mental models, building a shared vision, and team learning must exist. Known largely as systems thinking, principals and teachers should understand their role in engaging each other in conversations to promote professional growth and school improvement.

Conditions for Teacher Growth

A systems thinking model informs what empirical research reveals about ideal conditions for teacher growth. In their meta-analysis of teacher evaluation and school improvement, Hallinger, Heck, and Murphy (2013) establish evaluations work best when four conditions are met. These conditions are

1. Providing actionable feedback to teachers
2. Creating professional communities
3. Support for teacher work
4. Opportunities for ongoing professional learning

Teacher Evaluation Profile

Finally, using the building blocks from systems thinking to conditions for teacher growth in an evaluation process, the core elements of the Teacher Evaluation Profile (TEP) designed by Stiggins and Nickel (1989) and discussed in studies conducted by Stiggins and Duke (1988) are a critical component to the guiding theoretical framework of this study. After reliability and validations studies, the TEP provides 44 survey items to correlate teacher perceived quality and impact experiences in an evaluation environment. The TEP has been used in previous dissertations in Iowa, Wisconsin, and New Jersey to establish baseline data on perceived teacher and administrator experiences within a standards-based evaluation process.

For my study I employ a modified version of the TEP, like the survey used in Huckstadt (2011), to gather baseline data on the perceived quality and impact of the Tulsa Model Evaluation from teachers and principals across the state of Oklahoma. Alignment of the problem of practice, to the TEP, research questions, and the cross-sectional survey are shown in Table 1.1.

Table 1.1

Guiding Theoretical Framework Connected to Problem of Practice and Research Study

Problem of Practice	Teacher Evaluation Profile (Stiggins and Nickel, 1989)	Research Question	Teacher and Principal Survey Section
Lack of understanding of procedures or purpose of implementation		1. Do principals and teachers differ in their perceptions of knowledge and understanding of the Tulsa Model Evaluation process and if so, to what degree?	Section A: Knowledge and Understanding of the Evaluation System
	Teachers asked to rate the overall quality of the evaluation experience.	2. Do principals and teachers differ in their perceptions of the quality of the Tulsa Model Evaluation process and if so, to what degree?	Section B: Perceptions Regarding Quality
Lack of teacher input or feedback	Teachers describe the evaluator based on certain attributes.	3. Do principals and teachers differ in their perceptions of the qualification of the evaluator and if so, to what degree?	Section C: Qualifications and Evaluator Leadership
	Teachers asked to describe the nature of feedback.	4. Do principals and teachers differ in their perceptions of the quality of data and feedback in the Tulsa Model Evaluation process and if so, to what degree?	Section D: Data and Feedback
Lack of connection to professional development and professional community	Teachers asked to describe procedures during the evaluation.	5. Do principals and teachers differ in their perceptions of the impact on professional practices for the Tulsa Model Evaluation process and if so, to what degree?	Section E: Impact on Professional Practice
Lack of understanding for authenticity of the evaluation instrument	Teachers asked to describe the context in which the evaluation took place.	6. Do principals and teachers differ in their perceptions of the context of professional conversations during the Tulsa Model Evaluation process and if so, to what degree?	Section F: Personal Impressions of the Tulsa Model Evaluation System

Research Questions

The following research questions were developed based on survey responses gathered from principals and teachers in school districts across the state of Oklahoma who utilize the Tulsa Model Evaluation instrument. An analyses of these questions explored the perceptions of principals and teachers to determine the quality and impact of the evaluation process to create conditions for teacher growth.

1. Do principals and teachers differ in their perceptions of knowledge and understanding of the Tulsa Model Evaluation process and if so, to what degree?
2. Do principals and teachers differ in their perceptions of the usefulness of the Tulsa Model Evaluation process and if so, to what degree?
3. Do principals and teachers differ in their perceptions of the qualification of the evaluator and if so, to what degree?
4. Do principals and teachers differ in their perceptions of the quality of data and feedback in the Tulsa Model Evaluation process and if so, to what degree?
5. Do principals and teachers differ in their perceptions of the impact on professional practices for the Tulsa Model Evaluation process and if so, to what degree?
6. Do principals and teachers differ in their perceptions of the context of professional conversations during the Tulsa Model Evaluation process and if so, to what degree?

Methodology

The study employed a cross-sectional survey of Oklahoma teachers and administrators currently using the Tulsa Model Evaluation instrument from the 2019-2020 academic year. The survey was distributed by permission from September to December of 2019. Survey responses were collected using a web-based survey platform. Answers were provided based on a Likert scale. Data from the survey were analyzed in two phases.

Phase I included an analyses of demographic statistics to determine the characteristics of the sample group and distinguish the differences in perceptions of usefulness from teachers and administrators.

Phase II measured the variance of perceptions between principal evaluators and teachers. A Welch alternative ANOVA was used to determine what degree of differences occur in six areas based on growth condition attributes defined in the Stiggins and Nickel (1989) Teacher Evaluation Profile and modified by Huckstadt (2011). These six areas were

1. knowledge and understanding of the Tulsa Model Evaluation process;
2. perception of the quality of the evaluation process;
3. qualifications of the evaluator;
4. quality of the data collection and feedback;
5. impact on professional practices;
6. and personal impression of the context of the evaluation process.

Significance of the Study

My study adds to the research on implementation of standards-based evaluation processes across the nation. By surveying teachers and principals in the state of Oklahoma currently utilizing the Tulsa Model Evaluation instrument, I intended to explore teacher and principal evaluator perceptions about the quality and impact of the state mandated Tulsa Model Evaluation process used in 484 of the 551 schools across the state. The Tulsa Model Evaluation instrument is a standards-based evaluation instrument like tools developed in Iowa, Wisconsin, New Jersey, and other states around the country during the accountability reform movement. Studying the perceived effect of the implementation of these tools can shed light on their use and the development of principal and teacher collaborative conversations during the evaluation process, but more importantly, on whether these state mandated tools create conditions for teacher growth and are perceived as useful and impactful for teacher effectiveness.

No statewide baseline data on the quality and impact of the Tulsa Model Evaluation process have been collected since the model was fully implemented in 2014. The data collected and subsequent analyses of teacher and principal evaluator perceptions will increase educational professionals and policy maker understanding of the impact of standards-based evaluation processes used in Oklahoma schools. This study will also aid in understanding strengths or gaps in conditions for improving teacher professional growth for schools utilizing the Tulsa Model Evaluation process.

Limitations

Although this study intended to gather data from all school districts across Oklahoma using the Tulsa Model Evaluation instrument, participants self-selected based

on their willingness to join in the study. Teachers and principals who currently have a bias toward either favoring or disfavoring the evaluation process were most likely to complete the survey. The survey instrument used in this study has been used as part of previous research. Developed by Stiggins and Nickel (1989) to determine usefulness of teacher evaluation in certain schools, the survey, in various forms, has been used in dissertations to determine the perceived utility of standards-based teacher evaluations in other states (Huckstadt, 2011; Schumacher, 2004; Towe, 2012). The lack of random assignment of participants presented an additional threat. Since this study employed a cross-sectional survey rather than a randomized control experiment, there is an inherent threat to the reliability of results. Additional limitations include Type I errors due to the lack of homogeneity of variance in the sample size. The number of teachers in the sample included 718 (n=718) compared to the number of principal evaluators which included 137 (n=137). Welch alternative ANOVA was used to mitigate Type I error.

Definitions of Terms

No Child Left Behind (NCLB) – A 2001 amendment to the 1965 Elementary and Secondary Education Act. Enacted during the President George W. Bush administration, the amendment required states and local education agencies to report student achievement goals based on the number of students reaching proficiency with the overall goal being 100% of students proficient within a specified time (USDE, 2017).

Race to the Top (RttT) – Race to the Top is a term used to describe a series of grants offered to state education agencies during the President Barack Obama administration. These grants were issued as incentives to schools through the American Recovery and Reinvestment Act of 2009. Criteria for receiving the grants included six

areas of school improvement, one of which is the state's ability to develop "Great Teachers and Leaders" (USDE, 2009).

Every Student Succeeds Act (ESSA) – This legislation serves as the reauthorization of the 1965 Elementary and Secondary Education Act (ESEA). Signed into law in 2015, ESSA shifts the focus of student achievement goals from 100% proficiency to indicators of student growth. It also allows state and local education agencies to utilize other measures of success including student attendance. ESSA eliminates requirements for schools to employ what NCLB defined as highly qualified teachers.

Teacher Leader Effectiveness (TLE) – Legislation passed in Oklahoma in 2010 to establish a committee to oversee implementation of a unified statewide teacher and leader evaluation system (OSDE, 2014; Oklahoma Teacher Leader Effectiveness Evaluation System, 2016). The law came into effect in 2011 and required schools to use a state approved teacher evaluation model to measure teacher effectiveness.

Tulsa Model Evaluation (TME) – One of two statewide standards-based evaluation processes in Oklahoma used to evaluate teachers and comply with state TLE law. The Tulsa Model Evaluation system was developed by Tulsa Public Schools in 2009 through assistance from the Bill and Melinda Gates Foundation (Tulsa Public Schools, 2017).

Summary

Schools are now free from stiff accountability regulations of NCLB and grant driven educational reforms of Race to the Top. Continued obligations for system improvement are required with the Every Student Succeeds Act, albeit with greater

flexibility for states to determine how to apply research-based practices. States must learn to use the tools mandated by law under Race to the Top in a new context, a context of a learning organization which creates conditions for teacher growth, to get results (Stronge & Tucker, 1999; Van Clay & Soldwedel, 2013; Stiggins & Nickel, 1989; Stiggins & Duke, 1988). Otherwise, rhetoric about actors within the system being incompetent to correctly implement reforms will persist (Weisberg et al., 2009). Research has been done on leadership behaviors but not on mechanisms for successful organizational change using systems thinking and learning organizational theory (Honig & Louis, 2007; Heck & Hallinger, 2010). A gap exists in quantitative research for merging the reforms of the effective school movement and the accountability reform movement with local school contexts.

CHAPTER 2
REVIEW OF THE LITERATURE

Introduction

The nature and purpose of this review of the literature is to connect ideas generated from over 30 years of research on teacher evaluation and school improvement. In the era of the Every Student Succeeds Act (ESSA), schools face the challenge of merging two competing philosophical educational movements, effective schools and accountability reform. The current call for an integrated learning approach to school improvement resonates with well-intentioned educators who are still struggling to use accountability reform tools in ever evolving, change-oriented school contexts. The evolution of reform through pre-No Child Left Behind (NCLB), Race to the Top, and now moving into the age of revised ESSA has provided a rich organizational and social perspective of what works in schools. It is time to use our understanding of collaborative environments proven effective during the effective schools movement and the tools, such as standards-based teacher evaluation instruments, generated during the accountability reform movement as part of one system to increase conditions for teacher growth.

What we knew about standards and effective practice was just beginning to flourish until the NCLB era swept the nation and interrupted the effective schools movement (Fullan, 2009; Hallinger & Leithwood, 1998; Lezotte, 2001). Although politicians and education reformers debate about qualitative and quantitative measures for improving teacher effectiveness, researchers have concluded qualitative feedback in the form of teacher evaluation is effective for improving student performance (Kane, et al., 2011; Hill & Grossman, 2013; Taylor & Tyler, 2012). When teachers participate in

conversations with instructional leaders about specific learning strategies and how those strategies merge with standards of practice, student learning improves (Darling-Hammond et al., 2012; Hattie, 2012; The Danielson Group, 2013). However, feedback must be part of a system which encourages continuous learning, not only for the student but also for the teacher and principal (Bryk & Schneider, 2002; Connally & Tooley, 2016; Dufour & Fullan, 2012; Dufour & Marzano, 2011; Fullan, 2009; Hattie, 2015b; Kane, Kerr, & Pianta, 2014; Stronge & Tucker, 1999; Tucker & Stronge, 2005; Van Clay & Soldwedel, 2013). Yet, as teacher evaluation systems are introduced, whether required by law or encouraged by central offices as part of collaborative learning or professional learning communities, the purpose of evaluation must be clear to principals. In practice, principals and teachers see the evaluation as 1. quantifying teacher practice to "weed out bad teachers;" or 2. contributing to the professional learning of the teacher (Marzano & Toth, 2013; Darling-Hammond, et al., 2012; Fullan, 2013). School leaders struggle to achieve both results concurrently.

Principals may choose keeping teachers happy over enforcing a mandate, focusing on maintaining a stable or positive school climate, or worse, may fail to fully implement the evaluation instrument because they have a negative opinion about the process (Tuytens & Devos, 2011; Hallinger, Heck, & Murphy, 2014). Principals are actors within a social and political system. When they are not trained, or teachers are not trained, how to leverage actions within a social system, structural mandates fail (Weisberg, et al., 2009). A systems approach in which professionals are encouraged to be learners within an organization is ideal for creating conditions for professional growth.

Leaders must understand the importance of a systems approach to school improvement, and a principal should receive training and continued professional development in the role of the instructional leader to orchestrate multiple components of school improvement (Anderson & Reynolds, 2015; Heck & Hallinger, 2010; Honig, et al., 2010; Manna, 2015). Teacher evaluations are most effective within an overall system of school improvement culture (Goe, Biggers, & Croft, 2012; Leithwood & Louis, 2012). In a school site, a system of school improvement includes not only strengthening engagement of parents, the community, and the curriculum, but also building teacher capacity (Byrk, et al., 2010; Augustine, et al., 2009). Proper drivers used to build school culture must be understood by principals and teachers (Fullan, 2014; Honig & Louis, 2007).

The way a principal spends time as the lead learner building trust, structuring adult learning, and engaging in purposeful conversations about instruction is important to the overall outcome of teacher evaluation (Darling-Hammond, 1986; Darling-Hammond, 2013; Fullan, 1995; Herman, et al., 2016; Bryk & Schnieder, 2002; Hoy, Tarter, & Kottkamp, 1991; Lezotte, 1992; Marzano, 2003; Sebastian & Allensworth, 2012; Stronge, 2013; Wallace Foundation, 2012). Without understanding how the evaluation tool fits into an overall system, a principal will not achieve desired results for student learning (Bryk, et al., 2010; Fullan, 1995; Hallinger & Heck, 2010; Hallinger, Heck, & Murphy, 2014; Honig; 2008; Leithwood & Louis, 2012; Louis, et al., 2010; Van Clay & Soldwedel, 2013).

The content of this review will bring together themes from the effective schools movement and the accountability movement by discussing the reasons teachers are

evaluated, how schools encourage the work of a learning organization, how principals and teachers behave as actors in learning conversations, previous research in conditions for teacher growth, and finally, the next steps in quantitative research for leveraging will and skill in the system to improve student learning.

Why Evaluate Teachers?

Beginning in the 1980s, early conversations about teaching revolved around developing standards of best practice to reflect the complexity of teaching. Pioneers of the field like Linda Darling-Hammond and Charlotte Danielson discuss using frameworks to have conversations about how teachers impact student learning. Darling-Hammond (1986) stresses the importance of teacher input in the evaluation process. Teachers should be part of a collaborative process in discussing learning for students. During the height of the Race to the Top movement, her work *Getting Teacher Evaluation Right* (2013) urges educational leaders to move away from hardline teacher evaluation systems aiming only to quantify the complexity of teaching to fire "bad" teachers. Her research strongly supports using teacher evaluations as part of a system of school improvement.

Charlotte Danielson's the *Framework for Teaching* (The Danielson Group, 2013) developed from the same need to create indicators of effective teaching. The *Framework* came to be used by districts to have conversations about teaching, but it also began to be adopted as an evaluation tool, especially as states and districts competed for Race to the Top grants. Oklahoma adopted a version of *Framework* through its Tulsa Model Evaluation instrument (Oklahoma State Department of Education, 2014). The use and adaptations of Danielson's *Framework* illustrate how conversations about effective

practice during the effective schools movement became shoehorned into the needs of the accountability movement.

For teacher evaluation to work, it should be used as part of a system of feedback to teachers. Teachers are learners, and learners retain and apply more of their learning when feedback is given. Understanding the impact of specific teaching strategies has assisted in improving student achievement around the world. John Hattie (2003; 2015a) uses over 15 years of educational research to analyze specific strategies and their effects on student learning. His meta-analysis has informed educational leaders of what works and what does not work in schools.

Tucker and Stronge (2005) link teacher feedback to ongoing professional development. Using case studies of teacher evaluation systems in four states, they suggest teacher evaluation works best in a system of school improvement. All four states use some form of professional development linked to the evaluation tool. Tucker and Stronge (2005) highlight the impact a teacher can have on a student by quoting a Tennessee study which claims a teacher can impact a student's ability to learn for up to three years after the time of the teacher's instruction. Results such as these became fuel for arguments in the accountability movement to expose ineffective teachers and eliminate them from the system. Teacher evaluation frameworks became the mechanism of choice to identify ineffective teaching practices, enumerate those practices, and force teachers out of the profession.

A key focus of the accountability movement became improving teacher quality. Under the pressure of improving schools as part of a national movement, the Measures of Effective Teaching Project (METS) makes a case for pinpointing and measuring effective

standards and using these data points to provide feedback to teachers (Kane, Kerr, & Pianta, 2014). A drawback to this research is it prompted more intense quantitative processes such as Value-Added Measures (VAM) and these measures, or data points, were often used to "weed out bad teachers" (Weisberg, et al., 2009). In response to the use of data which fail to see the complexity of teaching practice, The New Teacher Project published *The Widget Effect* (2009). This report acknowledges the misuse of evaluation systems in the accountability movement and offers recommendations based on studies of twelve school districts across the country. Recommendations include understanding teachers must have differentiated professional development to improve practice and training principals to use the evaluation system and holding them accountable if they do not.

The findings in *The Widget Effect* provide guidance for moving forward with teacher evaluation tools developed before, and refined during, the accountability movement. Although there is a call for eliminating these systems because they are considered a waste of time and resources (Dynarski, 2016), researchers are finding when evaluation systems are understood and valued by actors within the system, they can have a significant impact on student learning. Merging evaluation instruments within a system of school improvement is the next step in best utilizing standards-based teacher evaluation systems (Hill & Grossman, 2013; Hallinger, Heck, & Murphy, 2014; Honig, 2008; Leithwood & Louis, 2012; Marzano & Toth, 2013).

Systems Thinking and the Learning Organization

Learning organizations exhibit characteristics or disciplines outlined by Peter Senge (2006). These characteristics include 1. building a shared vision to build a genuine

commitment; 2. personal mastery and reflecting on a personal vision; 3. creating mental models and sharing these with others; 4. team learning and collaborative communication; 5. systems thinking which fuses all the previous disciplines and brings them into action (Evans, Thornton, & Usinger, 2012). Applied in a public school context, administrators can create conditions which require adult learners to reflect on each of the five elements of a learning organization. Principals and teachers together create structures and norms to modify behaviors and support a more holistic way of thinking (Evans, Thornton, & Usinger, 2012; Fillion, Koffi, & Booto, 2015). When school leaders understand how to develop personal mastery for adult learning, mental models around mandates, a shared vision for generative learning, and habits for team learning, they can more smoothly manage school improvement initiatives. Understanding systems thinking aides school leaders in creating conditions for adult learning and school improvement.

Organizations in their efforts to be responsive to the complexity of change and human behaviors attempt to put in place structures which organize learning (Bolman & Deal, 2013; Evans, Thornton, & Usinger, 2012). Education leaders, striving to manage the complexities of schools, often rely on structural techniques to enforce change. Anthony Muhammad and Sharroky Hollie (2011) describe this practice as "institutionalizing cultural health" (p. 3). In a report for McREL, McIver, et al. (2009) define highly structured learning organizations as High Reliability Organizations (HRO). The McREL report emphasizes the importance of structure in preventing systemic failure to achieve certain outcomes. Qualities of HROs include a preoccupation with operational breakdowns, a reluctance to simplify operations and interpretations of lessons learned, and a commitment to organizational resilience. Schools may aspire to be collaborative

learning organizations within highly structured environments, but these highly structured environments have not delivered on results based on the promises of either the effective schools movement or the accountability movement (Hallinger & Heck, 2014).

Fullan (1995; 2009; 2013) has been an early proponent of merging structures of learning organization theory to the politics of social learning theory. Building on social cognitive theory (Bandura, 2001), Fullan connects the ideas of will and skill within the context of tri-level reform: districts, communities, and states working in concert (2009). Fullan asserts school reforms have failed because the politics of school cultures are not considered when implementing change (1995). His work suggests education reform movements ignore human behaviors within the dynamics of reciprocal work, and educational leaders have been trained to privilege fragmented initiatives rather than the synergy of a solutions-oriented system (2009). In a learning organization, teachers and leaders should be measuring learning, not teaching (2013). All actors within the organization are learners and capacity and clarity can be reached by leveraging structure within social relationships.

Honig, et al. (2010) brings organizational and sociocultural learning theories into a framework of practice for district leaders and explains the importance of educating leaders in their role as actors withing a learning organization. Despite over a decade of pressure for schools to be accountable and effective within a systems culture, very little actionable guidance has been given to leaders about how to accomplish systemwide school improvement reform. A logic model describing assistance relationships in schools sets itself apart from the linear pathways models of previous research (Bryk, et al., 2010; Hallinger & Leithwood, 1998). Proposing that central office administrators aid school

sites by modeling behaviors, creating and sustaining social engagement, developing learning tools, and supporting engagement in joint work, district leaders act as boundary crossers, relaying the cultural and structural language of right action across the district. The list of actions required in assistance relationships spans both the structural elements of organizational learning and the socio-political aspects of social learning theory. All actors in the system are encouraged to co-create the learning with district leaders acting as centralized leaders of learning.

Honig, et al., (2010) combines the lenses of organizational learning theory and social learning theory to explore the role of the principal within a synergized learning organization. Principals should be trained not just in habits of mind or abstract behaviors of trust and vision but in specific actions to use tools of social capital. Honig, et al., (2010) define these tools as leveraging a network and explicitly teaching the ways of working in an organization. Specifically, principals should be trained in "differentiating supports; modeling ways of thinking and acting; developing and using tools; and serving as a broker between principals and external resources" (p. 19).

Current research emphasizes principals be motivated toward right action and right work. At issue is how right action continues to be defined. At the closing of NCLB era, educational research proposes a principal should be a lead learner (Fullan, 2010; Strong, 2013), a lead innovator (Christensen, et al., 2008), promote collective leadership (Leithwood & Louis, 2012); promote trust (Bryk & Schneider, 2002; Sebastian & Allensworth, 2012), and closely monitor student progress (Hattie, 2015). However, in a learning organization, the principal is not the only actor influencing student learning.

Leaders in a Learning Organization

The reciprocity of behaviors within an organization aide the learning of the actors within the organization (Bandura, 2001). Applying systems thinking and learning organizational theory to public school settings clarifies the specific contextual dynamics at play, not only of student learners, but of the adult learners as well. The culture of a school site determines the ability of the principal and teachers to navigate interaction around cognitive, behavioral, and contextual factors. Social cognitive theory asserts people shape their own learning, which may translate into agency and efficacy; however, translation is not guaranteed. Because we are exposed to learning but may not fully demonstrate mastery of the learning, it is important to also view the implementation of mandates such as Oklahoma's Teacher Leader Effectiveness Law through the theory of social politics.

Within this context, action is created and exchanged to build capacity within an organization through intentionality, forethought, will, and efficacy (Bandura, 2001). Most research supports schools working in this way because collective work is more efficient at getting desired results than educators working alone (Dufour & Marzano, 2011; Fullan, 2010; Fulmer & Keys, 1998; Marzano, Waters, & McNulty, 2005). Previous research supports coaching conversations within schools that employ a systems thinking approach (Bryk & Schneider, 2007; Bryk, et al., 2010; Fullan, 2009; Fulmer & Keys, 1998; Leithwood & Louis, 2012; McIver, et al., 2009).

A goal of my research is to explore the perceived quality and impact of the Tulsa Model Evaluation process through data collected by surveying teachers and principals. As actors in the system, principals and teachers should exhibit reinforcing skills which

impact the overall learning environment (Augustine, et al., 2009; Bowers, 2016; Dufour & Marzano, 2011; Neumerski, 2013; Marks & Printy, 2003; Stronge & Tucker, 1999; Van Clay & Soldwedel, 2013). Principals and teachers should be trained and receive mental models for using tools within the organization, leveraging their networks to improve student learning, and getting results (Honig, 2010; Honig & Louis, 2007; Kane, Kerr, & Pianta, 2014).

Principals play a critical role in the success of the evaluation process to create conditions for teacher professional growth. Communication about instructional practices, creating a democratic school culture, and influencing change are all aspects of the role of school principal. The effective schools movement began to outline qualities of practice for school leaders in the 1990s (Lezotte, 1992). Transformational leadership is defined as leaders who tend to follow the collaborative leadership process (Stronge, Richard, & Catano, 2008) and who can navigate the democratic political and social aspects (Lezotte, 1992) of the school. Instructional leadership is defined by habits or actions of the principal (Hallinger & Leithwood, 1998), which can be distilled into a checklist of skills (Stronge, Richard, & Catano, 2008). Strong, Richard, and Catano (2008) continue the delineation of instructional versus transformational leader in *Qualities of Effective Principals*. In their review of the literature, checklists of principal skills are highlighted. Multiple lists of red flags of principal leadership are also outlined, which include traits such as “lacks a clear vision; is rarely seen in classrooms; and fails to provide feedback to teachers in the spirit of coaching as well as evaluation” (p. 172).

Early work focuses on specific practices or behaviors of a principal rather than helping a principal navigate a political context (Hallinger & Murphy, 1985). Quantitative

studies attempting to create pathways of effectiveness from principal to student learning emerge in the 2000s. These studies begin to discover the differences between principal effectiveness actions and the context in which a principal leverages these actions.

Linking standards-based teacher evaluation to student learning and school improvement brings to light the issue with the current Tulsa Model teacher evaluation system. It is fragmented and alienated from the actual practice of teaching. Darling-Hammond (2013) asserts an evaluation system should be used as part of ongoing professional development and conversations within the learning organization. She describes research to support four features of professional development: it should be intensive, ongoing and connected; it should focus on the teaching and learning of specific academic content; it should be connected to other school initiatives; and it should build strong working relationships among teachers. In light of the *Widget Effect* (2009), the benefit of contextual professional development for teachers is evident after linking teacher evaluation to student learning (Borko, Liston, & Whitcomb, 2009; Taylor & Tyler, 2012; Weisberg, et al., 2009). When operating within a culture of continuous conversations about learning, teachers and principals exhibit the same needs as any learner, a need for connection and application of the skills to be learned.

Fauske and Raybould (2005) link learning organization theory to a technology integration initiative in an elementary school. The study uses the concept of mental models to highlight the unsuccessful implementation and communication of a school reform. The breakdown occurs due to teacher and administrator misunderstandings of the technology framework and the lack of communication from administrators of the mental model, or the example of technology change in its final form. The mental model aids the

development of new routines. Routines build the daily habits needed for an initiative to become common practice within the organizational culture. The researchers found that a lack of communication creates a fuzzy mental model, and without the mental model, the learner cannot adopt or adapt the habit. An organization which cannot integrate new habits will fail to implement or garner expected results from new initiatives. When applied to an evaluation framework, this study may be recreated to examine the context of the Oklahoma Teacher and Leader Evaluation system, and the ability of schools to create mental models of the evaluation framework for successful results.

The work of Fauska and Raybould (2005) implies the importance of explicitly training leaders in systems thinking theory, but it does not clearly define the process or skills leaders need to build a learning organization. When leaders do not have the knowledge to implement change yet are mandated to use a framework or tool for change, the results are often disappointing. Murphy, Hallinger, and Heck (2013) point to the flaws in implementing an evaluation system without proper administrator training. In fact, they assert implementing an evaluation tool without proper understanding of the "muscle" the tool welds can have a detrimental impact on a school's culture. Overall, their research found little link to the teacher evaluation process and student achievement. Based on these findings, it seems doubtful a school leader will create conditions for teacher growth if he or she does not understand how the evaluation system aligns to the context of a learning organization.

Conditions for Teacher Growth

Hallinger, Heck, and Murphy (2014) provide a logic model synthesizing the theory of action behind the teacher evaluation policy movement. In their critical examination of the literature on teacher evaluation, they tackle assumptions about its effectiveness: the link between teacher effects and student learning; the evidence of teacher evaluations on student learning growth; and sociological theories about schools as organizations for learning. The logic model suggests teachers have a significant impact on student learning and researchers should be able to measure variables of interaction along pathways to student results. This review points to the failed assumptions that principals, using teacher evaluation, have any impact on student learning. The complexity of factors associated with principal impact on teacher effectiveness which results in improved student learning comes to light. The role of the principal moves beyond the binary theories of transformational leadership or instructional leadership. A question arises about why principals are not effective in influencing teacher practice through evaluations. If the evaluation system is seen as a "threat" (p. 20) within the politics of the system, principals will often choose to undervalue, misuse, or disregard the evaluation instrument. Hallinger, Heck, and Murphy (2014) do point out a promise for teacher evaluation within a context of "nonevaluative channels" (p. 22). This promise elicits further study for educational leaders who must use these state mandated evaluation instruments. More research is needed to determine the effect of a teacher evaluation process when it is used as part of a learning organization striving to create conditions for teacher growth.

Teacher Effectiveness

At least since the 1990s, researchers have attempted to quantify the qualitative components of effective teaching. The results have been a set of indicators, skills, standards, or behaviors which can be placed in a framework or model of effective practice (Darling-Hammond, 2013; The Danielson Group, 2013). Before the requirements of NCLB, leaders in the field tended to agree teacher inputs in student learning were more important to measure than the outputs (The Danielson Group, 2013). With the emphasis on high stakes testing, the focus on outputs, in the form of student achievement, has taken the conversation from measures of behavior to measures of standardized results. Since standardize assessments are a blunt instrument to measure the full scope of human learning and behavior, economists have reassured education reformers they have developed formulas which account for variables such as poverty, race, and gender (Green, 2014; Hanushek & Rivkin, 2012). The quantitative algorithms may satisfy proponents of market-based solutions for education reform, but professional educators, who have cut their teeth on evidence-based practices which influence human behavior, see no place in practice for algorithms. Most educators prefer fostering relationships, and as a result, focus school improvement efforts on the dynamics of collaborative learning environments.

Quantitative studies have been developed to measure the effects of collaboration and teacher effectiveness without focusing solely on outputs. Results from a quantitative study on the *Learning How to Learn* initiative in England by James and McCormick (2009) suggest a strong statistical relationship between school policy, teacher professional learning, and their capacity to promote student ownership for learning in

their classrooms. Because a goal of the initiative was to promote learning autonomy, teachers experienced greater success with the initiative when they directed their own learning through teacher networks. Teacher learning was supported by school administrators and collective inquiry was identified as a key factor in the success of the initiative.

Shared leadership and teacher voice in evaluation systems are established through specific activities or action steps within a learning organization. Professional learning goals and professional growth plans develop from specific feedback while evaluating teacher performance and provide reflective practice for teachers (Marzano & Toth, 2013). Self-reflection is a key factor in developing cognition and human agency (Bandura, 2001). Creating a structure in which teachers are setting their own goals for learning deepens the discussions and can be a part of what Marzano (2003) outlines as action steps for professionalism in *What Works in Schools: Translating Research into Action*.

Using meta-analysis of research dating back to the *Coleman Report* in the 1960s, Marzano (2003) suggests teachers should have content knowledge as well as pedagogical knowledge. To ensure teacher professional development is effective, four action steps should be followed: establish norms of conduct and behavior that engender collegiality and professionalism; establish governance structures that allow for teacher involvement in decisions and policies for the school; and engage teachers in meaningful staff development activities (p. 65). Teacher professional development should provide opportunities for active learning and be consistent with the learning goals established by the school's mission. Professional development should be subject-specific, tied to

context, and be related to on-going conversations about learning. Collaborative learning opportunities are also regarded as effective.

My study built on previous research and gathered principal and teacher perception data to understand the degree to which the standards-based evaluation instrument meets the teachers' need for professional growth. Using a cross-sectional survey for this study provided insight into the engagement of teachers in mandated accountability reform efforts, especially those efforts claiming to define teacher effectiveness and provide opportunities for professional development.

Chicago and Cincinnati

A growing segment of inquiry connects principal leadership to teacher evaluation and feedback, which ultimately impacts student learning and achievement results. A study conducted in Chicago Public Schools creates a pathway from principal leadership to student outcomes on the ACT Education Planning and Assessment System (Sebastian & Allensworth, 2012). Through multilevel equation modeling and analysis of teacher survey results, researchers created a link from principal actions to student achievement. The results found although a principal's direct impact on student learning is small, the overall school climate and learning environment created by the principal makes the biggest difference for student achievement.

In addition to the Chicago study, researchers in Cincinnati Public Schools found the qualitative data collected via classroom observation tools may also have positive influence on student achievement (Sebastian & Allensworth, 2012). Researchers in this study limited their data to the established Cincinnati Teacher Evaluation System

instrument, which is based on Charlotte Danielson's framework for effective teaching (Taylor & Tyler, 2012). In a comparative analysis of teacher evaluation scores and student achievement scores over time, the results suggest the act or the process of evaluation has a positive impact on student achievement, especially in math. Taylor and Tyler (2012) also uncover evidence that critical feedback to teachers improves student performance. The Cincinnati study makes a case for an evaluation instrument, if used to provide specific feedback to teachers, making a difference in student learning.

Teacher Evaluation Profile

In the 1980s, Stiggins and Nickel (1989) developed an instrument to assess the conditions for teacher growth via an evaluation system in a school context. The theory of action for this instrument is under the right conditions, teachers will grow professionally through the correct use of an evaluation tool. The Teacher Evaluation Profile (TEP) was created after multiple reliability and validity studies to measure the potential of an evaluation instrument to create the proper conditions for teacher professional growth. The TEP measures 44 key attributes of their evaluation experience to include not only conditions of the teaching environment but also qualifications or qualities of the evaluator. Participants were asked to rate on a 0 to 9 scale the quality and impact of their most recent evaluation experience.

The instrument was tested for reliability and validity in conjunction with Northwest Regional Education Laboratory profiling services in 1989. It was distributed to teachers in 27 schools across the United States during the 1987-1988 school year. Correlation studies were completed to determine the relationships between individual attributes and teacher perceptions of the quality and impact of their evaluation

experiences. Researchers found a high correlation between certain conditions such as "quality of ideas in feedback" to the overall perception of quality and impact of the evaluation process. Because the TEP is easy to administer and has the potential to detect the faults of an evaluation environment, it can be used to assess the effectiveness of an evaluation system (Stiggins & Nickel, 1989). A body of research has been established around the key attributes found in the TEP. The emergence of standards-based evaluations systems under the accountability reform movement has brought new use for the TEP's key areas of study, in particular the knowledge and understanding of the evaluation system; perceptions regarding usefulness; qualifications of the evaluator; quality of the data and feedback from principals; the impact of the evaluation process on professional practices; and personal impressions of the evaluation system on professional conversations. These key areas form the basis for the research questions in my study.

Next Steps in Quantitative Research on Evaluation Frameworks

Previous dissertation studies measuring the perceived effect of standards-based evaluation tools were completed in Iowa, Wisconsin, and New Jersey. The Schumacher (2004) study uses a mixed methods approach to collect teacher and principal perceptions of the implementation of a standards-based teacher evaluation tool in one school district in Wisconsin. The results highlight the mixed responses of teachers when answering survey questions compared to more definitive opinions during personal interviews. In survey responses, teachers report an understanding of the evaluation tool and believe indicators used in a standards-based evaluation model correlate to improved teacher practice. However, when interviewed teachers did not believe the evaluation process

improved their practice. The results of this study hint at a void between the intent of the evaluation tool and its implementation to improve learning outcomes.

A study conducted by Towe (2012) in urban schools in New Jersey utilized a mixed methods approach to determine the impact of a standards-based evaluation model on teacher effectiveness. Results determined perceptions between principals and teacher were similar in that the evaluation process had little effect on improving teacher practice; however, the summative evaluation did have some impact on targeting professional development for teachers. Findings suggested the best use of the evaluation tool is through a collaborative, coaching conversation in which a common language of effective practice is established. Recommendations from this study included measuring specific training needs of teachers and administrators to drive a stronger "growth-oriented" teacher evaluation process.

A study of Iowa teachers and principals attempted to measure the impact of the state's mandated standards-based evaluation system and the difference in perceived use of the system based on the perceptions of teachers and principals (Huckstadt, 2011). Findings from this study indicated administrators consistently rated the evaluation system as more useful than teachers in most categories. However, recommendations from the study suggested a statistically significant impact within isolated sub-questions of the survey. For instance, both administrators and teachers reported a small positive impact on connecting the evaluation system to professional development. Findings also indicated both administrators and teachers achieved an understanding of the evaluation tool, but they ranked the quality of feedback and the impact on effective teaching low. Overall, administrators ranked the impact of the evaluation tool higher than teachers. Teachers

saw less connection between evaluation feedback conversations and their teaching practices. The summative evaluation seemed to be the most impactful event in the evaluation process for administrators and teachers. Among the recommendations from Huckstadt's (2011) study was a suggestion for training administrators and teachers in coaching conversation skills and in self-evaluation skills to build agency. The Iowa study suggests more work is needed to truly understand the instructional practices around using a standards-based evaluation system rather than simply implementing a system and merely training principals in procedural logistics.

The Oklahoma Model: Theory of Action and History of the Teacher Leader Effectiveness Law

The theory of action presented by the Oklahoma State Department of Education during the implementation of the Teacher Leader Effectiveness Law assumed principal leadership through certain actions influenced teacher effectiveness (Oklahoma State Department of Education, 2014). These actions were thought to improve student achievement via a process of change in teacher performance. The Oklahoma theory mimics research supporting actions of a school leader to make a difference on standardized assessment results. What remained missing from the Oklahoma model, however, was the attention to feedback and support.

A major component of implementation of the Teacher Leader Effectiveness Law was mandating all public schools in Oklahoma adopt a standards-based evaluation process such as the Tulsa Model Evaluation. The Tulsa Model Evaluation process was developed in Tulsa Public Schools in 2009 in conjunction with research funded by the Bill and Melinda Gates Foundation (Oklahoma State Department of Education, 2014).

According to a document on the Oklahoma State Department website, the evaluation instrument was included in a 2011 MET Validation Engine Analysis which "confirmed that the Tulsa model measures what matters – that it captures practices that are empirically associated with gains in student achievement." (Tulsa Public Schools, 2017).

Since its implementation, no study has been conducted to evaluate the impact of the Tulsa Model Evaluation process on effective teaching across Oklahoma. My study explored perceptions of teachers and principals on the quality and impact of the Tulsa Model Evaluation process and provides recommendations for future policy and research based on the results of a cross-sectional survey of 54 Oklahoma schools.

Current Conversations about Standards-Based Teacher Evaluation Models

A statistic from *The Widget Effect* (2009) became an alarm bell for researchers of teacher evaluation during the accountability movement: 94% of teachers receive the highest or second highest rating on a multi-domain system. Kane, et al., (2011) use this statistic to set the stage for their study of teachers in Cincinnati Public Schools. The Cincinnati study provides an argument for continuing the use of standards-based teacher evaluation systems. The results indicate that observations, when used in conjunction with specific feedback from evaluators, can make a difference in school improvement. The Cincinnati study begins to point researchers in the direction of differentiated variable effects rather than only using broad domain ratings as links to student learning. When considering variables within the context of teaching strategies and principal behaviors and feedback, researchers can think about the impact of collaborative conversations and the role systems thinking within a learning organization plays in teacher professional growth. Schools are not just organizational structures. They are political systems.

Principals and teachers each play a role in these systems, and how each view their work has an impact on student learning.

Using quantitative and mixed methods provides a deeper dive into the nuances of coaching conversations or relationships built around learning conversations and is gaining ground in empirical research. Heck and Hallinger (2010) built models on pathways of principal behaviors which lead to student achievement. Results on student achievement are positively correlated to organizational changes around collaborative leadership. Using three statistical models to correlate the effects of reciprocal behaviors or change initiatives, they target and isolate actions in the process of school improvement. The findings suggest future research which can quantify the impact of certain high leverage leadership behaviors and effects on student learning.

As the evolution of teacher evaluation moved through the school reform movement of the early 2000s, evaluation instruments were tools for disruptive educational reform. Quantitative studies have emerged linking teacher evaluation scores to student achievement through Value Added Measures (VAM) (Kane, et al., 2014). Although results of VAM studies are mixed, dismissing teachers due to low performance on value added measures has fallen out of favor in the new era of the Every Student Succeeds Act (Connally & Tooley, 2016; Hanushek & Rivkin, 2012). A call for quantitative studies linking teacher evaluation to collaborative cultures is emerging (Bowers, 2016; Heck & Hallinger, 2010; Henson, Hull, & Williams; 2010). Because researchers know more about quantifying the complexity of social relationships in schools, using factors of climate as independent variables, a case can be made for

increased quantitative research to discover the effect of leveraging social and structural behaviors within specific school contexts (Honig & Louis, 2007).

Quantitative research attempts to draw a hard line, or pathway, from teacher actions to school improvement or principal action to teacher action to school improvement. Multiple logic models have been developed to explain the factors which have the least or greatest impact on the linear progression of leadership behaviors as an input passed through teacher behaviors to effect student outcomes (Hallinger & Heck, 2014; Herman, et al., 2016; Leithwood & Jantzi, 1999; Louis, et al., 2010; Marzano, 2003). Recent calls for additional studies focus on the use of an evaluation instrument as part of a system or overall culture of school improvement (Hallinger & Heck, 2014; Hill & Grossman, 2013; Goe, Biggers, & Croft, 2012; Hallinger, Heck, & Murphy, 2014; Murphy, Hallinger, & Heck, 2013; Taylor & Tyler, 2012).

Multiple studies have measured the impact of the principal as the agent or vehicle for delivering feedback to teachers based on standards of effective practice. When used as a tool for coaching rather than for weeding out bad teachers, the evaluation instrument provides a pathway to increased school improvement (Hattie, 2015; Kane et al., 2014; Murphy, Hallinger, & Heck, 2013; Stronge & Tucker, 1999; Taylor & Tyler, 2012; Tucker & Stronge, 2005; Wallace Foundation, 2012). The Cincinnati study found teachers can gain skill in effective practice through principal input via evaluation (Kane et al., 2011; Taylor & Tyler, 2012). This increased skill shows a positive effect on student learning over time.

The bulk of early research on teacher evaluation focuses on teacher perception of the evaluation instrument and whether this perception impacts student achievement.

When the standards-based evaluation movement began in the 1990s, studies focused on the teacher's relationship to student achievement because little was known about the principal connection to student learning (Darling-Hammond, et al., 2012; Darling-Hammond, 1998; Hill & Grossman, 2013; Hallinger & Heck, 2014). The premise behind many of these studies emphasizes using the evaluation instrument as part of a collaborative conversation between principal and teacher. A strong argument for professionalism on the part of the evaluator is made. To link student learning to the evaluation process, a teacher must respect the principal and the principal's ability to create an atmosphere for learning. The respect gained by the principal builds a willingness to hear feedback which can result in an optimal learning environment for students (Darling-Hammond, et al., 2012; Green, 2014; Heck & Hallinger, 2010; Hattie, 2015; Marzano & Toth, 2013; Murphy, Hallinger, & Heck, 2013; The Danielson Group, 2013; Weisberg, et al., 2009).

Summary

By studying the perceived impact of the Tulsa Model Teacher Evaluation process in Oklahoma, I add to the emerging research on the possible effects of mandated teacher evaluation reforms. Exploring the relationship between attitudes of teachers and principals using a cross-sectional survey provides data yet to be collected in the state of Oklahoma. Analyses of the data established insight into the quality and perceptions of the learning conversations between principals and teachers while using the Tulsa Model Evaluation instrument.

The purpose of my study was to measure the effect of the Tulsa Model Evaluation instrument within an integrated system of learning for principals and teachers.

Researchers have provided results which point to the significance of integrating structure and politics when implementing school improvement measures (Bryk & Schneider, 2002; Fullan, 2009; Honig et al., 2010). Multiple studies employ pathways models, linking variables of self-reinforcing activities within social interactions in school cultures (Heck & Hallinger, 2010; Hallinger, Heck, & Murphy, 2014; Herman et al., 2016; Honig, 2008; Leithwood & Louis, 2012; Louis et al., 2010; Pierson, 2000). Pathways models are important to isolate specific social conditions which lead to student learning; however, more research is needed to isolate high leverage behaviors which result in improved outcomes for all actors within a learning organization.

Because recent changes to the Elementary and Secondary Education Act (ESEA) will loosen a state's urgency to tie student achievement to principal and teacher evaluation scores, law makers are especially interested in this topic now because of budgetary restrictions. However, because ESEA gives way to open-ended requirements on states to build their own accountability systems, policy makers will want to ensure the most effective measures based on research-based practices are utilized. States must learn to use the tools mandated by law in a new context, a context of a learning organization.

CHAPTER 3

RESEARCH METHODS

Overview

Even before the emphasis on standards-based teacher evaluation during the Obama administration's Race to the Top era (Viteritti, 2012), multiple empirical studies focused on the impact of teacher evaluation (Borko, et al., 2009; Connally & Tooley, 2016; Darling-Hammond, et al., 2012; Dynarski, 2016; Hallinger & Heck, 2014; Kane, et al., 2011; Kane, et al., 2014; Stronge & Tucker, 1999; Taylor & Tyler, 2012; Tuytens & Devos, 2011; Weisberg, et al., 2009). Although the intent of standards-based evaluation instruments is to either provide greater accountability for effective teaching or to provide targeted feedback to teachers about effective teaching practices, results from the research provides modest encouragement that standards-based evaluation tools have a positive impact on creating collaborative learning cultures, on overall school improvement, or for creating conditions for teacher growth.

The purpose of my study was to explore the quality and impact of the mandated implementation of the Tulsa Model Evaluation process in Oklahoma schools through a broad sample of principals and teachers. Using a post-positivists approach, this study employed a cross-sectional survey to measure the difference in perceived quality and impact of the evaluation instrument on conditions for teacher growth. The survey was constructed based on the Teacher Evaluation Profile (Stiggins & Nickel, 1989) and previous dissertation work in Iowa to gather information on the degree to which principals and teachers differ in their understanding of the Tulsa Model Evaluation

instrument's impact on use in six areas: knowledge and understanding of the evaluation system; perceptions regarding usefulness; qualifications and evaluator leadership; quality of data and feedback from principals; impact on professional practices; and personal impressions of the evaluation system on professional conversations.

The Definition of Teacher Evaluation in Oklahoma

While there are certain required components for mandated evaluation systems in Oklahoma, school districts choose between two models: Marzano's Causal Teacher Evaluation Model or the Tulsa Model Evaluation. Evaluation procedures differ within school districts depending on the model evaluation system chosen. The Tulsa Model Evaluation instrument utilizes a framework of 5 domains and 20 dimensions. For this study, teacher evaluation procedures included all steps and tools applied within a participating school district to implement the Oklahoma Teacher Leader Effectiveness Law. This included but was not limited to the following: 1. review of personal learning focus; 2. pre-observation conferences; 3. classroom observations; 4. post-observation conferences; 5. peer observations; 6. examination of lesson plans, materials or other related artifacts including student achievement data; and 7. summative evaluation conferences. When reference is made in the survey to teacher evaluation, it should be understood to encompass any of the procedures followed during the implementation of the Oklahoma Teacher and Leader Effectiveness Law and the Tulsa Model Evaluation within a school district.

The Tulsa Model Evaluation process was developed in Tulsa Public Schools in 2009 in conjunction with research funded by the Bill and Melinda Gates Foundation. (Oklahoma State Department of Education, 2014). According to a document on the

Oklahoma State Department website, the evaluation instrument was included in a 2011 MET Validation Engine Analysis which "confirmed that the Tulsa model measures what matters – that it captures practices that are empirically associated with gains in student achievement." (Tulsa Public Schools, 2017). The process was adapted after the Teacher Leader Effectiveness Law came into effect in 2011 and was fully implemented in 484 schools across the state. For this study 54 school districts were sampled after superintendents in those districts agreed to participate in the study.

Quantitative Paradigm for Understanding Human Behaviors

The basis for this study design lies in quantitative research on school districts around the nation. Evidence suggests school improvement is positively impacted in schools where systems approaches infuse teacher evaluation instruments within a model where climate is measured through specific conditions of a learning organization (Bryk, et al., 2010; Leithwood & Louis, 2012; Heck & Hallinger, 2010; Kane, Kerr, & Pianta, 2014; Honig, 2008).

Using a quantitative statistical approach, the purpose of this study was to explore the effect of the use of the Tulsa Model Evaluation process in Oklahoma schools. By measuring teacher perceptions of not only the evaluation instrument but also principal effectiveness in creating a collaborative context and ability to deliver effective feedback around the instrument, this study explored principal and teacher perceptions as they relate to the effective behaviors employed when creating conditions for teacher growth. Previous research has identified at least four main problems of practice in teacher evaluation:

1. Lack of understanding of the purpose of the evaluation or ineffective principal training and implementation. Principals do not receive adequate professional development for building collaborative cultures in which feedback is a form of coaching. Principals do not see how an evaluation tool fits into a whole system of improvement or part of a learning organization. In addition, principals may lack the skill to properly understand what should be recorded as effective practice (Bryk & Schneider, 2002; Connally & Tooley, 2016; Darling-Hammond, et al., 2012; Fullan, 2009; Hill & Grossman, 2013; Hallinger & Heck, 2014; Hallinger, Heck, & Murphy, 2014; Honig, et al., 2010).
2. Lack of teacher input or feedback. Teachers should be treated as learners and leaders in the system (Darling-Hammond et al., 2012). Without attention to teachers as adult learners who need feedback, mental models, and reflection, effective strategies will not bridge the knowing/doing gap. When teachers do not apply what they have learned about effective practice, student achievement will suffer (Hattie, 2015a; Stronge & Tucker, 1999; Taylor & Tyler, 2012; James & McCormick, 2009).
3. Lack of connection to professional development and professional community. Teacher evaluation instruments work best as a tool within schools as learning organizations. When principals and teachers use evaluation to provide focused conversations about specific strategies, student achievement results are likely to increase (Reeves, 2010; Robinson, Lloyd, & Rowe, 2008; Taylor & Tyler, 2012; Goe, Biggers, & Croft, 2012).

4. Limited range and authenticity in evaluation scores. Authenticity can be defined as validity or legitimacy. An evaluation of scores from several evaluation tools ranks most teachers a 3 on a five-point scale. Most systems use a Likert scale rubric based on standards of effective teaching practice. Principals tend to choose mid-range scores resulting in overall ratings of "effective" or "highly effective" either because they believe the rating scale to be too blunt an instrument to measure effective practice or they believe teachers should not be rated poorly if proper supports are not given for improvement (Borko, Liston, & Whitcomb, 2009; Connally & Tooley, 2016; Dynarski, 2016; Taylor & Tyler, 2012; Weisberg, et al., 2009). The lack of validity is especially problematic for schools attempting to use the evaluation process as a vehicle to measure effective practice as it is linked to school improvement. Although a regulation can become internalized, leading to sustained change, if the regulation is seen to have value (Gagne & Deci, 2005), one way to add value is to provide an environment for proper feedback. If evaluation ratings consistently yield a 3 rating, the value or usefulness of the tool comes into question.

The survey instrument used for this study contained specific questions to measure teacher and administrator shared understanding and perceived value of the Tulsa Model Evaluation process. Survey responses explored teacher and principal perceptions of a common understanding of the evaluation instrument, the quality of feedback from the principal, opportunities for professional growth, and the impact of the context of the

evaluation process to create conditions for teacher growth. Each of these criteria directly relate to the problems of practice addressed in previous research.

Instructional leadership has been acknowledged as having significant effects on student learning; however, research is moving beyond the step-by-step linear models of leadership's direct link to student achievement (Marks & Printy, 2003; Heck & Hallinger, 2010). Some researchers have drawn conclusions that for principals to be effective using teacher evaluation instruments, the principal's expertise in a content area must match the teacher being observed (Hill & Grossman, 2013). These conclusions take the idea of instructional leadership to the extreme.

A quantitative approach has been employed in many studies investigating correlations between teacher evaluation and student achievement. Multiple logic models explaining the pathways relationship between principal, teacher, and community effects on student learning have been developed producing quantitative results (Bryk, et al., 2010; Hallinger, Heck, & Murphy, 2014; Leithwood & Louis, 2012; Heck & Hallinger, 2010; Honig, 2008). Quantitative methods add to the body of research on the topic of integrating learning organization theory and social learning theory approaches to school improvement. Acknowledging the complexity of educational research in the field, quantitative methods emphasize critical thinking and deeper questioning (Henson, Hull, & Williams, 2010). Studies using quantitative methods have a responsibility to check assumptions and report magnitude of results. By understanding the quantified effect of educational interventions, such as teacher evaluations, we can assess what works in education. Quantitative studies move the research forward to better understand the broader impact of policy mandates. A survey approach lends itself to a larger sample size

and highlights specific problems of practice within the day-to-day conditions of a school environment and can form the basis for future research.

Foundations for the Research Design

Foundational research to support a quantitative approach include studies analyzing system levers which influence school improvement. Bryk et al. (2010) conducted extensive studies on whole system improvement of over 100 elementary school sites in Chicago Public Schools. These studies isolated subsystems of supports: leadership as the driver for change; parent-community ties; professional capacity; student-centered learning climate; and instructional guidance. By aggregating each school's data based on support indicators and pairing the data with each school's student achievement data, researchers could determine an overall school organizational capacity score. This research is important to building understanding of leadership and school climate pathways toward teacher effectiveness and increased student learning. Although individual effects of each of the five subsystems is studied, the overall conclusion is the most successful elementary schools employ all five subsystems in total for sustained change. Systems approaches exhibit evidence of success for improving professional practice and improving student learning.

Leithwood and Louis (2012) build on the existing research by employing quantitative and qualitative methods to explore the effects of leadership on student achievement and teacher perceptions of learning. Data collected from 43 states include interviews, surveys, and student achievement scores. Results support principals and teachers building strong professional communities for increased student learning. Of importance to my study is the link between professional community, teacher leadership,

and improved student achievement. Researchers state that professional community is a strong predictor of instructional effectiveness which in turn is associated with increased student learning.

A study by Heck and Hallinger (2010) utilizes reciprocal-effects models to determine the relationship between principal leadership and student achievement in a collaborative environment. Using a sample of 195 elementary schools, measures of leadership process and communication chains make up three models in which pathways of behaviors can be linked to student achievement and school improvement. The results of the study point to the need and the importance of quantitative studies which measure pathways of leadership behaviors and the significance of specific behaviors on a learning organization.

Kane, Kerr, and Pianta (2014) emphasize principals as coaches during the evaluation process and the use of professional development targeted toward specific teaching strategies. In extensive studies using METS data, researchers connect principal feedback to teacher actions. Within a collection of 15 studies, teacher evaluation is linked to student outcomes using not only principal observation data but also Value-Added Measures. Employing a variety of statistical models, researchers use analyses ranging from linear regression to spline regression to delineating thresholds based on observation categories and applying analysis of variance. The argument for quantitative analyses is clear: assessment of teaching performance using reliable data is the only way to determine teacher impact on student learning and to discover actions which yield results. The struggle for reliable data gathered with a valid evaluation tool remains a factor in future quantitative research. In Oklahoma, my study begins to collect data from a

statewide sample of the quality and impact of mandated evaluation processes using quantitative methods.

These studies emphasize the importance of tools codified during the accountability reform movement, such as a standards-based teacher evaluation instrument, being examined for usefulness and impact within a system that supports a culture of learning. Although research in the social sciences is never black and white, as data are refined through reliable instruments and as researchers home in on variables related to high leverage social/political behaviors in schools, quantitative methods are important in gaining understanding of the complexity of getting results in schools.

Questions remain about how principals and teachers within the system utilize the evaluation instrument as a communication device or vehicle to link standards of behavior or instructional strategies to school improvement through professional conversations. Studies focused on a systems approach to school improvement suggest achieving successful student outcomes using an evaluation instrument requires evaluating multiple social and political factors as well as organizational structures within a classroom and a school community (Bryk, et al., 2010; Dufour & Fullan, 2012; Honig, 2008). If we consider a school a complex network of social behaviors and political norms, then we must determine if the evaluation process is assisting the development of professional conversations to achieve better conditions for teacher growth and student results. Or is it cancerous to it? The purpose of my study was to further explore the effectiveness of mandated standards-based evaluations on creating conditions for teacher growth.

Research Design

This study used a cross-sectional survey design to explore perceived quality and impact of the Tulsa Model Evaluation process to create conditions for teacher growth in Oklahoma schools for the 2019-2020 school year. Survey data is best utilized when attempting to explore a phenomenon at a specific point in time (Cohen, et al., 2013). Principals and teachers were asked to complete surveys designed to investigate their perceived understanding of the evaluation process and its relationship to questions about quality of feedback, professional growth, and the context of the evaluation process related to impact. Data collected were analyzed in two phases. Phase I outlined descriptive statistics such as gender, school size, and teacher and principal years of experience. Phase II determined relationships between principal and teacher responses using statistical analyses in the form of a Welch alternative ANOVA. The overarching question for this study was do principals and teachers differ in their perceptions of the quality and impact of the Tulsa Model Evaluation process based on Teacher Evaluation Profile conditions for teacher growth?

Results were analyzed to determine statistically significant effects in six areas: knowledge and understanding of the evaluation system; perceptions regarding usefulness; qualifications and evaluator leadership; quality of data and feedback from principals; impact on professional practices; and personal impressions of the evaluation system on professional conversations.

Previous research supports coaching conversations within schools that employ a systems or learning organization approach (Bryk, et al, 2010; Leithwood & Louis, 2014). Understanding previous work on teacher evaluation systems focused on separate

components of the system, my study attempted to explain the use of the Oklahoma Tulsa Model Evaluation process as it is used within a public school context in real-world school settings. The Tulsa Model Evaluation instrument is based on Danielson's *Framework for Teaching* and is one of two choices schools have for the state's standards-based evaluation instrument as required by the Teacher Leader Effectiveness Law. My study reflects a post-positivist point of view in that I attempt to tease out the variables contributing to positive or negative differences on the perception of the Tulsa Model Evaluation process from teachers and principals who use the model across the state. Post-positivism best suits this quantitative study because the nature of this viewpoint privileges the objective reality of determining the true usefulness of teacher evaluation in Oklahoma (Butin, 2010). Post-positivism mirrors my worldview based on over 30 years as an educator. By nature, educators conduct action-research in classroom environments. Our positions, close to the daily dynamics of the social complexity of learning, require continuous reflection. The ever-changing and fluid aspects of learning, especially in American classrooms with 20 or more students, depend on a practitioner's ability to adjust based on observations and data, and this positionality is important to knowing ourselves, so we may understand the larger world of the profession (Salzman, 2002). I acknowledge there is an art in teaching, a craft to be shaped over years of experience; however, there are also outcomes which can be analyzed using scientific methods. Teachers understand for every action there is an equal and opposite reaction. Understanding the underlying causes for results through experimentation is the holy grail pursuit of our profession.

This study originated through an attempt to understand a learning organization's relationship to student outcomes and overall school improvement. In addition, I am concerned with the impact of politically driven school reform mandates on improving professional practice. The theoretical framework for this study was built on the work of behavioral and social scientists who endeavor to understand learning organization theory and systems thinking in a public school context. It employs the lenses of both theories based on previous empirical studies linking effects of principal or teacher leadership to school improvement (Honig, 2008; Hallinger, Heck, & Murphy, 2014; Louis, et al., 2010). The survey was derived from the work of Stiggins and Nickel (1989) and Stiggins and Duke (1989) in which the Teacher Evaluation Profile was developed based on 44 attributes to determine optimal conditions for teacher growth. Using the Teacher Evaluation Profile as a foundation, multiple dissertation studies applied versions of the Teacher Evaluation Profile survey questions to explore the perceived value and usefulness of standards-based evaluation instruments in the wake of states across the country competing for Race to the Top grants (Huckstadt, 2011; Schumacher, 2004; Towe, 2012). Cross-sectional survey data are gathered in the form of a Likert scale and used for quantitative analyses. Permission was granted from Dr. Kim Huckstadt in 2018 to modify the survey for the Oklahoma study.

Research Questions

The purpose of this study was to explore the perceived quality and impact of the Tulsa Model evaluation process on conditions for teacher growth when the evaluation instrument is used as a part of an overall system. Using a Welch alternative analysis of variance, the study measured data collected through a cross-sectional survey of teachers

and principals in the state of Oklahoma. The following research questions were addressed:

1. Do principals and teachers differ in their perceptions of knowledge and understanding of the Tulsa Model Evaluation process and if so, to what degree?
2. Do principals and teachers differ in their perceptions of the usefulness of the Tulsa Model Evaluation process and if so, to what degree?
3. Do principals and teachers differ in their perceptions of the qualification of the evaluator and if so, to what degree?
4. Do principals and teachers differ in their perceptions of the quality of data and feedback in the Tulsa Model Evaluation process and if so, to what degree?
5. Do principals and teachers differ in their perceptions of the impact on professional practices for the Tulsa Model Evaluation process and if so, to what degree?
6. Do principals and teachers differ in their perceptions of the context of professional conversations during the Tulsa Model Evaluation process and if so, to what degree?

Each question was formed based on previous studies which examine the dynamics of teacher evaluation in schools and the ability of evaluation instruments to create conditions for teacher growth. These questions accomplished building additional understanding of the dynamics of principal and teacher interactions using the Tulsa Model Evaluation process as a vehicle for effective instructional practice. This study

expanded on previous studies using quantitative research by explaining effects of collaborative conversation on conditions for teacher growth.

The main claim of my study is professional growth and agency improves when principals and teachers understand how the Tulsa Model Evaluation process fits within a system functioning as a learning organization. The following hypotheses were addressed:

H_0 There is no significant difference in principal and teacher perceptions of knowledge and understanding of the Tulsa Model Evaluation process.

H_1 There is no significant difference in teacher and principal perceptions of the usefulness of the Tulsa Model Evaluation process.

H_2 There is no significant difference in principal and teacher perceptions of the qualification of the evaluator.

H_3 There is no significant difference in principal and teacher perceptions of the quality of data and feedback in the Tulsa Model Evaluation process.

H_4 There is no significant difference in principal and teacher perceptions of the impact on professional practices for the Tulsa Model Evaluation process.

H_5 There is no significant difference in principal and teacher perceptions of the context of professional conversations during the Tulsa Model Evaluation process.

Participants and Sampling

Participants of this study included teachers and principal evaluators across the state of Oklahoma who utilize the Tulsa Model Evaluation process. The sample included teachers from Pre-K through grade 12 and who teach in various subject areas.

Demographic information collected included gender, number of years' experience, and the size of the school. According to the Oklahoma State Department of Education (2018), there were 41,047 certified teachers in the state in the 2017-2018 school year. Currently there are 551 schools using the Tulsa Model Evaluation system and 34 districts in the state not using the system. Superintendents were emailed asking permission for their district to be included in the survey. Of 551 schools asked to participate in this study, 54 school district superintendents agreed to allow their principals and teachers to participate. Teachers and principals were then forwarded a link to their perspective surveys by their district administration. After consenting to participate in the survey, participants completed the survey using a web-based survey instrument. A sample return of 10% can be expected from a web-based survey (Cohen, et al., 2013). Eliminating teachers and principals who do not use the Tulsa Model Evaluation, a reasonable parameter for the statewide sample can be established at 2,500 teachers and principals. For rigorous study design, the total collected sample needed to include over 484 participants to ensure a confidence level of 95% with a confidence interval of 4% (Cohen, et al., 2013). The total number of teacher participants included 718 (n=718). The total number of principal evaluator participants included 137 (n=137).

Instrumentation

This study employed a cross sectional survey design. The survey instrument was adapted from the Teacher Evaluation Profile and was used through permission based on previous dissertation work by Dr. Kim Huckstadt. Huckstadt (2011) conducted a study of Iowa teachers and principals to determine their perceptions of the genuine utility and value of its standards-based teacher evaluation system. My survey collected teacher

perceptions of principal effectiveness and overall perceptions of the standards-based evaluation instrument. Utilizing a Likert scale, the survey gathered opinions from teachers and principals using parallel questions in the following areas: demographic information, knowledge and understanding of the evaluation system, perceptions regarding implementation, qualifications and evaluator training, data and feedback, impact on professional practice, and personal impressions of the standards-based evaluation system. Overall, the survey included 7 limited response demographic items to determine gender, years of experience, school size, ethnicity, grade level assignment, and school use of collaborative teams. The bulk of the survey included 48 Likert-scale items with a scale from 1 Strongly Disagree to 5 Agree, and the survey was delivered through a web-based survey system.

The survey used in my study has a history in the Teacher Evaluation Profile designed by Stiggins and Nickel (1989) and Stiggins and Duke (1988). The intent of the Teacher Evaluation Profile is to determine the relationship between teacher evaluation and professional growth (Stiggins & Nickel, 1988). The Teacher Evaluation Profile establishes the context that conditions for teacher growth can be improved within certain established surroundings. The theory of action (Table 1.1) presented is the structural or contextual environment within the system of teacher evaluation must hold certain criteria to produce growth.

Reliability and validity are addressed via modification of the Teacher Evaluation Profile instrument in other studies (Huckstadt, 2011). The survey has been used in previous studies to measure teacher and principal perceptions of standards-based evaluation frameworks in Wisconsin and Iowa. Initial use of the survey in Wisconsin

measured the effect of a Danielson-based evaluation model on teaching practices utilizing expectancy theory (Schumacher, 2004). The survey was modified and combined with the Teacher Evaluation Profile to study the perceived utility of a standards-based evaluation system in Iowa (Huckstadt, 2011). For my study, permission to use the survey was granted to explain the perceived impact and difference in perceptions of the Tulsa Model Evaluation process in Oklahoma schools in the areas of shared understanding, quality of feedback, opportunities for professional growth, and teacher and principal impressions of the evaluation process.

The survey used for the Oklahoma study was modified from the one used in Iowa due to the differences in the two evaluation systems. Although both systems are standards-based, terminology varies slightly. The Tulsa Model Evaluation instrument uses 5 domains and 20 dimensions to capture effective teaching practice. The Iowa system incorporates 8 standards and 42 criteria. Both systems are based on Danielson's model of effective teaching. The Oklahoma survey included the same demographic questions as the Iowa study with the addition of a question "My school uses collaborative teams or professional learning communities to inform instruction and make decisions about student learning," which elicited a yes or no response. The reason for the addition of this question is to determine which participants in the sample identify themselves as part of a collaborative learning culture.

The Oklahoma survey retained the same section titles as the Iowa survey. An analysis of common language used in the Iowa survey and the TEP was conducted to create the Oklahoma survey and ensure the integrity of the Iowa survey remained intact.

Each section contained questions designed to be answered using a Likert scale and to gather perceptions on the following topics:

1. Knowledge and Understanding of the Evaluation System
2. Perceptions Regarding Implementation
3. Qualifications and Evaluator Training
4. Data and Feedback
5. Impact on Professional Practice
6. Personal Impressions of the Tulsa Model Evaluation System

Previous validation of the survey was done through a pilot test of Iowa teachers and principals (Huckstadt, 2011). Changes to language in the instrument do not impact the validity and reliability of the instrument.

Research Procedures

The data collected were analyzed in two phases. Phase I determined the descriptive statistics including gender, school size, and number of years' teaching experience. Phase II employed a Welch alternative ANOVA to determine the variance and the degree of difference between perceptions of principals and teachers in the effectiveness of the Tulsa Model Evaluation process. The between group variances of perceptions of quality of feedback, professional growth, and professional agency based on understanding of the evaluation process were measured between teachers and principal evaluators. Survey questions were designed to capture the climate of each school site and included language to measure support, feedback, and personal satisfaction (Huckstadt, 2011).

Data Analysis

Data were collected using a web-based survey instrument. Phase I and Phase II analyses were conducted using SPSS to describe the sample population and compare means. Data were analyzed to determine sum totals for both principals and teachers, frequency distribution, mean, and standard deviation. Phase II explored the variance of means between principals and teachers. Data were sorted by category based on criteria established in the research questions: understanding, quality of feedback, professional growth, and agency. A Welch alternative ANOVA and Cohen's *d* were used to determine if perceptions between principals and teachers were significantly different. A guide to effect size using Cohen's *d* establishes 0.2 as small effect, 0.5 as medium effect, and .0.8 as large effect (Cohen, 1992).

Descriptive statistics assist in determining control variables such as teacher and principal years of experience and grade level. Reporting descriptive statistics provide necessary information about means, standard deviations, and ranges (Creswell, 2014, American Educational Research Association, 2006). However, the scope of this study does not include measurement of discreet variable effects.

Because the sample of principal evaluators (n=137) and sample teachers (n=720) was uneven, a Welch alternative ANOVA was used to determine the perceived effect of quality of feedback, professional growth, and professional agency. Cohen's d effect size for this sample was at $\alpha = .05$ requires $N = 64$ for a power of .80 to be significant (Cohen, 1992). Assumptions of normality and homoscedasticity were addressed using SPSS. Because of the unevenness of the sample, a Welch alternative ANOVA assisted to

mitigate Type I errors. Variation was found to be within normal range. Within the data set the following process was used:

1. Clean data, replace missing values, and prepare for analyses
2. Run and extract descriptive data for the population of teachers and principals
3. Run and extract ANOVA for 6 comparison categories of
    a. knowledge and understanding of the evaluation system;
    b. perceptions regarding usefulness;
    c. qualifications and evaluator leadership;
    d. quality of data and feedback from principals;
    e. impact on professional practices; and
    f. personal impressions of the evaluation system on professional conversations.

Limitations

In any cross-sectional survey study design challenges may occur with the study sample. Insufficient sampling, poor quality of survey questions, lack of response, and reporting bias may all hinder results (Cohen, et al., 2013). To address these limitations a large sample size allows for aggregate data to be categorized based on complete responses. Research questions were closely aligned to the research literature and survey questions to ensure participant responses from teachers and principals addressed research questions in a deliberate manner (Butin, 2010). Table 1.1 aligned the study design to the problems of practice, research questions, Teacher Evaluation Profile attributes, and the survey questions.

Assumptions for ANOVA were checked to ensure the quality of the data was suitable for study. A Welch alternative ANOVA addressed the lack of homogeneity in the sample (Tomarken & Serlin, 1986; Jan & Sheih, 2013).

Summary

The purpose of this study was to measure the perceptions of principals and teachers on the effectiveness of the successful implementation in terms of quality and impact of the Tulsa Model Evaluation process across the state of Oklahoma. The overarching goal of the study was to explore the extent and degree principals and teachers differ in their perceptions of the impact of the Tulsa Model Evaluation. An extensive survey study design is best suited to analyze differences in principal and teacher perceptions. Survey data from teachers and principals were collected from 54 schools in the state utilizing the Tulsa Model Evaluation process. In Phase I of the analyses, demographic data, such as grade level taught, years of experience, teaching assignment, and gender, were gathered. Phase II of the analyses used a Welch alternative ANOVA to determine the degree of differences in the perceptions of principal evaluators and teachers. Results from the survey and analyses of the date are presented in Chapter 4.

CHAPTER 4

RESULTS, ANALYSIS, AND DISCUSSION

Tulsa Model Evaluation Overview

The Tulsa Model Evaluation tool was developed in Tulsa Public Schools in 2009 in conjunction with research funded by the Bill and Melinda Gates Foundation. (Oklahoma State Department of Education, 2014). According to a document on the Oklahoma State Department website, the evaluation instrument was included in a 2011 MET Validation Engine Analysis which "confirmed that the Tulsa model measures what matters – that it captures practices that are empirically associated with gains in student achievement." (Tulsa Public Schools, 2017).

The Tulsa Model Evaluation as a standards-based evaluation process is based on Charlotte Danielson's Framework for Teaching (The Danielson Group, 2013; The Oklahoma State Department of Education, 2014). The process of evaluation includes a series of events to be conducted by certified administrators who have completed evaluation training. The handbook for administrators states the system values "feedback, analysis, and refinement" (Tulsa Public Schools, 2015). These evaluation events include at least two observations lasting 20 to 30 minutes minimum, a summative evaluation, and conferences between observer and teacher (Tulsa Public Schools, 2015). Modest research has been completed on the use of the Tulsa Model Evaluation across the state of Oklahoma. A survey design dissertation conducted by Boothe in 2019 gathered some insight into principals' perceptions of the use of the Tulsa Model compared to other evaluation instruments. The results indicated of the 347 Oklahoma principals responding

to the survey, 96% employ the Tulsa Model Evaluation to evaluate teachers. When asked if principals believed the evaluation model they use is helpful in providing teachers with appropriate feedback to improve student learning, over 75% of principals said they agree with that statement (Boothe, 2019). However, research from teachers and evaluating principals comparing perceptions of use has not been conducted across schools in Oklahoma.

Study Overview

Utilizing a cross-sectional survey, this study explored the perceptions of principal evaluators and teachers in Oklahoma on the quality and impact of the Tulsa Model Evaluation process which has been mandated across the state as a component of the Teacher Leader Effectiveness Law. A broad representative sample was collected by first receiving permission from superintendents whose districts utilize the Tulsa Model Evaluation system. Once permission was given, a survey was sent to principal evaluators and teachers who self-selected to complete 44 questions through a web-based survey platform. The survey gathered data in six key areas based on previous similar studies: knowledge and understanding of the evaluation system; perceptions of quality of the process; perceptions of the qualifications of the evaluator; the quality of data and feedback; perceived impact of the evaluation system on professional practice; and impressions of the evaluation context for professional conversations (Huckstadt, 2011; Stiggins and Duke, 1988).

Survey questions were derived from previous research designed to explore the usefulness of standards-based teacher evaluations. Huckstadt (2011) provided permission to replicate the survey used in his study on the impact of standards-based teacher

evaluation in Iowa. The survey utilized in this study has foundations in earlier research conducted around the Teacher Evaluation Profile developed by Stiggins and Duke (1988). Reliability and validity tests were completed, resulting in Cronbach's Alpha scores of .963 for the teacher survey and .957 for the administrator survey (Huckstadt, 2011). The survey's internal consistency for Cronbach's Alpha is considered in the acceptable range of measurement error (Tavakol and Dennick, 2011). Questions from my survey were derived with permission from the Huckstadt study and modified slightly to provide an Oklahoma context for district using the Tulsa Model Evaluation.

The study was conducted in two phases. Phase I described the sample of Oklahoma principals and teachers utilizing the Tulsa Model Evaluation instrument. Participants included a broad sample of 134 principals (n=134) and 718 teachers (n=718) from 54 school districts.

Phase II included statistical analyses using a Welch alternative ANOVA to determine statistical significance and to what degree principal evaluators and teachers differ in response to six research questions:

1. Do principals and teachers differ in their perceptions of knowledge and understanding of the Tulsa Model Evaluation process and if so, to what degree?
2. Do principals and teachers differ in their perceptions of the usefulness of the Tulsa Model Evaluation process and if so, to what degree?
3. Do principals and teachers differ in their perceptions of the qualification of the evaluator and if so, to what degree?

4. Do principals and teachers differ in their perceptions of the quality of data and feedback in the Tulsa Model Evaluation process and if so, to what degree?
5. Do principals and teachers differ in their perceptions of the impact on professional practices for the Tulsa Model Evaluation process and if so, to what degree?
6. Do principals and teachers differ in their perceptions of the context of professional conversations during the Tulsa Model Evaluation process and if so, to what degree?

This study builds on the research from previous studies on the quality and impact of standards-based evaluations which have been mandated across the nation in response to the Race to the Top grant competitions.

Previous quantitative research around effective teaching and school improvement points to student learning being positively impacted in schools where systems approaches infuse teacher evaluation tools within a model where climate is measured through specific conditions of a learning organization (Bryk, et al., 2010; Leithwood & Louis, 2012; Heck & Hallinger, 2010; Kane, Kerr, & Pianta, 2014; Honig, 2008).

The quantitative approach used in my study adds to the research by measuring teacher perceptions of not only the evaluation instrument, but also principal effectiveness in creating a collaborative context and ability to deliver effective feedback around the instrument. This study explores principal and teacher perceptions as they relate to the effective behaviors to create conditions for teacher growth outlined specifically in the research presented in the Teacher Evaluation Profile (Stiggins and Nickel, 1989).

The results presented below discuss how the assumptions of ANOVA were met as well as description of demographic statistics for the sample. In addition, variances are addressed for each of the six areas explored through the research questions.

Assumptions of ANOVA

Because the assumption of homogeneity was not met using the large sample in this study, a Welch alternative ANOVA is used to control for Type I error. The Welch test has been used in multiple empirical studies to counter the effects of heteroscedasticity (Tomarken and Serlin, 1986; Jan and Sheih, 2013). An additional step to test the assumption of homogeneity was completed. The results of the multiple analysis of covariance can be seen in Table 4.1 below.

A multiple analysis of Covariance (MANCOVA) was conducted as part of the assumption of homogeneity of regression. A Box Test of Equality result as presented below is significant, therefore, the Pillai's Trace statistic is used.

Table 4.1

Box's Test of Equality of Covariance Matrices[a]

Box's M	339.753
F	15.861
df1	21
df2	146685.216
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + Group + Gender + Race + Group * Gender + Group * Race + Group * Gender * Race

Results

Phase I included an analysis of demographic data. Participants were asked to respond to a series of demographic questions to determine years of experience, district size, current teaching or principal grade level assignment, gender, race, and the participants' school use of collaborative teams.

Within the sample of principals and teachers a summary of demographic data for experience is shown in Table 4.2. Most principals in the sample have between two to ten years of experience as an administrator, whereas most teachers in the sample have 21 or more years of experience.

Table 4. 2

Principal and Teacher Years of Experience as a Percentage of the Sample

	Principals	Teachers
This is my first year	8.3	2.6
2 to 5 years	24.8	15.9
6 to 10 years	24.8	17.9
11 to 15 years	17.9	16.8
16 to 20 years	11.7	14.6
21 or more years	12.4	32.2

Note. Principals n = 134, Teachers n = 718

The broad sample contained principals and teachers from various size school districts as shown in Table 4.3. Most principals, 35.2%, in the sample work in mid-size districts with enrollment between 1,000 to 2,500 students, and 24.4% of teachers

responded they work in the same size districts. Most teachers in the sample work in districts of over 10,000 students with 32.5% responding they work in a large district. Larger districts are least represented in the principal sample with only 12.4% of the respondents saying they work in a large district.

Table 4. 3

Principal and Teacher District Size as a Percentage of the Sample

	Principals	Teachers
Less than 400	10.3	3.7
400 to 1,000	9.0	9.3
1,000 to 2,500	35.2	24.4
2,500 to 5,000	21.4	17.9
5,000 to 9,000	11.7	12.4
Over 10,000	12.4	32.5

Note. Principals n = 134, Teachers n = 718

Table 4.4 shows principal and teacher grade level assignments where most of the teacher sample represents grades 9 through 12 at 30.9% and principals represent grades 2 through 6 at 28.3%.

Table 4. 4

Principal and Teacher Grade Level Assignment as a Percentage of the Sample

	Principals	Teachers
Pre-K through Grade 1	13.8	12.0
Grades 2 through 5 or 6	28.3	28.0

Grades 6 through 8	20.0	22.9
Grades 9 through 12	22.1	30.9
K through 12	15.9	6.1

Note. Principals n= 134, Teachers n = 718

Overall, participants perceived their districts utilize some form of collaborative culture in their work. When asked if their district uses collaborative teams or professional learning communities, 93.1% of principals responded "Yes" and 6.9% of principals responded "No". Similarly, 90.2% of teachers responded "Yes" and 9.8% of teachers responded "No."

Principals were also asked to categorize the number of evaluations they conducted in the 2019-2020 school year. Table 4.5 describes the numbers of evaluations for principals surveyed. Most principals at 37.2% say they conducted more than 20 teacher evaluations in the school year.

Table 4. 5

Number of Evaluations Conducted by Principals Each Year as a Percentage of the Sample

1 to 4	13.1
5 to 10	13.1
11 to 15	19.3
16 to 20	17.2
More than 20	37.2

Note. Principals n = 134

Teachers were asked to state their latest evaluation year with 81.6% of teachers responding their latest evaluation was completed in the 2018-2019 school year. Because the survey was given during the fall of 2019, teacher responses correspond with the Tulsa Model Evaluation cycle. The cycle would not be complete for the 2019-2020 school year until the spring of 2020. Teacher responses to this question establish most of the teacher sample had completed the latest cycle possible.

In Phase II, a one-way Analysis of Variance (ANOVA) was used to determine the differences in perceived quality and impact of the Tulsa Model Evaluation process between principals and teachers based on six criteria: 1. knowledge and understanding of the evaluation system; 2. perceptions regarding usefulness; 3. qualifications and evaluator leadership; 4. quality of data and feedback from principals; 5. impact on professional practices; and 6. personal impressions of the evaluation system on professional conversations. Table 4.6 shows the demographic information for the sample, and Table 4.8 shows descriptive statistics for the two groups of respondents based on the six tested criteria.

Table 4. 6

Sample Demographic Information

	Administrators (n)	Teachers (n)	% of Sample
Gender			
Female	77	579	76.9
Male	57	139	23.0
Ethnicity			

White	116	611	84.9
African American	3	3	0.7
American Indian/Alaska Native	11	47	6.8
Asian	0	1	0.1
Native Hawaiian/Pacific Islander	0	0	0.0
Multi-Racial/Other	7	57	7.5

Note. Gender N = 852, Ethnicity N = 856

Based on the unequal sizes of the sample and the *Levene's F* test result, which revealed that the homogeneity of variance assumption was not met ($p < .001$), the conservative *Welch's F* test was used. Table 4.7 illustrates the equality of means in the sample. No significant differences in means were found; therefore, the assumption of homogeneity in variance is met.

Table 4. 7

Results for Robust Tests of Equality of Means

		Statistic[a]	df1	df2	Sig.
Knowledge and Understanding	Welch	152.479	1	344.653	.000
Usefulness	Welch	133.180	1	315.982	.000
Evaluator Leadership	Welch	70.291	1	418.658	.000
Quality of Data	Welch	122.995	1	479.419	.000
Impact on Practice	Welch	65.633	1	216.508	.000
Personal Impressions	Welch	97.646	1	228.964	.000

After employing a Bonferroni correction, an alpha level of .05 was used for all subsequent analyses. The one-way ANOVA of the differences in perceived quality and impact of the Tulsa Model Evaluation instrument between principals and teachers revealed statistically significant main effects for all tested six criteria as follows:

Knowledge and understanding of the evaluation system

Welch's F (1, 344.65) = 152.45, $p < .001$, indicating that principals and teachers did not have the same average score on their knowledge and understanding of the evaluation system. The estimated omega squared ($\omega^2 = .15$) indicated that approximately 15% of the total variation in knowledge and understanding of the evaluation system is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 4.73, SD = 0.55) had a significantly higher average score on their knowledge and understanding of the evaluation system than teachers (M = 3.98, SD = 1.02).

Perceptions regarding usefulness

Welch's F (1, 315.98) = 133.18, $p < .001$, indicating that principals and teachers did not have the same average score on their perceptions regarding the usefulness of the valuation system. The estimated omega squared ($\omega^2 = .14$) indicated that approximately 14% of the total variation in the usefulness of the evaluation system is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 4.57, SD = 0.52) had a significantly higher average score on their perceptions regarding the usefulness of the evaluation system than teachers (M = 3.90, SD = 0.93).

Qualifications and evaluator leadership

Welch's F (1, 418.66) = 70.29, $p < .001$, indicating that principals and teachers did not have the same average score on their perceptions of the qualifications and leadership of the evaluator. The estimated omega squared ($\omega^2 = .08$) indicated that approximately 8% of the total variation in perceptions of the qualifications and leadership of the evaluator is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 4.66, SD = 0.39) had a significantly higher average score on their perceptions of the qualifications and leadership of the evaluator than teachers (M = 4.23, SD = 0.91).

Quality of data and feedback from principals

Welch's F (1, 479.42) = 122.99, $p < .001$, indicating that principals and teachers did not have the same average score on their perceptions of the quality of data and feedback from principals. The estimated omega squared ($\omega^2 = .14$) indicated that approximately 14% of the total variation in perceptions of the quality of data and feedback from principals is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 4.64, SD = 0.38) had a significantly higher average score on their perceptions of the quality of data and feedback from principals than teachers (M = 4.05, SD = 0.99).

Impact on professional practices

Welch's F (1, 216.51) = 65.63, $p < .001$, indicating that principals and teachers did not have the same average score on their perceptions of the impact of the evaluation system on professional practices. The estimated omega squared ($\omega^2 = .08$) indicated that

approximately 8% of the total variation in perceptions of the impact of the evaluation system on professional practices is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 3.89, SD = 0.82) had a significantly higher average score on their perceptions of the impact of the evaluation system on professional practices than teachers (M = 3.17, SD = 1.14).

Personal impressions of the evaluation system on professional conversations

Welch's F (1, 228.96) = 97.65, $p < .001$, indicating that principals and teachers did not have the same average score on their personal impressions of the evaluation system on professional conversations. The estimated omega squared ($\omega^2 = .12$) indicated that approximately 12% of the total variation in personal impressions of the evaluation system on professional conversations is attributable to differences between the respondent's role as a teacher or principal. Based on the results given in Table 4.8, principals (M = 4.19, SD = 0.77) had a significantly higher average score on their personal impressions of the evaluation system on professional conversations than teachers (M = 3.34, SD = 1.16).

Table 4. 8

Descriptive Statistics – Means and Standard Deviations

	n	M	SD
Principals			
Knowledge and Understanding	137	4.7	0.6
Usefulness	131	4.6	0.5

Evaluator Leadership	124	4.7	0.4
Quality of Data	120	4.6	0.4
Impact on Practice	118	3.9	0.8
Personal Impressions	115	4.2	0.8
Teachers			
Knowledge and Understanding	720	3.9	1.0
Usefulness	671	3.9	0.9
Evaluator Leadership	638	4.2	0.9
Quality of Data	614	4.1	0.9
Impact on Practice	599	3.2	1.1
Personal Impressions	587	3.3	1.2

Note. Knowledge and Understanding N = 857, Usefulness N = 802, Evaluator Leadership N = 762, Quality of Data N = 734, Impact on Practice N = 717, Personal Impressions N = 702

Analysis

The purpose of this study was to explore the differences in perceptions between principals and teachers in the quality and impact of the Tulsa Model Evaluation process. Principals and teachers were asked to answer separate surveys with parallel items. Survey items aligned to surveys in previous studies on the usefulness of standards-based teacher evaluation instruments (Schumacher, 2004; Huckstadt, 2011), and based on the reliability and validity studies of the Teacher Evaluation Profile (Stiggins and Nickel, 1989; Huckstadt, 2011). The intention of my study was to gather baseline information about the variation between principal evaluators and teachers on the quality and impact of the Tulsa Model Evaluation process to create conditions for teacher growth. Future research could

explore the ability of principals as evaluators to create conditions for teacher growth based on correlations between school size, teacher experience, or gender differences using the available survey data.

The largest variance at 15% in principal and teacher perceptions is in knowledge and understanding of the evaluation system. Table 4.9 provides insight into the survey responses between principals and teachers to explain the variance. When asked if the participant has a thorough understanding of the 5 domains and 20 dimensions within the Tulsa Model Evaluation, only 36% of teachers agreed with that statement compared to 73% of principals. A wider gap can be seen in the responses to the statement about a thorough understanding of the intended purposes of the evaluation system. Teachers responded they understood impact with only 36% agreeing. Principals responded to the same statement with 80% in agreement.

Table 4. 9

Principal and Teacher Responses of Knowledge and Understanding as a Percentage of the Sample

Describe your knowledge and understanding of the Tulsa Model Evaluation system.										
	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
I have a thorough understanding of the 5 domains and 20 dimensions that formulate the basis of the teacher evaluation system. I know what they are, and I know how they relate to my teaching responsibilities.	73.0	36.0	24.8	42.5	0.0	8.9	0.0	6.8	1.5	5.8

I have a thorough understanding of the evaluation tool and procedures utilized in my district.	75.2	43.5	21.9	38.1	0.0	8.8	0.0	5.8	2.2	3.9
I have a thorough understanding of the intended purposes of the Tulsa Model Evaluation system.	80.3	36.3	16.8	36.4	0.7	12.1	0.7	8.6	0.7	6.7

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Responses between principals and teachers also differ in the area of quality of data and feedback as seen in Table 4.10. Data and feedback are critical components of successful evaluation systems (Stiggins and Nickel, 1989; Stiggins and Duke, 1988). In response to the statement about the quality of feedback teachers receive during the evaluation process being directly linked to teaching practice, 45% of teachers responded they agreed while 70% of principals responded they agreed with the same statement.

Table 4.10

Principal and Teacher Perceptions Regarding Quality of Data and Feedback as a Percentage of the Sample

Describe your perceptions regarding the quality of the data collected and the feedback provided during your most recent evaluation.

	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
The feedback I receive from my evaluator is clearly linked to the Tulsa Model domains and dimensions.	61.7	52.1	35.0	32.6	3.3	8.0	0.0	4.7	0.0	2.6
The feedback I receive from my evaluator is an	70.8	45.6	26.7	30.8	0.8	12.7	0.8	6.0	0.8	4.9

accurate reflection of my teaching practice.										
The feedback I receive from my evaluator includes discussion and connection to student learning and monitoring the progress of the students in my classes.	73.3	48.0	25.8	29.0	0.0	13.4	0.0	6.2	0.0	3.4
The feedback I receive from my evaluator is valuable information I have used to improve my teaching practice.	67.5	44.3	30.8	27.4	0.8	13.0	0.0	10.3	0.8	5.0
The feedback I receive from my evaluator is supported by opportunities to participate in professional development or training designed to improve my teaching practice.	58.3	40.6	35.8	27.4	2.5	15.5	3.3	10.9	0.0	5.7

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Table 4.11 demonstrates the differences in survey responses around impact of the evaluation process on professional practices. The variance for this area is one of the lowest at 8%. When asked if the alignment of expectations for teachers, professional development, and student achievement has improved significantly because of the evaluation process, only 23.7% of principals and 13.9% of teachers agreed with this statement.

Table 4. 11

Principal and Teacher Perceptions of Outcomes of the Evaluation Process as a Percentage of the Sample

Describe your perceptions of the outcomes of the evaluation process related to professional development activities and professional practices.										
	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
I have changed instructional methodologies for the benefit of my students as a result of participating in the Tulsa Model Evaluation.	37.3	20.5	47.5	33.1	9.3	16.7	5.1	18.5	0.8	11.2
I can show data that indicates the achievement level of my students has improved as a result of my participation in the Tulsa Model Evaluation process.	17.8	15.4	38.1	24.5	28.0	23.7	13.6	22.5	2.5	13.9
As a result of my participation in the Tulsa Model Evaluation process, I now spend a greater amount of time reflecting on my teaching practice.	32.2	15.9	40.7	31.1	17.8	19.4	7.6	19.5	0.8	14.2
As a result of the implementation of the Tulsa Model Evaluation process, the alignment between expectations for teachers, professional development practices and student achievement goals has improved significantly in my school.	23.7	13.9	44.1	31.2	21.2	24.0	9.3	17.7	1.7	13.2
The Tulsa Model domains and dimensions have provoked discussion regarding effective teaching practices among teachers at my school.	35.6	17.5	35.6	27.7	17.8	22.9	7.6	19.2	3.4	12.7
The Tulsa Model Evaluation system has challenged me to improve my teaching practices.	42.4	23.5	40.7	28.9	7.6	18.7	6.8	17.9	2.5	11.0

Participation in the Tulsa Model Evaluation process is valuable to me as a professional development activity.	33.9	19.9	44.1	26.2	11.0	20.9	8.5	16.0	2.5	17.0
I have gained a great deal of personal satisfaction as a result of participating in the Tulsa Model Evaluation process.	22.0	13.9	39.0	25.7	18.6	21.2	14.4	19.4	5.9	10.9

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Teacher and principal perceptions of the overall impact of the evaluation system on professional conversations reveal a variance of 12%. Survey responses as shown in Table 4.12 provide examples to explain the variance between principals and teachers. When teachers were asked to agree with the statement, *I am generally supportive of the Tulsa Model Evaluation system*, only 21% agreed compared to 57% of principals.

Table 4.12

Principal and Teacher Impressions of the Tulsa Model Evaluation System as a Percentage of the Sample

Describe your perceptions regarding the impact of the Tulsa Model Evaluation system on professional conversations between teachers and administrators in your school.

	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
The Tulsa Model domains and dimensions have encouraged more in-depth discussions between administrators and teachers regarding effective teaching practices.	36.5	19.8	43.5	34.4	12.2	20.4	7.0	15.5	0.0	9.9
Working with my administrator through the Tulsa Model Evaluation system	34.2	24.4	48.2	32.2	14.0	19.6	3.5	14.1	0.0	9.7

helped me to learn how I can improve my teaching practices.										
I believe by working with my administrator and by meeting the expectations identified in the Tulsa Model Evaluation, my instructional practices will progress to a higher level.	39.1	24.9	47.0	30.8	8.7	20.6	5.2	14.5	0.0	9.2
The professional growth through administrator and teacher discussions I have gained from participation in the Tulsa Model Evaluation process make it worth the time and effort invested.	37.4	19.9	40.0	27.4	15.7	22.0	5.2	16.4	0.9	14.3
Participating in the Tulsa Model Evaluation motivates me to improve my teaching performance.	38.3	20.8	36.5	29.0	15.7	17.9	7.0	18.2	1.7	14.1
I have learned a lot from participating in the Tulsa Model Evaluation process.	41.7	15.8	35.7	26.9	13.0	23.2	7.8	19.1	0.9	15.0
I am generally supportive of the Tulsa Model Evaluation system.	57.4	21.6	27.8	37.3	10.4	17.0	2.6	10.6	0.9	13.5
I support the continued use the Tulsa Model Evaluation system in Oklahoma.	55.7	22.8	29.6	33.2	8.7	17.7	3.5	10.4	1.7	15.8

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Based on the results of this study and general observations of the survey data, the null hypothesis can be rejected on all six research questions. The variance of principals

and teachers in their perceptions of the Tulsa Model Evaluation process is significant in each area to varying degrees.

In the area of knowledge and understanding of the Tulsa Model Evaluation system, principals and teachers have a 15% degree of variance depending on their role with principals responding in agreement at a higher average. The null hypothesis can be rejected based on the significant statistical difference in means.

In the area of perceptions regarding usefulness of the Tulsa Model Evaluation system, principals and teachers have a 14% degree of variance depending on their role with principals responding in agreement at a higher average. The null hypothesis can be rejected based on the significant statistical difference in means.

In the area of qualifications and evaluator leadership within the Tulsa Model Evaluation system, principals and teachers have an 8% degree of variance depending on their role with principals responding in agreement at a higher average. The null hypothesis can be rejected based on the significant statistical difference in means.

In the area of data and feedback within the Tulsa Model Evaluation system, principals and teachers have a 14% degree of variance depending on their role with principals responding in agreement at a higher average. The null hypothesis can be rejected based on the significant statistical difference in means.

In the area of impact on professional practices within the Tulsa Model Evaluation system, principals and teachers have a 12% of variance depending on their role with principals responding in agreement at a higher average. The null hypothesis can be rejected based on the significant statistical difference in means.

Summary

This chapter provided results of a Welch alternative ANOVA analyses of data collected through a cross-sectional survey of teacher and principal perceptions of the quality and impact of the Tulsa Model Evaluation system in 54 Oklahoma schools. Research was conducted in two phases which included a description of demographic statistics and provided the analysis of variance to determine significance. Results showed statistical significance between the means of survey responses. Principals and teachers differ in their responses in all six areas of study. Conclusions are presented in Chapter 5.

CHAPTER 5

FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Summary of Study Findings

A study design utilizing a cross-sectional survey may create insufficient sampling, poor quality of survey questions, lack of response, and reporting bias which may impact results (Cohen, et al., 2013). These limitations were addressed in my study by gathering a large sample size to aggregate data and test assumptions. The research questions were closely aligned to the research literature and survey items to ensure participant responses from teachers and principals addressed research questions in a deliberate manner (Butin, 2010). Table 1.1 exhibits the alignment of research questions to research and problems of practice.

Assumptions for ANOVA were checked to ensure the quality of the data was suitable for study. Because of the uneven variance in the large sample size, a Welch alternative ANOVA addressed the lack of homogeneity in the sample to mitigate Type I errors (Tomarken and Serlin, 1986; Jan and Sheih, 2013).

The findings of this study point to significant differences in perceptions of the quality and impact of the Tulsa Model Evaluation instrument depending on the participant's role in the evaluation process. Of the six areas explored in the research questions, all six results show a variance in means with principal responses rating each of the six areas higher than teachers in the survey.

Participant responses on perceptions of their knowledge and understanding of the evaluation system vary by 15%. Principals responded with a higher average score

(M=4.73, SD=0.55) than teachers (M=3.98, SD=1.02). The difference is further illustrated in Figure C.1 in Appendix C.

Results are consistent throughout the other five areas of the evaluation survey. For perceptions regarding usefulness, a variance of 14% between principals and teachers is found in the data. The same is true in qualifications of the evaluator with 8% variance, quality of feedback and data at 14% variance, impact on professional practice at 8% variance, and personal impressions of the Tulsa Model Evaluation system at 12% variance.

These variances can be seen in the survey data between principals and teachers. The highest rating for principals is 80.3% agree they have a thorough understanding of the intended purposes of the Tulsa Model Evaluation system; however, only 36.3% of teachers agree with this statement. Similar differences were found in the principal and teacher responses to questions about usefulness of the evaluation process. Of the principals surveyed, 80.2% agree their school district has implemented the evaluation system with reasonable expectations that are reasonably attainable. Yet only 45.9% of teachers agree.

Principals and teachers are more closely aligned in their responses to rating the quality of the feedback during the evaluation process as tied to the Tulsa Model domains and dimensions. However, when asked whether feedback is an accurate reflection of teaching practice, principals agreed 70.8% of the time but teachers responded in agreement 45.6% of the time.

The results of this survey of perceptions of the Tulsa Model Evaluation system in Oklahoma represent a cross section of principals and teachers in the 2019-2020 school year. No survey of educator perceptions of the Tulsa Model Evaluation process has been conducted since its full implementation in 2014. Results from this study align with results in previous studies of standards-based teacher evaluation systems. Principals and teachers differ in their perceptions of the quality and impact of the Tulsa Model Evaluation process. The degree of difference should assist educators and policy makers in improving Oklahoma evaluation systems to promote more effective professional development and school improvement initiatives aimed at creating conditions for teacher growth.

Conclusions

Feedback and a principal's ability to provide a larger, more purposeful context for the teacher evaluation process can create feelings of competence and relatedness needed to turn the evaluation from a mandate to more internalized motivation to improve student learning (Stiggins and Nickel, 1988).

The problems with state mandated standards-based teacher evaluation tools are many.

1. Lack of understanding of the purpose of the evaluation or ineffective principal training and implementation. Principals do not receive adequate professional development for building collaborative cultures in which feedback is a form of coaching. Principals do not see how an evaluation tool fits into a whole system of improvement or part of a learning organization. In addition, principals may lack the skill to properly understand what should be recorded as effective practice (Bryk & Schneider, 2002; Connally & Tooley, 2016;

Darling-Hammond, et al., 2012; Fullan, 2009; Hill & Grossman, 2013; Hallinger & Heck, 2014; Hallinger, Heck, & Murphy, 2014; Honig, et al., 2010).

2. Lack of teacher input or feedback. Teachers should be treated as learners and leaders in the system (Darling-Hammond et al., 2012). Without attention to teachers as adult learners who need feedback, mental models, and reflection, effective strategies will not bridge the knowing/doing gap. When teachers do not apply what they have learned about effective practice, student achievement will suffer (Hattie, 2015a; Stronge & Tucker, 1999; Taylor & Tyler, 2012; James & McCormick, 2009).
3. Lack of connection to professional development and professional community. Teacher evaluation instruments work best as a tool within schools as learning organizations. When principals and teachers use evaluation to provide focused conversations about specific strategies, student achievement results are likely to increase (Reeves, 2010; Robinson, Lloyd, & Rowe, 2008; Taylor & Tyler, 2012; Goe, Biggers, & Croft, 2012).
4. Limited range and authenticity in evaluation scores. Authenticity can be defined as validity or legitimacy. An evaluation of scores from several evaluation tools ranks most teachers a 3 on a five-point scale. Most systems use a Likert scale rubric based on standards of effective teaching practice. Principals tend to choose mid-range scores resulting in overall ratings of "effective" or "highly effective" either because they believe the rating scale to be too blunt an instrument to measure effective practice or they believe

teachers should not be rated poorly if proper supports are not given for improvement (Borko, Liston, & Whitcomb, 2009; Connally & Tooley, 2016; Dynarski, 2016; Taylor & Tyler, 2012; Weisberg, et al., 2009). The lack of validity is especially problematic for schools attempting to use the evaluation process as a vehicle to measure effective practice as it is linked to school improvement. Although a regulation can become internalized, leading to sustained change, if the regulation is seen to have value (Gagne & Deci, 2005), one way to add value is to provide an environment for proper feedback. If evaluation ratings consistently yield a 3 rating, the value or usefulness of the tool comes into question.

A gap exists in quantitative research for merging organizational and systems thinking theory within local school contexts where accountability reform mandates, such as standards-based evaluation processes, are applied. Research has been done on leadership behaviors but not on mechanisms for successful organizational change (Honig & Louis, 2007; Heck & Hallinger, 2010). Schools are now free from the stiff federal accountability regulations of NCLB and grant driven educational reforms of Race to the Top. Continued obligations for system improvement are required with revisions to the Every Student Succeeds Act., albeit with greater flexibility for states to determine how to apply research-based practices. Even under new federal guidelines, principals and teachers are left attempting to successfully use mandated evaluation frameworks which were adopted during the accountability reform onslaught.

After collecting perceptions from Oklahoma principals and teachers, it is evident educators in this state do not fully understand the purpose, usefulness, or potential impact

of the Tulsa Model Evaluation process on conditions for teacher growth or, as promised by policy makers in its initial adoption, for improving learning outcomes for students. In this study the variance between principal and teacher responses is clear for all six research questions. In the area of knowledge and understanding of the Tulsa Model Evaluation instrument, principals were more likely to answer in agreement they have a thorough understanding of the structure, procedures, and purpose of the process. When asked if participants had a thorough understanding of the purpose of the Tulsa Model Evaluation system, 80.3% of principals agreed compared to 36.3% of teachers.

There was another large gap in the perception of implementation of the system. When asked if their district has implemented the Tulsa Model Evaluation system with reasonable expectations that are realistically attainable, principals responded with 80.2% in agreement with teachers responding at 45.9% agreement. The disparity in these responses point to the need for increased collaborative conversations between teachers and principals to set clearer expectations of intended outcomes when engaged in the evaluation process.

Perhaps not surprisingly, overall principals believe they bring knowledge, expertise, and support to their teachers during the evaluation process. Principals responded that they have the technical expertise as an instructional leader with 70.2% in agreement. Teachers responded to a parallel question about their evaluator's technical expertise with 52.4% in agreement. Although the variance in this area was not as wide as in other areas of the survey, it does raise the question of what is being lost in communication between principals and teachers when it comes to discussion about improving teaching practices.

Another key area of disparity is in the answers given for the perceptions regarding the quality of data collected and feedback provided during the evaluation process. Principals agreed at 70.8% the feedback they provide is an accurate reflection of a teacher's teaching practice. Teachers, however, only agreed with this statement at 45.6%. In addition, 73.3% of principals in the sample agreed the feedback they provide includes discussion and connection to student learning and the monitoring of student progress, whereas 48% of teachers agreed with this statement. Feedback is a critical component in improving teacher conditions for growth (Stiggins and Nickel, 1988), yet based on the results of this survey, the quality of feedback given to Oklahoma teachers can be improved.

One area with the least amount of variance between principals and teachers is in the perceptions of outcomes of the Tulsa Model Evaluation process on professional development and professional practice. The variance between principals and teachers in this section of the survey is 8%. The variance can be attributed to responses on multiple questions. When asked if the participant can show data indicating the achievement level of students has improved because of their participation in the Tulsa Model Evaluation process, only 17.8% of principals agreed and only 15.4% of teachers agreed with this statement. When asked if the alignment of professional development practices and teacher expectations has improved due to the implementation of the Tulsa Model Evaluation process, only 23.7% of principals and 13.9% of teachers agreed. Finally, when asked if participants have gained a great deal of personal satisfaction from participating in the Tulsa Model Evaluation process, principals agreed at 22% and teachers agreed at 13.9%. Principals and teachers in this sample agree the evaluation

process has had very little impact on their professional development and professional practices.

Principals have a slightly better response to personal impressions regarding the impact of the Tulsa Model Evaluation system on professional conversations. When asked if the Tulsa Model domains and dimensions have encouraged more in-depth conversations between principals and teachers about effective teaching practice, 36.5% of principals agreed, but only 19.8% of teachers agreed. A further divide can be seen when asked if the professional growth provided through principal and teacher discussions about the Tulsa Model Evaluation process make the time and effort invested worthwhile. Principals agreed with this statement at 37.4% yet teachers agreed only at 19.9%. Although 55.7% of principals support the continued use of the Tulsa Model evaluation system in Oklahoma, only 22.8% of teachers do.

The results of this study align with results from previous research. Schumacher (2004) concluded based on his study of standards-based teacher evaluation in Wisconsin that although teachers know and understand the process of the evaluation, they were uncertain whether the evaluation process motivated them to improve their teaching practices. Additional findings from the study show teachers felt they did not gain personal satisfaction from the evaluation and believed it should be discontinued (Shumacher, 2004). Recommendations from this study are to clearly define and connect the evaluation system to the expected outcomes.

A defined purpose of the evaluation system is also emphasized in the findings of a similar study in Iowa. Huckstadt (2011) found Iowa teachers' and administrators' responses varied depending on their role in the standards-based evaluation process.

Administrators consistently ranked every survey question higher than teachers in seven areas: 1. reported knowledge of the Iowa Teaching Standards; 2. perceptions regarding the fair and ethical implementation of the evaluation system; 3. perceived qualifications and expertise of those conducting evaluations; 4. the quality of data and feedback provided; 5. the observed impact on teaching and professional development practices; 6. the reported impact on motivation and personal satisfaction; and 7. the impact specific elements of the evaluation process have had on professional practices.

The conclusions of the Huckstadt (2011) study were teachers perceptions varied significantly from principals, and principals believed the evaluation system had more significant impact on professional practices than teachers. Survey responses revealed that the support provided by principals may not be as helpful as they perceive according to teacher responses. A major recommendation from this study is the teacher evaluation process should be more strongly connected to teacher professional development.

Recommendations for Future Research

Previous dissertation studies measuring the perceived effect of standards-based evaluation tools have been completed in Iowa, Wisconsin, and New Jersey. The Huckstadt (2011) study collected quantitative survey data from administrators and teachers in Iowa. Findings in this study indicated that both administrators and teachers achieved an understanding of the evaluation tool, but they ranked the quality of feedback and the impact on effective teaching low. Overall, administrators ranked the impact of the evaluation instrument higher than teachers. Teachers saw less connection between evaluation feedback conversations and their teaching practices. The summative evaluation seemed to be the most impactful event in the evaluation process for

administrators and teachers. Among the recommendations from Huckstadt's (2011) study was a suggestion for training administrators and teachers in coaching conversation skills and in self-evaluation skills to build agency around the evaluation process.

Recommendations from my study include using the Oklahoma Tulsa Model Survey data to disaggregate variable effects. Survey data collected variables for gender, ethnicity, school size, educator years of experience, principal number of evaluations each year, grade bands, and use of collaborative teams in the district. It was not within the scope of this study to explore multiple regressions, but a cursory examination of data may show correlations between teacher or principal years of experience and positive or negative perceptions of the evaluation system.

An additional recommendation is using this survey to sample a larger population. Because this survey has been used to gather information in two studies on standards-based teacher evaluations, one in Iowa and one in Oklahoma, and because the survey has proved to be both valid and reliable, it is possible to use this survey to gather information from a larger sample of Oklahoma teachers through education agencies. The agencies might be interested in gathering information on how the Tulsa Model Evaluation process impacts mentoring and induction programs. Components within the evaluation system such as pre-observation conferences, post-observations conferences and the compilation of artifacts gathered during the evaluation process can also be studied to determine their individual impact on teacher professional growth.

Implication for Policy and Practice

A growing segment of inquiry connects principal leadership to teacher evaluation and feedback, which ultimately impacts student learning and achievement results. A

study conducted in Chicago Public Schools creates a pathway from principal leadership to student outcomes on the ACT Education Planning and Assessment System (Sebastian & Allensworth, 2012). Through multilevel equation modeling and analysis of teacher survey results, researchers created a link from principal actions to student achievement. The results found that although a principal's direct impact on student learning is small, the overall school climate and learning environment created by the principal makes the biggest difference for student achievement.

In addition to the Chicago study, researchers in Cincinnati Public Schools found qualitative data collected via classroom observation tools may also have positive influence on student achievement (Sebastian & Allensworth, 2012). Researchers in this study limited their data to the established Cincinnati Teacher Evaluation System instrument, which is based on Charlotte Danielson's framework of effective teaching (Taylor & Tyler, 2012). In a comparative analysis of teacher evaluation scores and student achievement scores over time, the results suggest the act or the process of evaluation has a positive impact on student achievement, especially in math. Taylor and Tyler (2012) also uncover evidence that critical feedback to teachers improves student performance. The Cincinnati study makes a case that an evaluation instrument, if used to provide specific feedback to teachers, can make a difference for student learning.

A quantitative approach has been employed in many studies about correlations between teacher evaluation and student achievement. Multiple logic models explaining the pathways relationship between principal, teacher, and community effects on student learning have been developed through quantitative results ((Bryk, et al., 2010; Hallinger, Heck, & Murphy, 2014; Leithwood & Louis, 2012; Heck & Hallinger, 2010; Honig,

2008). Quantitative methods continue to add to the body of research on the topic of integrating learning organization theory and social learning theory approaches to school improvement. Acknowledging the complexity of educational research in the field, quantitative methods emphasize critical thinking and deeper questioning (Henson, Hull, & Williams, 2010). Quantitative studies move the research forward to better understand the broader impact of policy mandates. A survey approach for quantitative analyses lends itself to a larger sample size and highlights specific problems of practice within the day-to-day conditions of a school environment.

Throughout the surveying of teachers and principals for this study, I received emails from teachers with general comments about their experiences with the Tulsa Model Evaluation process. One vocational teacher wrote, "The Tulsa Model is managed in a way that does not fully recognize what we do. It makes teaching of the trades cumbersome and time consuming. Most of my colleagues and I teach for effect. The districts know who are good teachers and who are not, the community knows who are good teachers and who are not. Administration needs to find ways to remove bad teachers from the equation without making the entire educational field jump through hoops like they are bad teachers."

Another pre-K teacher said, "I think because I have been teaching so long it is NOT the tool or the feedback that improves my teaching. Rather, it is my personal drive for what I LOVE to do that motivates me to read, research, debate, seek outside resources and professional training opportunities, etc. I think for new teachers it is more effective; providing reflection, discussion, collaboration and opportunities to ask for and receive help."

Conclusions and implications for this study connect directly to the problems of practice found in the literature. Teachers should be more involved in the training and potential outcomes of the Tulsa Model Evaluation process. When teachers are not engaged in collaborative conversations about their own professional performance, they can take no agency in creating conditions for their own professional growth. In the worst instances, teachers are removed from the profession without fully understanding their role in student learning, and in the very worst instances, teachers are shamed out of the profession and their civil rights are violated (Amrein-Beardsley, 2019).

Teachers often feel these standards-based mandated frameworks are oppressive, and a system that fails to engage teachers in discussing the very conditions which should encourage their growth will not deliver on promised outcomes. When the Teacher Leader Effectiveness Law was introduced in Oklahoma, the theory of action described to justify its use was frameworks such as the Tulsa Model Evaluation would find strengths and weakness, provide quality professional development, and gather ongoing feedback to improve the system. Based on the results of this study, the mandated evaluation system has not delivered on its promises.

School districts and policy makers should consider re-tooling these frameworks in collaboration with teachers. Prior to these deficit-oriented accountability movement reform mandates, the effective schools movement championed conversations explicitly teaching adults to get results based on tools integrated into the school system. Educating teachers and school leaders on the conditions for teacher growth could increase understanding of the Tulsa Model Evaluation instrument and could improve outcomes for teachers and students

Conversations should continue and education agencies and organizations as well as policy makers should use responses from educators in the field to refine practices on mandated school reforms. When comparing principal responses to teacher responses, there is a significant gap between perceptions of quality and impact of the Tulsa Model Evaluation system. In addition, when it comes to understanding evaluation overall, work can be done to improve the purpose of the instrument and how it can be used for continuous school improvement. Principals favor future use by 55.7%, but teachers do not. Only 22.8% agree with the statement to continue using the evaluation process. Perhaps better conversations will assist in improving its usefulness.

Districts could also use the survey from this study to create their own evaluation systems in conjunction with the state mandated accountability process. Stiggins and Duke (1988) state the best conditions for teacher growth include perceived credentials of evaluator (.65), quality of feedback (.59), depth of information on feedback (.58), and persuasiveness of rationale for suggested changes (.58). Hallinger, Heck, and Murphy (2013) update this research on conditions for teacher growth to include providing actionable feedback to teachers, creating professional communities, support for teacher work, and opportunities for ongoing professional learning. Improved professional development for principals about how their specific behaviors influence conditions for teacher growth should be included in any district level change.

It is important for any successful evaluation process to be perceived as non-threatening (Hallinger, Heck, & Murphy, 2014), and teachers assisting in developing their own evaluation process could create more favorable outcomes toward their own professional growth. Principals as well as central office leaders would do well to

understand previous research which supports leaders working alongside teachers to create processes and learning tools which assist teachers in fully participating in a school organization that creates a healthy learning environment for all (Honig, et al, 2010; Stiggins and Nickel, 1989).

Summary

A recent Gallup Poll (2014) argues for a more human approach to school. The nationwide poll of teachers and students reveals students and teachers suffer from a lack of engagement. A key recommendation from this poll is that teachers should have greater voice in decision making. Using a survey for this study to explore the perceptions of teachers is an attempt to gather more information about how engaged they are in school reform efforts, especially those which claim to build shared knowledge and opportunities for professional development.

Principals and teachers in certain school settings continue to lack motivation, will, and skill (Leithwood & Louis, 2012). Simply legislating implementation of an evaluation process does not result in higher test scores. The problem for practitioners is to understand that climate and culture are critical to garnering any benefit from an evaluation system. In addition, to be most effective, mandated evaluation processes should be part of an integrated learning organization which honors and attempts to maximize relationships promoting positive school cultures. It is time to re-examine accountability reform mandates and ensure these tools achieve the outcomes promised.

REFERENCES

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. Educational Researcher, 35(6), 33-40. Retrieved from http://www.jstor.org.ezproxy.lib.ou.edu/stable/pdf/3876756.pdf

Amrein-Beardsley, A. (2019). The education value-added assessment system (EVAAS) on trial: a precedent-setting lawsuit with implications for policy and practice. *eJEP: eJournal of Education Policy*. Retrieved from https://eric.ed.gov/?id=EJ1234497

Anderson, E., & Reynolds, A. (2015). The state of state policies for principal preparation program approval and candidate licensure. *Journal of Research on Leadership Education*, 10(3), 193-221. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1942775115614292

Augustine, C. H., Gonzalez, G., Ikemoto, G. S., Russell, J., & Zellman, G. L. (2009). Improving school leadership: The promise of cohesive leadership systems. Rand Corporation. Retrieved fromhttp://www.rand.org/pubs/monographs/MG885.html

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review Psychology*, 52, 1-26. Retrieved from http://web.b.ebscohost.com.ezproxy.lib.ou.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=baffb08a-9e15-494c-a999-fc9abada3934%40sessionmgr111&hid=110

Bolman, L. G. & Deal, T. E. (2013). *Reframing Organizations: Artistry, Choice, and Leadership*. San Francisco, CA: Jossey-Bass.

Boothe, M. S. (2019). *Principals' perceptions of oklahoma's teacher and leader effectiveness model* (Order No. 27540053). Available from ProQuest Dissertations & Theses Global. (2300268934). Retrieved from https://login.ezproxy.lib.ou.edu/login?url=https://www-proquest-com.ezproxy.lib.ou.edu/docview/2300268934?accountid=12964

Borko, H., Liston, D., & Whitcomb, J. (2009). Humility and wisdom: Necessary ingredients to reverse the widget effect. *Journal of Teacher Education*, 60(4): 361-363. Retrieved from doi: 10.1177/0022487109342690

Bartholo, R., Tunes, E., & Tacca, M. C. V. R. (2010). Vygotsky's and Buber's pedagogical perspectives: Some affinities. *Educational Philosophy and Theory*, 42(8), 867-880. doi: 10.1111/j.1469-5812.2007.00367.x

Bowers, A. J. (2016). Quantitative research methods training in education leadership and administration preparation programs as disciplined inquiry for building school improvement capacity. *Journal of Research on Leadership Education*, 1942775116659462. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1942775116659462

Bozeman, B. & Kingsley, G. (1998). Risk culture in public and private organizations. *Public Administration Review*, 58(2): 109-118. Retrieved from DOI: 10.2307/976358

Bryk, A., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S. & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Buber, M., & Smith, R. G. (1947). *Between Man and Man; Translated by Ronald Gregor Smith*. Routledge and Kegan Paul.

Butin, D. W. (2010). *The education dissertation: A guide for practitioner scholars*. Thousand Oaks, CA: Corwin.

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns* (Vol. 98). New York: McGraw-Hill.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155. Retrieved from http://www2.psych.ubc.ca/~schaller/528Readings/Cohen1992.pdf

Cohen, L., Manion, L., & Morrison, K. (2013). Research methods in education. New York. Routledge.

Connally, K. & Tooley, M. (2016). Beyond Ratings: Re-envisioning State Teacher Evaluation Systems as Tools for Professional Growth. *New America*. Retrieved from https://www.newamerica.org/education-policy/policy-papers/beyond-ratings/

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: Sage.

Darling-Hammond, L. (1986). A Proposal for Evaluation in the Teaching Profession. *The Elementary School Journal*, 86(4), 531-551. Retrieved from http://www.jstor.org/stable/1001145

Darling-Hammond L. (1998) Policy and Change: Getting Beyond Bureaucracy. In: Hargreaves A., Lieberman A., Fullan M., Hopkins D. (eds) International Handbook of Educational Change. Kluwer International Handbooks of Education, vol 5. Springer, Dordrecht doi.org/10.1007/978-94-011-4944-0_32

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. Teachers College Press.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), 8-15. Retrieved from http://www.jstor.org/stable/41497541

Deci, E. L. and Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: The University of Rochester Press.

DuFour, R., & Fullan, M. (2012). *Cultures Built to Last: Systemic PLCs at Work*. Bloomington, IN: Solution-tree Press.

Dufour, R. & Marzano, R. J. (2011). Leaders of learning: How district, school, and classroom leaders improve student achievement. Bloomington, IN: Solution-tree Press.

Dynarski, M. (2016). Teacher observations have been a waste of time and money. Policy Brief. Brookings. Retrieved from https://www.brookings.edu/research/teacher-observations-have-been-a-waste-of-time-and-money/

Evans, Lisa, Thornton, Bill, & Usinger, Janet. (2012). Theoretical frameworks to guide school improvement. (Report). *NASSP Bulletin, 96*(2), 154-171.

Eyal, O. and Roth, G. (2011). Principals' leadership and teachers' motivation: Self-determination theory analysis. *Journal of Educational Administration,* 49(3),

256-275 doi.org/10.1108/09578231111129055

Fauske, J. R. & Raybould, R. (2005). Organizational learning theory in schools. *Journal of Educational Administration*, 43(1), 22 – 40 doi.org/10.1108/09578230510577272

Fillion, G., Koffi, V. & Booto, J. (2015). Peter Senge's learning organization: A critical view and the addition of some new concepts to actualize theory and practice. *Journal of Organizational Culture, Communications and Conflict*, 19(3), 73-102.

Fullan, M. (1995) The school as a learning organization: Distant dreams, Theory into Practice, 34:4, 230-235, DOI: 10.1080/00405849509543685

Fullan, M. (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10:101-113. Retrieved from doi:10.1007/s10833-009-9108-z

Fullan, M. (2010). *Motion leadership: The skinny on becoming change savvy*. Thousand Oaks, CA: Corwin.

Fulmer, R. M. & Keys, J. B. (1998). A conversation with Peter Senge: New developments in organizational learning. *Organizational Dynamics*, 27(2), 33-42 doi:10.1016/S0090-2616(98)90022-5

Gagne, M. and Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4), 331-362 https://www.jstor.org/stable/4093832

Gallup. (2014). State of America's schools: The path to winning again in education. Policy Report. Retrieved from http://www.gallup.com/services/178709/state-america-schools-report.aspx

Goe, L., Biggers, K., & Croft, A. (2012). Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning. Research & Policy Brief. National Comprehensive Center for Teacher Quality. Retrieved from http://www.gtlcenter.org/products-resources/linking-teacher-evaluation-professional-development-focusing-improving-teaching

Green, E. (2014). *Building a better teacher: How teaching works (and how to teach it to everyone)*. New York: W. W. Norton & Company, Inc.

Hallinger, P. & Heck, R. H. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 26, 5-28. doi: 10.1007/s11092-013-9179-5.

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 26(1), 5-28. doi: 10.1007/s11092-013-9179-5

Hallinger, P., & Leithwood, K. (1998). Unseen forces: The impact of social culture on school leadership. *Peabody Journal of Education*, 73(2), 126-151. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15327930pje7302_6

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annu. Rev. Econ.*, *4*(1), 131-157. doi.org/10.1146/annurev-economics-080511-111001

Hattie, J. (2002). School research: How, what, and why: Six things school leaders should know about educational research. Retrieved from https://cdn.auckland.ac.nz/assets/education/hattie/docs/six-things-school-leaders-should-know-(2002).pdf

Hattie, J. (2003). Teachers make a difference: What is the research evidence? Retrieved from http://www.educationalleaders.govt.nz/Pedagogy-and-assessment/Building-effective-learning-environments/Teachers-Make-a-Difference-What-is-the-Research-Evidence

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.

Hattie, J. (2015a). What works best in education: The politics of collaborative expertise. Open ideas at Pearson, Pearson, London. Retrieved from https://www.pearson.com/hattie/solutions.html

Hattie, J. (2015b). What doesn't work in education: The politics of distraction. Open ideas at Pearson, Pearson, London. Retrieved from https://www.pearson.com/hattie/distractions.html

Heck, R., & Hallinger, P. (2010). Collaborative Leadership Effects on School Improvement: Integrating Unidirectional-and Reciprocal-Effects Models. *The Elementary School Journal*, 111(2), 226-252. doi:10.1086/656299

Henson, R. K., Hull, D. M., & Williams, C. S. (2010). Methodology in our education research culture toward a stronger collective quantitative proficiency. *Educational Researcher*, 39(3), 229-240. Retrieved from https://brainmass.com/file/1571233/EDUCATIONAL+RESEARCHER-2010-Henson-229-40.pdf

Herman, R., Gates, S. M., Arifkhanova, A., Bega, A., Chavez-Herrerias, E. R., Han, E., & Wrabel, S. (2016). School Leadership Interventions Under the Every Student

Succeeds Act: Evidence Review. Retrieved from http://www.rand.org/pubs/research_reports/RR1550-2.html

Hill, H. & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*. 83 2), 371-384. Retrieved from http://dx.doi.org.ezproxy.lib.ou.edu/10.17763/haer.83.2.d11511403715u376

Honig, M. I. (2008). District central offices as learning organizations: How sociocultural and organizational learning theories elaborate district central office administrators' participation in teaching and learning improvement efforts, 114(4): 627-664. Retrieved from http://www.jstor.org/stable/10.1086/589317

Honig, M. I., Copland, M. A., Rainey, L., Lorton, J. A., & Newton, M. (2010). Central office transformation for district-wide teaching and learning improvement. Center for the Study of Teaching and Policy, University of Washington. Retrieved from http://www.wallacefoundation.org/knowledge-center/Pages/Central-Office-Transformation-District-Wide-Teaching-and-Learning.aspx

Honig, M. I., & Louis, K. S. (2007). A new agenda for research in educational leadership: A conversational review. *Educational Administration Quarterly*, 43(1), 138-148. Retrieved from https://www.researchgate.net/profile/Karen_Louis/publication/249684598_A_New_Agenda_for_Research_in_Educational_Leadership_A_Conversational_Review/links/0deec5399fdff151f2000000.pdf

Hoy, W. K., Tarter, C. J., & Kottkamp, R. B. (1991). *Open schools, healthy schools: Measuring organizational climate*. Newbury Park, CA: Corwin Press.

Huckstadt, K. P. (2011). The impact of standards-based teacher evaluation on instruction and professional practices in Iowa (Order No. 3457019). Available from ProQuest Dissertations & Theses Global. (873394673). Retrieved from https://search-proquest-com.ezproxy.lib.ou.edu/docview/873394673?accountid=12964

Israel, G. D. (1992). Determining sample size. Gainesville: University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.

Jan, S., & Shieh, G. (2014). Sample size determinations for Welch's test in one-way heteroscedastic ANOVA. *British Journal of Mathematical & Statistical Psychology, 67*(1), 72-93.

James, M. & McCormick, R. (2009). Teachers learning how to learn. *Teaching and Teacher Education*. 25, 973-982. Retrieved from http://dx.doi.org.ezproxy.lib.ou.edu/10.1016/j.tate.2009.02.023

Kane, T. J., Kerr, K. A., & Pianta, R. C. (2014). Designing Teacher Evaluation Systems: New Guidance from the Measures of Effective Teaching Project. Retrieved from http://k12education.gatesfoundation.org/learning/designing-teacher-evaluation-systems-new-guidance-from-the-measures-of-effective-teaching-project/

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Evaluating teacher effectiveness: can classroom observations identify practices that raise achievement? *Education Next*, 11(3), 54-61.

Leithwood, K., & Jantzi, D. (1999). The relative effects of principal and teacher sources of leadership on student engagement with school. Educational administration

quarterly, 35(5), 679-706. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0013161X99355002

Leithwood, K. & Louis, K. S. (2012). *Linking leadership to student learning.* San Francisco, CA: Jossey-Bass.

Lezotte, L. W. (1992). 'Principal' insights from effective schools. (Cover story). Education Digest, 58(3), 14. Retrieved from http://libraries.ou.edu/access.aspx?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=9302030408&site=ehost-live

Lezotte, L. (2001). Revolutionary and evolutionary: The effective schools movement. Okemos, MI: Effective Schools Products, Ltd. Retrieved from http://www.edutopia.org/pdfs/edutopia.org-closing-achievement-gap-lezotte-article.pdf

Louis, K. S., Leithwood, K., Wahlstrom, K. L., Anderson, S. E., Michlin, M., & Mascall, B. (2010). Learning from leadership: Investigating the links to improved student learning. Center for Applied Research and Educational Improvement/University of Minnesota and Ontario Institute for Studies in Education/University of Toronto, 42, 50. Retrieved from http://www.wallacefoundation.org/knowledge-center/Pages/Investigating-the-Links-to-Improved-Student-Learning.aspx

Neumerski, C. M. (2013). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly*, 49(2), 310-347.

Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0013161X12456700

Manna, P. (2015). Developing Excellent School Principals to Advance Teaching and Learning: Considerations for State Policy. Wallace Foundation. Retrieved from http://www.wallacefoundation.org/knowledge-center/Pages/Developing-Excellent-School-Principals.aspx

Marks, H. M., & Printy, S. M. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 39(3), 370-397. Retrieved from http://journals.sagepub.com/doi/10.1177/0013161X03253412

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.

Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: ASCD.

McIver, M. C., Kearns, J., Lyons, C., & Sussman, M. (2009). Leadership: A McREL Report Prepared for Stupski Foundation's Learning System. Mid-continent Research for Education and Learning (McREL). Retrieved from http://files.eric.ed.gov/fulltext/ED544625.pdf

Mourshed, M., Chijioke, C., & Barber, M. (2010). How the world's most improved school systems keep getting better. McKinsey. Chicago. Retrieved from

http://www.mckinsey.com/industries/social-sector/our-insights/how-the-worlds-most-improved-school-systems-keep-getting-better

Muhammad, A., & Hollie, S. (2011). *The will to lead, the skill to teach: Transforming schools at every level*. Bloomington, IN: Solution-tree Press.

Murphy, J. Hallinger, P., & Heck, R. H. (2013). Leading via teacher evaluation: The case of the missing clothes. Educational Researcher, 20(10), 1-6. doi: 10.3102/0013189X13499625

Oklahoma State Courts Network. (2015). Question submitted by: the honorable Ed Cannaday, State Representative, district 15. Retrieved from http://www.oscn.net/applications/oscn/DeliverDocument.asp?CiteID=465651

Oklahoma State Department of Education (2019). Educator effectiveness theory of action. Retrieved on May 2, 2019 from http://ok.gov/sde/tle.

Oklahoma State Department of Education (2018). Fast Facts FY 17. Retrieved on February 19, 2018 from http://sde.ok.gov/sde/sites/ok.gov.sde/files/documents/files/Fast%20Facts%20FY17%20 December%2021%202017.pdf

Oklahoma State Department of Education (2020). Oklahoma teacher and leader evaluation systems. Retrieved on December 11, 2020 from https://sde.ok.gov/newsblog/2015-01-29/oklahoma-teacher-and-leader-evaluation-systems-sreb-focus-group-findings

Oklahoma State Department of Education (2014). Teacher leader effectiveness (TLE). Retrieved on November 2, 2014 from http://ok.gov/sde/tle

Oklahoma Teacher and Leader Effectiveness Evaluation System. §70-6-1-1.16 (2016)

Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *The American Political Science Review*, 94(2): 251-267. Retrieved from http://www.jstor.org/stable/2586011

Reeves, D. B. (2010). *Transforming professional development into student results*. Alexandria, VA: ASCD.

Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*. doi: 10.1177/0013161X08321509

Salzman, P. C. (2002). On reflexivity. *American Anthropologist*, 104(3): 805-813. Retrieved from doi: 10.1525/aa.2002.104.3.805

Schumacher, G. T. (2004). Perceptions of the impact of a standards -based teacher evaluation system, based on the Danielson "Framework for teaching" model, on teaching and student learning (Order No. 3128070). Available from ProQuest Dissertations & Theses Global (305111804). Retrieved from https://search-proquest-com.ezproxy.lib.ou.edu/docview/305111804?accountid=12964

Sebastian, J. & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning. *Educational Administration Quarterly*, 48: 626-663. doi: 10.1177/0013161X11436273

Senge, P. M. (2006). *The fifth discipline: The art & practice of the learning organization*. New York: Crown Business.

Sheehan, B. & Sheehan, A. M. (2013, June). Successful district leadership interview with Michael Fullan. *AASA New Superintendent's E-Journal*. Retrieved from http://www.aasa.org/content.aspx?id=29254

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

Stiggins, R. J. & Nickel, P. (1989). The teacher evaluation profile: A technical analysis. *Journal of Personnel Evaluation* 2:151-165. https://doi.org/10.1007/BF00127177

Stronge, J. H. (2013). *Principal evaluation: Standards, rubrics, and tools for effective performance.* Alexandria, VA: ASCD.

Stronge, J. H., Richard, H. B., & Catano, N. (2008). *Qualities of effective principals*. ASCD.

Stronge, J. H., & Tucker, P. D. (1999). The politics of teacher evaluation: A case study of new system design and implementation. *Journal of Personnel Evaluation in Education*, 13(4), 339. Retrieved from https://search-proquest-com.ezproxy.lib.ou.edu/docview/201486904?accountid=12964

The Danielson Group. (2013). The framework. Retrieved from http://danielsongroup.org/framework/

Tavakol, Mohsen, & Dennick, Reg. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55.

Taylor, E. S. & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, 102(7): 3628-3651. Retrieved from http://doi.org/10.1257/aer.102.7.3628

Tomarken, A. J. & Serlin, R. C. (1986). Comparisons of ANOVA alternatives under variance of heterogeneity and specific noncentrality structures. *Psychological Bulletin*, 99(1), 90-99.

Towe, P., & Walker, Elaine. (2012). *An Investigation of the Role of a Teacher Evaluation System and Its Influence on Teacher Practice and Professional Growth in Four Urban High Schools,* ProQuest Dissertations and Theses.

Tucker, P. D. & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: ASCD.

Tulsa Public Schools. (2015). TLE observation & evaluation handbook for evaluators using the Tulsa model. Retrieved from https://www.nctq.org/dmsView/Handbook_TLE_Observation_and_Evaluation_System_74031

Tulsa Public Schools. (2017). Tulsa Public Schools' Teacher Observation and Evaluation System: Its Research Base and Validation Studies. Policy Brief. Retrieved from http://sde.ok.gov/sde/sites/ok.gov.sde/files/TLE-TPSEvalResearchBrief.pdf

Tuytens, M. & Devos, G. (2011). Stimulating professional learning through teacher evaluation: An impossible task for the school leader? *Teaching and Teacher Education*, 27, 891-899. Retrieved from http://dx.doi.org.ezproxy.lib.ou.edu/10.1016/j.tate.2011.02.004

Urick, A., & Bowers, A. J. (2011). What influences principals' perceptions of academic climate? A nationally representative study of the direct effects of perception on climate. *Leadership and Policy in Schools*, 10(3), 322-348. Retrieved from doi: 10.1080/15700763.2011.577925

U. S. Department of Education. (2009). *Race to the top executive summary.* Retrieved from https://www2.ed.gov/programs/racetothetop/executive-summary.pdf

U. S. Department of Education. (2017). *Transitioning to the every student succeeds act (ESSA): Frequently asked questions.* Retrieved from https://www2.ed.gov/policy/elsec/leg/essa/essatransitionfaqs11817.pdf

Van Clay, M., & Soldwedel, P. (2013). *Aligning school districts as PLCs*. Bloomington, IN: Solution-tree Press.

Viteritti, J.P. (2012). The Federal Role in School Reform: Obama's 'Race to the Top'. Retrieved from http://www.hunteruap.org/wpcontent/uploads/2012/09/NDLR_Federal-Role-In-School-Reform.pdf

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard university press.

Wallace Foundation. (2012). The School Principal as a Leader: Guiding Schools to Better Teaching and Learning. The Wallace Foundation, New York. Retrieved from http://www.wallacefoundation.org/knowledge-center/Pages/The-School-Principal-as-Leader-Guiding-Schools-to-Better-Teaching-and-Learning.aspx

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The Widget Effect. Education Digest, 75(2), 31-35. Retrieved from http://libraries.ou.edu/access.aspx?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=44479954&site=ehost-live

APPENDIX A

Letters of Permission

Duncan mail

Melonie Hau <melonie.hau@duncanps.org>

Request for permission to use survey instrument

3 messages

Melonie Hau <melonie.hau@duncanps.org> Thu, Aug 2, 2018 at 9:26 AM
To: Kim Huckstadt <kim.huckstadt@uni.edu>
Bcc: Melonie Hau <melonie.hau@duncanps.org>

Dr. Huckstadt,

I am a doctoral student at the University of Oklahoma, completing a dissertation in Educational Administration, Curriculum, and Supervision. I am writing to ask written permission to use the survey instrument titled Iowa Standards-Based Teacher Evaluation Survey from your dissertation in my research study. I am interested in understanding the perceived impact of standards-based evaluation in Oklahoma by using the same survey questions and Likert scale but adapting the title of the instrument for use in Oklahoma schools. Our state also adopted a standards-based evaluation tool in response to Race to the Top fever. My research is being supervised by my professor, Dr. Kathrine Gutierrez.

My study will include a sample of Oklahoma schools with enrollments from various sizes. Our state has not collected usable data about the impact of the evaluation tool which has been mandated since 2012. I intend to measure the variance of perceptions between administrators and teachers in the use of the Tulsa Model Evaluation tool in four areas: understanding of the purpose of the evaluation, quality of feedback, professional growth, and outcomes on professional agency.

Since I plan to distribute the survey this fall, I would also appreciate any insights you have in distributing, collecting, and synthesizing the data.

In addition to using the instrument, I ask your permission to reproduce it in my dissertation appendix.

I would like to use your survey instrument under the following conditions:

- I will use the instrument only for my research study and will not sell or use it for any other purposes.
- I will include a statement of attribution and copyright on all copies of the instrument. If you have a specific statement of attribution that you would like for me to include, please provide it in your response.
- At your request, I will send a copy of my completed research study to you upon completion of the study and/or provide a hyperlink to the final manuscript.

If you do not control the copyright for these materials, I would appreciate any information you can provide concerning the proper person or organization I should contact.

If these are acceptable terms and conditions, please indicate so by replying to me through e-mail at melonie.hau@duncanps.org.

Teacher evaluation has been a passion of mine since I completed National Board Certification as a teacher. I believe in a strong coaching/evaluation model, and as a superintendent, I try to instill that believe in principals in our district. I still see a place for effective teacher evaluation in school in spite of the challenges of state mandates and local implementation.

Thank you for your time and I hope to hear from you soon.

Sincerely,

Melonie Hau

Superintendent

(580) 255-0686
1706 W. Spruce
Duncan, OK 73533

An Ongoing Tradition of Excellence

a 1

Kim Huckstadt <kim.huckstadt@uni.edu> Fri, Aug 3, 2018 at 9:19 AM
To: Melonie Hau <melonie.hau@duncanps.org>

Dear Melonie,

I appreciate your request and I am glad to give you permission to utilize the survey I developed for my dissertation research focused on the impact of standards-based evaluation on professional practices in Iowa. When I completed my study in 2010, one of my greatest hopes was that it might help others down the road conduct parallel research that would focus on improving our evaluation system. So I am pleased that you feel my work will be relevant and useful in your study.

As you have perhaps noted, I am a strong proponent of elevating the formative aspects of our evaluation systems to create greater opportunities for meaningful collaboration focused on professional growth. All too often, the evaluation is viewed as "hoop jumping" event and not at all connected to our professional development plans. I look forward to hearing more about your research in the future.

You have my permission to use the survey instrument I adapted and developed in my 2010 dissertation study in accordance with the parameter described in your e-mail. I very much appreciate that you will include a statement of attribution regarding the instrument. I do not have a specific statement of attribution for you to include. Please use the citation and acknowledgement format that is customary for this kind of work.

Best of luck with your study. I am sure it will go well!

Sincerely,
Kim H.

PS - In 2010, I utilized Survey Monkey to distribute the 43 question surveys to participating schools. It took quite a bit of time to get the number of schools identified because I wanted a representative population of Iowa schools as described in the study. I tried to work with the Iowa Dept. of Education to help recruit schools but it would have taken years to get through all the red tape... and I did not want to take forever! So I ended up reaching out through my professional networks. Iowa is a relatively small state and people were very supportive.

Most of this is described in the methods chapter but I would be glad to visit about specifics if that would be helpful to you.
[Quoted text hidden]
--
Kim P. Huckstadt
Assistant Professor
Educational Leadership
University of Northern Iowa
Office: (319) 273-4525
Cell: (563) 212-4363

Melonie Hau <melonie.hau@duncanps.org> Sat, Aug 4, 2018 at 12:44 PM
To: Kim Huckstadt <kim.huckstadt@uni.edu>

Dr. Huckstadt,

Thank you so much for your permission and your help! I plan to use Survey Monkey also, and as a superintendent in this small state, I hope to pull on the kindness of a few of my fellow education warriors.

I believe there's an element of coaching in the evaluation system principals could utilize to impact continuous school improvement. I see feedback as a form of social currency that is often neglected. Most of the districts in our state use the Tulsa Model Evaluation, which was developed through Tulsa Public Schools via a Gates Foundation Grant. I am honing in on this evaluation tool because it is based on the Danielson Framework. Tulsa conducts some internal research on the impact of the evaluation, but our state hasn't explored its effectiveness at all, even though it is mandated by Oklahoma law.

Thank you again, and I will keep you posted on my progress.

Melonie Hau
Superintendent
(580) 255-0686
1706 W. Spruce

APPENDIX B

Survey Instruments

OKLAHOMA STANDARDS-BASED TEACHER EVALUATION SURVEY

TULSA MODEL EVALUATION IMPACT STUDY

(TEACHER VERSION)

Teacher Online Consent to Participate in Research

Would you like to be involved in research at the University of Oklahoma?

I am Melonie Hau from the Educational Leadership and Policy Studies Department and I invite you to participate in my research project entitled "The Perceived Quality and Impact of the Tulsa Model Evaluation in Oklahoma Schools." This research is being conducted at the University of Oklahoma. You were selected as a possible participant because you are a teacher or evaluator in Oklahoma. You must be at least 22 years of age to participate in this study.

Please read this document and contact me to ask any questions you may have BEFORE agreeing to take part in my research.

What is the purpose of this research? The purpose of this research is to survey teachers and principals on the perceived understanding and impact of the Tulsa Model Evaluation instrument.

How many participants will be in this research? Up to 500 people will take part in this research.

What will I be asked to do? If you agree to be in this research, you will take a short survey with multiple choice answers. You will be asked to complete some demographic information regarding your experience as a teacher or evaluator. You will also be asked to include some information about your use and implementation of the Tulsa Model Evaluation instrument in your school. Please note that several questions require an answer from you in order to continue the survey; however, the question regarding gender is not required but is requested.

How long will this take? Your participation will take 12 minutes to complete the survey including demographic data.

What are the risks and/or benefits if I participate? Although we do not anticipate any risks, there is a possibility, given the multiple demographics and employment questions we are asking, that someone may deductively re-identify you. The anticipated number of participants will minimize this risk as well. If you are concerned about this possibility, you can choose not to participate in the survey or terminate your participation at any point. There are no benefits for participation in this study.

Will I be compensated for participating? You will not be reimbursed for your time and participation in this research.

Who will see my information? In research reports, there will be no information that will make it possible to identify you. The possibility of deductive re-identification has been minimized by the anticipated number of participants to complete this survey. Research records will be stored securely and only approved researchers and the OU Institutional Review Board will have access to the records.

Data are collected via an online survey system that has its own privacy and security policies for keeping your information confidential. Please note no assurance can be made as to the use of the data you provide for purposes other than this research.

What will happen to my data in the future? After removing all identifiers, we might share your data with other researchers or use it in future research without obtaining additional consent from you. Data are intended to include baseline information for further studies and may be retained for future studies of teacher evaluation in Oklahoma. It is possible only data in de-identified form will be used in Legislative interim studies or to craft future education policy in Oklahoma through the Oklahoma State Department of Education or state professional education groups such as Oklahoma Education Association (OEA), Cooperative Council for Oklahoma School Administrators (CCOSA), or Oklahoma State School Board Association (OSSBA).

Do I have to participate? No. If you do not participate, you will not be penalized or lose benefits or services unrelated to the research. If you decide to participate, you do not have to answer any questions and can stop participating at any time.

Who do I contact with questions, concerns or complaints? If you have questions, concerns or complaints about the research please contact Melonie Hau at (405) 226-9263 or haumelonie@yahoo.com. You can also contact my faculty advisor Dr. John Jones at (405) 325-4165 or jrjones@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 if you have questions about your rights as a research participant, concerns, or complaints about the research and wish to talk to someone other than the researcher(s) or if you cannot reach the researcher(s).

Please print this document for your records. By providing information to the researcher(s), I am agreeing to participate in this research.

This research has been approved by the University of Oklahoma, Norman Campus IRB.

IRB Number: 11125 **Approval date:** 10/05/2019 Exempt

Tulsa Model Evaluation tool TEACHER Survey Instrument

Section 1: Demographic Information

1. Including the current year, how many years have you been teaching?
    1. This is my first year
    2. 2 to 5 years
    3. 6 to 10 years
    4. 11 to 15 years
    5. 16 to 20 years
    6. 21 or more years
2. What best describes the enrollment of your district?
    1. Less than 400
    2. 400 to 1,000
    3. 1,000 to 2,500
    4. 2,500 to 5,000
    5. 5,000 to 9,000
    6. Over 10,000
3. Your current teaching assignment grade level. Select the answer that best describes your current position.
    1. Pre-K through Grade 1
    2. Grades 2 through 5 or 6
    3. Grades 6 through 8
    4. Grades 9 through 12
    5. K through 12
4. Your gender
    1. Male
    2. Female

5. Choose one or more races that you consider yourself to be
    1. White
    2. Black or African American
    3. American Indian of Alaska Native
    4. Asian
    5. Native Hawaiian or Pacific Islander
    6. Other

6. Date of most recent summative evaluation
    1. During the academic year 2017-2018
    2. During the academic year 2016-2017
    3. During the academic year 2015-2016
    4. During the academic year 2014-2015
    5. Prior to 2014

7. My school uses collaborative teams or professional learning communities to inform instruction and make decisions about student learning.
    1. Yes
    2. No

Section 2: Participant Responses to 6 Key Areas Based on Research Questions. Participants respond based on a Likert scale: 1 = Strongly Agree through 5 = Agree

A. Knowledge and Understanding of the Evaluation System: Describe your knowledge and understanding of the Tulsa Model Evaluation system for teachers.

1. I have a thorough understanding of the 5 domains and 20 dimensions that formulate the basis of the teacher evaluation system in the state of Oklahoma. I know what they are, and I know how they relate to my teaching responsibilities.
2. I have a thorough understanding of the evaluation tool and procedures utilized in my district.
3. I have a thorough understanding of the intended purposes of the Tulsa Model Evaluation System.

B. Perceptions Regarding Usefulness: Describe your perceptions of and experiences utilizing the Tulsa Model Evaluation process in your school.

1. My school/district has implemented the Tulsa Model Evaluation System with reasonable expectations that are realistically attainable.
2. Because of the Tulsa Model Evaluation system, I have a clearer idea of what my district expects of me.
3. The procedures used during the Tulsa Model Evaluation are fair.
4. The procedures used during the Tulsa Model Evaluation promote ethical practices.
5. I believe my school has implemented the Tulsa Model Evaluation system in an effective manner.
6. The domains and dimensions identified in the state of Oklahoma are representative of best professional practice and help define effective teaching.

C. Qualifications and Evaluator Leadership: Describe your perceptions of the knowledge, expertise and supportiveness of the individual conducting your most recent performance evaluation.

1. My evaluator has a thorough knowledge of the Tulsa Model domains and dimensions.
2. My evaluator knows well the procedures to implement the Tulsa Model Teacher Evaluation System.
3. My evaluator spends adequate time observing my instruction in order to form a basis to assess my performance related to the Tulsa Model domains and dimensions.

4. I am generally satisfied with the discussions that I have had with my evaluator regarding my performance in relation to the Tulsa Model domains and dimensions.
5. I receive adequate support to meet the expectations identified in the Tulsa Model domains and dimensions.
6. It has been my experience that evaluations are conducted in a non-threatening manner with a focus on continuous improvement.
7. My evaluator has the knowledge and technical expertise as an instructional leader to guide the professional growth needed to improve my teaching practice.

D. Data and Feedback: Describe your perceptions regarding the quality of the data collected and feedback provided during your most recent evaluation.

1. The feedback I receive from my evaluator is clearly linked to the Tulsa Model domains and dimensions.
2. The feedback I receive from my evaluator is an accurate reflection of my teaching practice.
3. The feedback I receive from my evaluator includes discussion and connection to student learning and monitoring the progress of the students in my classes.
4. The feedback I receive from my evaluator is valuable information that I have used to improve my teaching practice.
5. The feedback I receive from my evaluator is supported by opportunities to participate in professional development or training designed to improve my teaching practice.

E. Impact on Professional Practices: Describe your perceptions of the outcomes of the evaluation process related to professional development activities and professional practices.

1. I have changed instructional methodologies for the benefit of my students as a result of participating in the Tulsa Model Evaluation.
2. I can show data that indicates the achievement level of my students has improved as a result of my participation in the Tulsa Model Evaluation process.
3. As a result of my participation in the Tulsa Model Evaluation process, I now spend a greater amount of time reflecting on my teaching practice.
4. As a result of the implementation of the Tulsa Model Evaluation process, the alignment between expectations for teachers, professional development practices and student achievement goals has improved significantly in my school.
5. The Tulsa Model domains and dimensions have provoked discussion regarding effective teaching practices among teachers at my school.
6. The Tulsa Model Evaluation system has challenged me to improve my teaching practices.
7. Participation in the Tulsa Model Evaluation process is valuable to me as a professional development activity.

8. I have gained a great deal of personal satisfaction as a result of participating in the Tulsa Model Evaluation process.

F. Personal Impressions of the Tulsa Model Evaluation System: Describe your perceptions regarding the impact of the Tulsa Model Evaluation System on professional conversations between teachers and administrators in your school.

1. The Tulsa Model domains and dimensions have encouraged more in-depth discussions between administrators and teachers regarding effective teaching practices.
2. Working with my administrator through the Tulsa Model Evaluation system helped me to learn how I can improve my teaching practices.
3. I believe that by working with my administrator and by meeting the expectations identified in the Tulsa Model Evaluation, my instructional practices will progress to a higher level.
4. The professional growth through administrator and teacher discussions I have gained from participation in the Tulsa Model Evaluation process make it worth the time and effort invested.
5. Participating in the Tulsa Model Evaluation motivates me to improve my teaching performance.
6. I have learned a lot from participating in the Tulsa Model Evaluation process.
7. I am generally supportive of the Tulsa Model Evaluation system.
8. I support the continued use the Tulsa Model Evaluation System in Oklahoma.

OKLAHOMA STANDARDS-BASED TEACHER EVALUATION SURVEY

TULSA MODEL EVALUATION IMPACT STUDY

(ADMINISTRATOR VERSION)

Administrator Online Consent to Participate in Research

Would you like to be involved in research at the University of Oklahoma?

I am Melonie Hau from the Educational Leadership and Policy Studies Department and I invite you to participate in my research project entitled "The Perceived Quality and Impact of the Tulsa Model Evaluation in Oklahoma Schools." This research is being conducted at the University of Oklahoma. You were selected as a possible participant because you are a teacher or evaluator in Oklahoma. You must be at least 22 years of age to participate in this study.

Please read this document and contact me to ask any questions you may have BEFORE agreeing to take part in my research.

What is the purpose of this research? The purpose of this research is to survey teachers and principals on the perceived understanding and impact of the Tulsa Model Evaluation instrument.

How many participants will be in this research? Up to 500 people will take part in this research.

What will I be asked to do? If you agree to be in this research, you will take a short survey with multiple choice answers. You will be asked to complete some demographic information regarding your experience as a teacher or evaluator. You will also be asked to include some information about your use and implementation of the Tulsa Model Evaluation instrument in your school. Please note that several questions require an answer from you in order to continue the survey; however, the question regarding gender is not required but is requested.

How long will this take? Your participation will take 12 minutes to complete the survey including demographic data.

What are the risks and/or benefits if I participate? Although we do not anticipate any risks, there is a possibility, given the multiple demographics and employment questions we are asking, that someone may deductively re-identify you. The anticipated number of participants will minimize this risk as well. If you are concerned about this possibility, you can choose not to participate in the survey or terminate your participation at any point. There are no benefits for participation in this study.

Will I be compensated for participating? You will not be reimbursed for your time and participation in this research.

Who will see my information? In research reports, there will be no information that will make it possible to identify you. The possibility of deductive re-identification has been minimized by the anticipated number of participants to complete this survey. Research records will be stored securely and only approved researchers and the OU Institutional Review Board will have access to the records.

Data are collected via an online survey system that has its own privacy and security policies for keeping your information confidential. Please note no assurance can be made as to the use of the data you provide for purposes other than this research.

What will happen to my data in the future? After removing all identifiers, we might share your data with other researchers or use it in future research without obtaining additional consent from you. Data are intended to include baseline information for further studies and may be retained for future studies of teacher evaluation in Oklahoma. It is possible only data in de-identified form will be used in Legislative interim studies or to craft future education policy in Oklahoma through the Oklahoma State Department of Education or state professional education groups such as Oklahoma Education Association (OEA), Cooperative Council for Oklahoma School Administrators (CCOSA), or Oklahoma State School Board Association (OSSBA).

Do I have to participate? No. If you do not participate, you will not be penalized or lose benefits or services unrelated to the research. If you decide to participate, you do not have to answer any questions and can stop participating at any time.

Who do I contact with questions, concerns or complaints? If you have questions, concerns or complaints about the research please contact Melonie Hau at (405) 226-9263 or haumelonie@yahoo.com. You can also contact my faculty advisor Dr. John Jones at (405) 325-4165 or jrjones@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 if you have questions about your rights as a research participant, concerns, or complaints about the research and wish to talk to someone other than the researcher(s) or if you cannot reach the researcher(s).

Please print this document for your records. By providing information to the researcher(s), I am agreeing to participate in this research.

This research has been approved by the University of Oklahoma, Norman Campus IRB.

IRB Number: 11125 **Approval date:** 10/05/2019 Exempt

Tulsa Model Evaluation tool ADMINISTRATOR Survey Instrument

Section 1: Demographic Information

1. Including the current year, how many years have you served as an administrator?
    1. This is my first year
    2. 2 to 5 years
    3. 6 to 10 years
    4. 11 to 15 years
    5. 16 to 20 years
    6. 21 or more years

2. What best describes the enrollment of your district?
    1. Less than 400
    2. 400 to 1,000
    3. 1,000 to 2,500
    4. 2,500 to 5,000
    5. 5,000 to 9,000
    6. Over 10,000

3. Your current administrative assignment grade level. Select the answer that best describes your current position.
    1. Pre-K through Grade 1
    2. Grades 2 through 5 or 6
    3. Grades 6 through 8
    4. Grades 9 through 12
    5. Pre-K through 12
4. Your gender
    1. Male
    2. Female

5. Choose one or more races that you consider yourself to be
    1. White
    2. Black or African American
    3. American Indian of Alaska Native
    4. Asian
    5. Native Hawaiian or Pacific Islander
    6. Other

6. List the number of formal/summative evaluations you will conduct this year
    1. 1 to 4
    2. 5 to 10

3. 11 to 15
4. 16 to 20
5. More than 20

7. My school uses collaborative teams or professional learning communities to inform instruction and make decisions about student learning.
    1. Yes
    2. No

Section 2: Participant Responses to 6 Key Areas Based on Research Questions. Participants respond based on a Likert scale: 1 = Strongly Agree through 5 = Agree

A. Knowledge and Understanding of the Evaluation System: Describe your knowledge and understanding of the Tulsa Model Evaluation system for teachers.

1. I have a thorough understanding of the 5 domains and 20 dimensions that formulate the basis of the teacher evaluation system in the state of Oklahoma. I know what they are, and I know how they relate to my responsibilities as an administrator.
2. I have a thorough understanding of the evaluation tool and procedures utilized in my district.
3. I have a thorough understanding of the intended purposes of the Tulsa Model Evaluation system.

B. Perceptions Regarding Implementation: Describe your perceptions of and experiences utilizing the Tulsa Model Evaluation process in your school.

1. My school/district has implemented the Tulsa Model Evaluation System with reasonable expectations that are realistically attainable.
2. Because of the Tulsa Model Evaluation system, I have a clearer idea of what my district expects of teachers under my supervision.
3. The procedures used during the Tulsa Model Evaluation are fair.
4. The procedures used during the Tulsa Model Evaluation promote ethical practices.
5. I believe my school has implemented the Tulsa Model Evaluation system in an effective manner.
6. The domains and dimensions identified in the Tulsa Model Evaluation are representative of best professional practice and help define effective teaching.

C. Qualifications and Evaluator Training: As an individual conducting performance evaluations for teachers, describe your perceptions of the knowledge, expertise and supportiveness in conducting your most recent performance evaluations.

1. As an evaluator, I have a thorough knowledge of the Tulsa Model domains and dimensions.

2. As an evaluator, I know well the procedures to implement the Tulsa Model Teacher Evaluation System.
3. As an evaluator, I spend adequate time observing instruction in order to form a basis to assessing teaching performance related to the Tulsa Model domains and dimensions.
4. I am generally satisfied with the discussions I have had with teachers regarding their performance in relation to the Tulsa Model domains and dimensions.
5. I provide adequate support in assisting teachers to meet the expectations identified in the Tulsa Model domains and dimensions.
6. It has been my experience that evaluations are conducted in a non-threatening manner with a focus on continuous improvement.
7. I have the knowledge and technical expertise as an instructional leader to guide the professional growth of teachers and to help improve their teaching practices.

D. Data and Feedback: Describe your perceptions regarding the quality of the data collected and feedback provided during your most recent evaluation process.

1. The feedback I provide to teachers is clearly linked to the Tulsa Model domains and dimensions.
2. The feedback I provide to teachers is an accurate reflection of each teacher's teaching practice.
3. The feedback I provide teachers includes discussion and connection to student learning and monitoring the progress of their students in their classes.
4. The feedback I provide to teachers is valuable information they are able to use to improve their teaching practices.
5. The feedback I provide teachers is supported by opportunities to participate in professional development or training designed to improve their teaching practices.

E. Impact on Professional Practices: Describe your perceptions of the outcomes of the evaluation process related to professional development activities and professional practices.

1. I have observed teachers changing instructional methodologies for the benefit of students as a result of participating in the Tulsa Model Evaluation System.
2. I can show data that indicates the achievement level of students has improved as a result of my participation in the Tulsa Model Evaluation process.
3. As a result of my participation in the Tulsa Model Evaluation process, I now spend a greater amount of time reflecting on my practices as an instructional leader.
4. As a result of the implementation of the Tulsa Model Evaluation process, the alignment between expectations for teachers, professional development practices and student achievement goals has improved significantly in my school.
5. The Tulsa Model domains and dimensions have provoked discussion regarding effective teaching practices among teachers at my school.

6. The Tulsa Model Evaluation system has challenged me to improve my understanding of teaching practices.
7. Participation in the Tulsa Model Evaluation process is valuable to me as a professional development activity.
8. I have gained a great deal of personal satisfaction as a result of participating in the Tulsa Model Evaluation process.

F. Personal Impressions of the Standards-Based Evaluation System: Describe your perceptions regarding the impact of the Tulsa Model Evaluation System.

1. The Tulsa Model domains and dimensions have encouraged more in-depth discussions between administrators and teachers regarding effective teaching practices.
2. Working with teachers through the Tulsa Model Evaluation system helped me to learn how I can improve teaching practices.
3. I believe that by meeting the expectations identified in the Tulsa Model Evaluation, instructional practices of teachers will progress to a higher level.
4. The professional growth through administrator and teacher discussions while participating in the Tulsa Model Evaluation process make it worth the time and effort invested.
5. Participating in the Tulsa Model Evaluation motivates me to improve as an instructional leader.
6. I have learned a lot from participating in the Tulsa Model Evaluation process.
7. I am generally supportive of the Tulsa Model Evaluation system.
8. I support the continued use the Tulsa Model Evaluation System in Oklahoma.

APPENDIX C

Supplemental Tables and Figures

Table C. 1

Principal and Teacher Responses Regarding Usefulness as a Percentage of the Sample

Describe your perceptions of and experiences utilizing the Tulsa Model Evaluation process in your school.										
	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
My school/district has implemented the Tulsa Model Evaluation system with reasonable expectations that are realistically attainable.	80.2	45.9	17.6	37.4	0.0	9.8	0.8	4.2	0.8	2.7
Because of the Tulsa Model Evaluation system, I have a clearer idea of what my district expects of me.	56.5	30.7	35.1	38.0	3.8	15.8	3.1	9.5	1.5	6.0
The procedures used during the Tulsa Model Evaluation are fair.	65.6	29.7	24.4	34.4	6.9	20.6	1.5	10.1	0.8	5.2
The procedures used during the Tulsa Model Evaluation promote ethical practices.	67.9	37.4	27.5	40.2	2.3	13.0	0.8	5.8	0.0	3.6
I believe my school has implemented the Tulsa Model Evaluation system in an effective manner.	61.8	36.4	31.3	36.5	3.1	15.8	3.1	6.7	0.0	4.6
The domains and dimensions identified in the Tulsa Model Evaluation are representative of best professional practice and help define effective teaching.	53.5	28.2	37.2	40.5	6.2	15.4	2.3	9.7	0.8	6.3

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Table C. 2

Principal and Teacher Responses on Qualifications of Evaluator Leadership as a Percentage of the Sample

Describe your perceptions of the knowledge, expertise and supportiveness of the individual conducting your most recent performance evaluation.

As an individual conducting performance evaluations for teachers, describe your perceptions of the knowledge, expertise and supportiveness in conducting your most recent performance evaluations.

	Agree		Somewhat Agree		Somewhat Disagree		Disagree		Strongly Disagree	
	P	T	P	T	P	T	P	T	P	T
My evaluator has a thorough knowledge of the Tulsa Model domains and dimensions.	74.2	58.0	24.2	30.1	0.8	6.1	0.0	3.6	0.0	2.2
My evaluator knows well the procedures to implement the Tulsa Model Teacher Evaluation system.	81.5	56.4	14.5	31.0	3.2	6.6	0.0	3.6	0.0	2.4
My evaluator spends adequate time observing my instruction in order to form a basis to assess my performance related to the Tulsa Model domains and dimensions.	61.3	45.9	26.6	26.8	7.3	11.3	3.2	9.9	0.8	6.1
I am generally satisfied with the discussions I have had with my evaluator regarding my performance in relation to the Tulsa Model domains and dimensions.	66.9	55.2	27.4	25.1	3.2	8.8	2.4	6.9	0.0	4.1
I receive adequate support from my evaluator to meet the expectations identified in the Tulsa Model domains and dimensions.	58.9	51.1	34.7	27.0	4.0	10.7	1.6	6.7	0.0	4.5
It has been my experience that evaluations are conducted in a non-threatening manner with a focus on continuous improvement.	77.4	64.3	20.2	21.3	0.8	7.7	0.0	3.0	0.8	3.8
My evaluator has the knowledge and technical expertise as an instructional	70.2	52.4	28.2	26.0	0.0	12.9	0.0	4.9	0.8	3.9

leader to guide the professional growth needed to improve my teaching practice.

Note. P = Principals, T = Teachers
Principals n = 134, Teachers n = 718

Figure C.1

Mean of Knowledge and Understanding of Tulsa Model Evaluation System

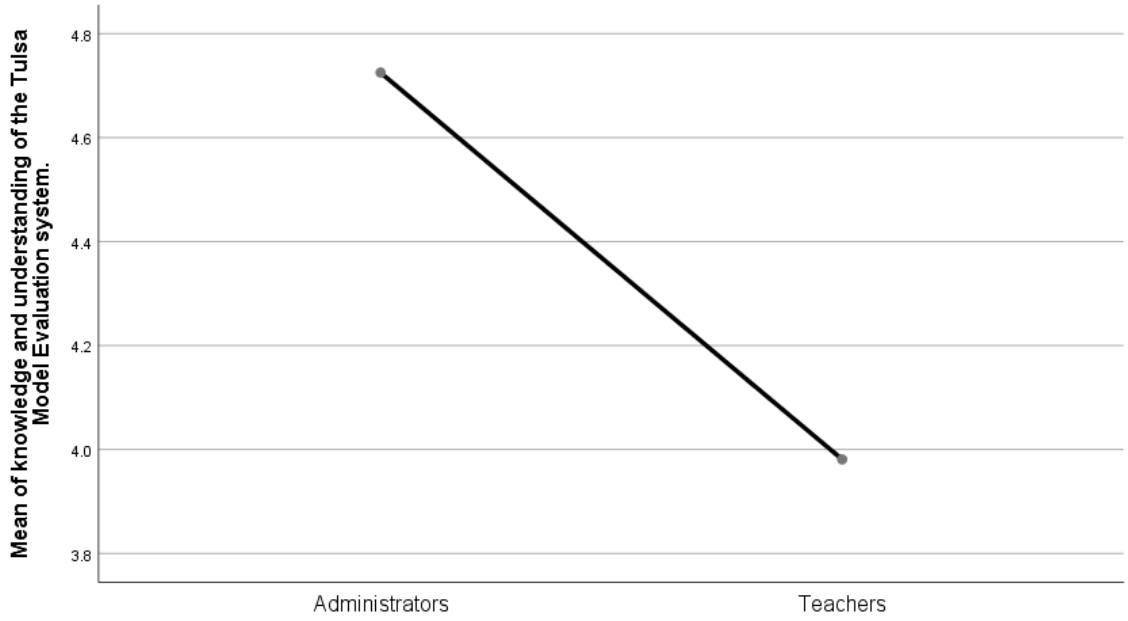

Figure C.2

Mean of Perceptions of Experiences Using the Tulsa Model Evaluation System

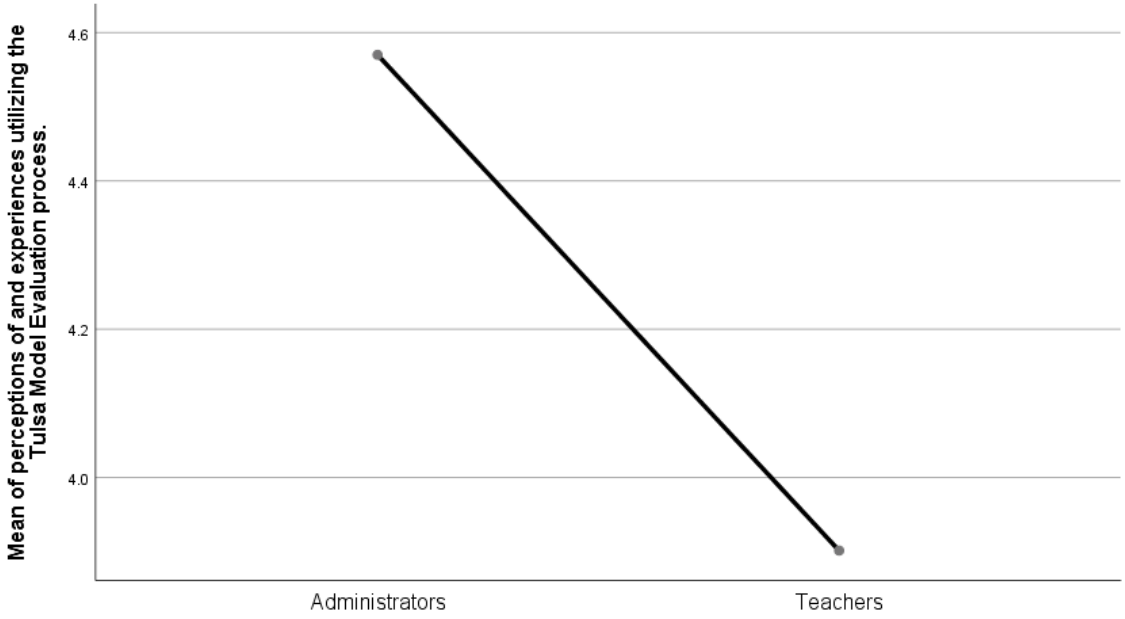

Figure C.3

Mean of Qualifications of the Evaluating Administrator

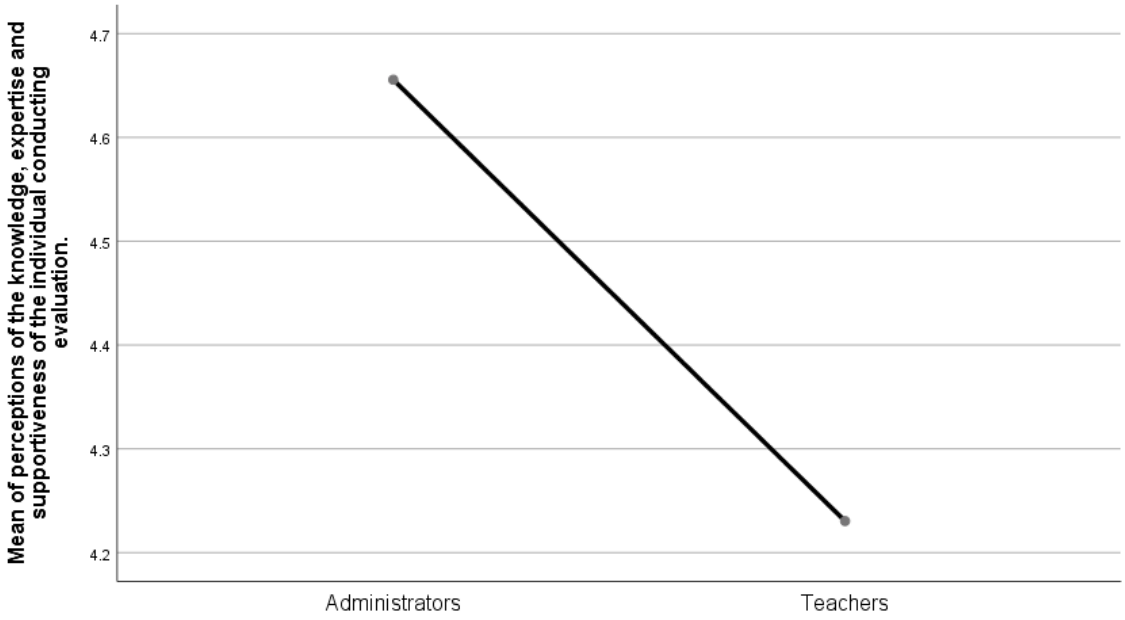

Figure C.4

Mean of Perceptions Regarding the Quality of Data and Feedback

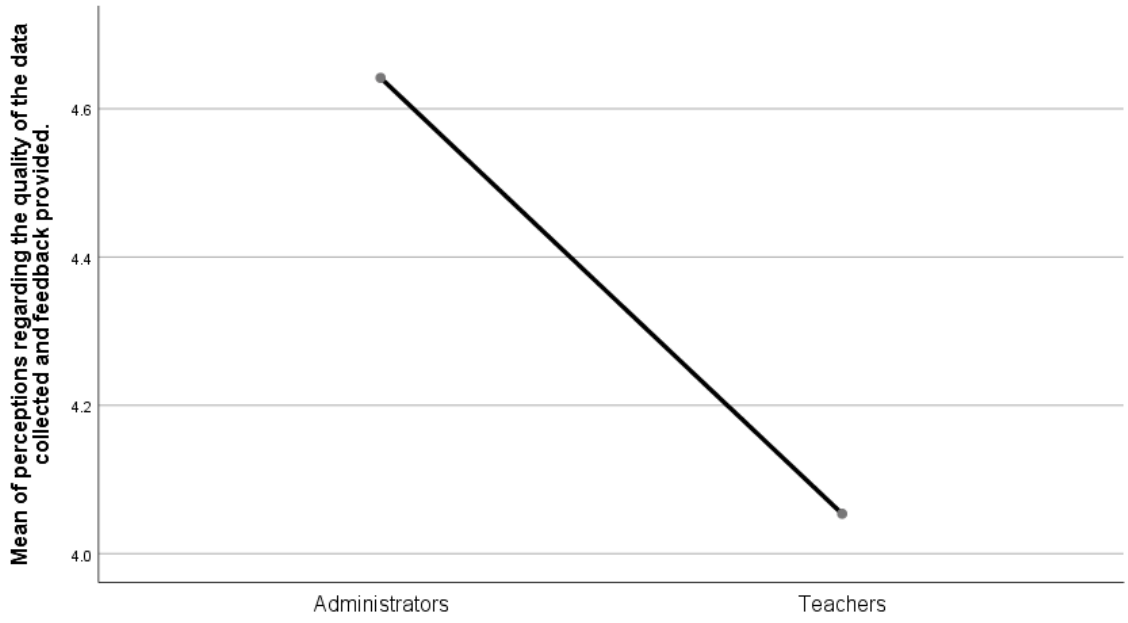

Figure C.5

Mean of Perceptions of Outcomes of the Evaluation Process

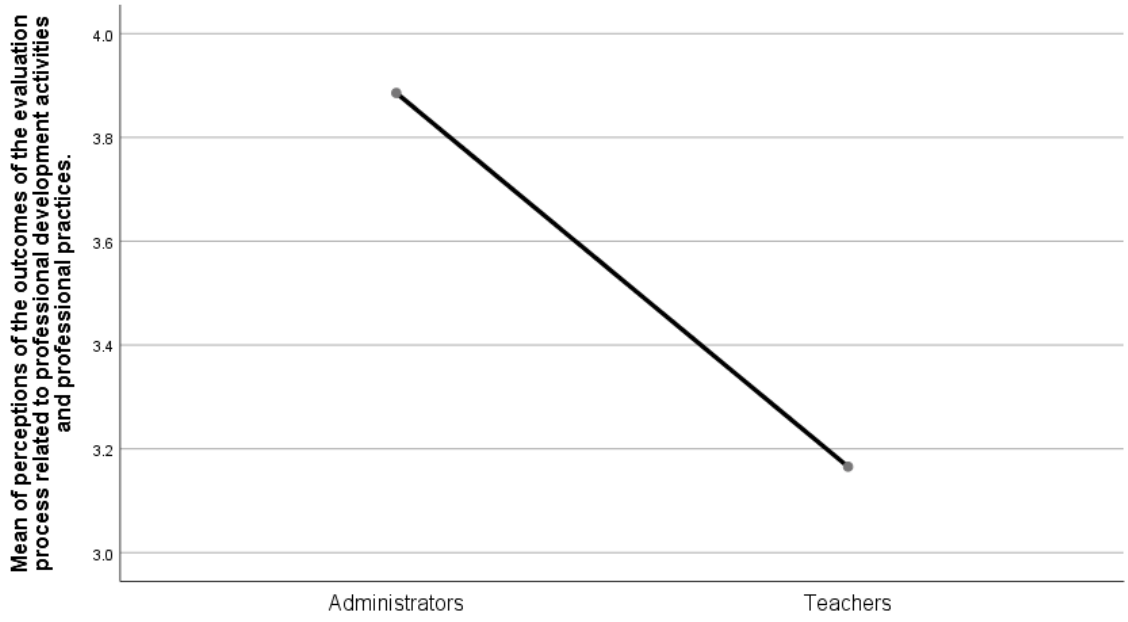

Figure C.6

Mean of Perceptions Regarding Impact of the Tulsa Model Evaluation System