EXAMINING TEACHER EFFECTIVENESS

THROUGH VALUE-ADDED SCORES AND

OBSERVED TEACHING PRACTICES

By

AMY CASTRO BRAUN

Bachelor of Science in Education
University of Oklahoma
Norman, Oklahoma
2002

Master of Science in Instructional Leadership and
Academic Curriculum
University of Oklahoma
Norman, Oklahoma
2004

Master of Science in Educational Leadership
University of Central Oklahoma
Edmond, Oklahoma
2013

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF EDUCATION
May, 2022

EXAMINING TEACHER EFFECTIVENESS

THROUGH VALUE-ADDED SCORES AND

OBSERVED TEACHING PRACTICES

Dissertation Approved:

Dr. Katherine Curry

Dissertation Adviser

Dr. Ed Harris

Dr. Mary Jo Self

Dr. Christine Ormsbee

.

ACKNOWLEDGEMENTS

Thank you to my incredibly supportive and loving parents, Mike and Debi Castro, for raising a strong and confident daughter who believes she can accomplish anything.

To my loving husband, Justin Braun, who lets me be me even when my never-ending commitments and adventures keep me away from spending quality time with you. To the most precious gifts I have ever been given, Jackson and Charles, you have never known a life without your mother either working on her National Boards or being in graduate school. You also share me every day with hundreds of other children and never complain. You both are my world and I hope I am setting a good example for you of how to live a life dedicated to serving others.

My sincere appreciation goes to my OSU OKC Cohort! What a blessing it was to be on this journey with you! Michael Branch, Lisa Crosslin, Jennifer Cruz, Brian Jones, Marcus Macias, Roni McKee, Vicki Nighswonger, Tony Rose, Gabriel Schmidt, Becky Slothower, Terry Williamson, and Melissa Yarbrough – I am so honored to have you as dear friends and esteemed colleagues.

Thank you to Dr. Curry, Dr. Harris, Dr. Ormsbee, and Dr. Self for serving on my dissertation committee and providing your feedback and expertise! I have much respect for all of my professors at OSU and truly value the wisdom and guidance you provided to me and will continue to bestow upon the next generation of education leaders!

To all my colleagues in Moore Public Schools and across the state of Oklahoma, I am indebted to you for constantly inspiring me to be the best teacher and leader that I can be and forever grateful for your constant support and friendship. Thank you for your unwavering dedication and service to our students!

To David Peak and Michael Rogers – most public school educators and administrators will never understand the enormity of the task we conquered with HB2957. I cannot fathom the detriment on the morale of teachers if student growth measures would have been a required component of TLE during the teacher walkout and COVID-19. I consider HB2957 one of our greatest achievements and was fortunate to play a role in its passage into state law.

Last but not least, I believe “I can do all things through Christ who strengthens me.”

Name: AMY CASTRO BRAUN

Date of Degree: MAY, 2022

Title of Study: EXAMINING TEACHER EFFECTIVENESS THROUGH VALUE-ADDED SCORES AND OBSERVED TEACHING PRACTICES

Major Field: SCHOOL ADMINISTRATION

Abstract: Value-added assessment is designed to measure teacher contributions to student achievement in order to promote effective teaching (Battelle for Kids, 2011b; Darling-Hammond et al., 2012; Lee, 2011). When value-added assessment is used, research indicates that in some cases effective teaching is promoted, but in other cases it is not (Amrein-Beardsley & Collins, 2012; Darling-Hammond et al., 2012; Betebenner et al., 2012; McCaffrey & Hamilton, 2007; Quattrochi & Chapman, 2010). One reason that effective teaching may not be promoted is value-added models are not specifically designed to be diagnostic tools of effective teaching (Betebenner et al., 2012; Darling-Hammond, et al., 2012; Goe, 2008; RAND Corporation, 2004). Empirical evidence about which specific teaching practices improve student learning is lacking and additional research is needed (Goe, 2008; RAND Corporation, 2004; Stronge, Ward, & Grant, 2011).

The purpose of this study was to examine the relationship between value-added scores for elementary, junior high, and high school teachers of English/language arts, reading, and/or math in a large suburban Oklahoma public school district and administrators' ratings of their specific teaching practices as measured by the Tulsa Model for Observation and Evaluation. There were small, positive correlations that were statistically significant for all teachers between overall value-added scores and overall evaluation scores on the Tulsa Model rubric for both school years. This result indicates that higher ratings of effective teaching were slightly associated with higher overall value-added scores and increased student achievement levels for this sample. Only two of the Tulsa Model rubric domains – classroom management and instructional effectiveness – had statistically significant relationships with value-added scores. This is attributed to the student-focused nature of these domains in contrast to the professional growth, interpersonal skills, and leadership domains. The specific teaching practices of effective teachers in the areas of Preparation, Discipline, Modeling, Monitoring, and Adjusts Based Upon Monitoring were linked to increased value-added scores.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Experts in the field of education agree that teacher effectiveness is the single most important factor in improving student achievement (Oklahoma State Department of Education, 2013; Weisberg, Sexton, Mulhern, & Keeling, 2009). However, determining exactly what practices constitute effective teaching has been the focus of research for decades with the goal of identifying factors that lead to a quality education for all students that enhances student performance (Stronge, Ward, & Grant, 2011). Because of the relationship that has been established through research findings between teacher effectiveness and student achievement (Goe, 2008; Lee, 2011), public school reform has shifted its focus from solely inspecting student achievement data to evaluating teacher effectiveness through students' standardized test scores. Teacher evaluation policy is now centered on the notion that effective teachers will improve student achievement (Lee, 2011), therefore, leading to an emphasis on the "added value" that a teacher contributes to student learning. After Race to the Top grant funding was touted by the Obama administration in 2009 to states who adopted teacher evaluation models linked to student achievement, states like Oklahoma rapidly implemented value-added assessment models to determine how much a particular teacher contributed to students' learning as demonstrated through standardized test scores (Duffrin, 2011; Fryer, 2011; Hovanetz &

Hellman, 2012; Rubin, Stuart, & Zanutto, 2004). However, just a few years after adoption in Oklahoma, state legislation reversed the demand for the use of value-added measures in the statewide evaluation system due to the enormous costs and made it an optional component for school districts (Oklahoma House of Representatives, 2016).

The purpose of value-added assessment is to use statistical methods to examine and assess learning trajectories as demonstrated by student progress on standardized tests over time through different classrooms, teachers, schools, and districts (Amrein-Beardsley, 2008; Battelle for Kids, 2011b; Doran & Fleischman, 2005; McCaffrey & Hamilton, 2007; Rubin et al., 2004; Yeh, 2012). In theory, value-added assessment models measure teaching effectiveness and reveal quantifiable differences among teachers yielding comparative results of classrooms that perform above, at, or below expected predictions and patterns of growth (Battelle for Kids, 2011a; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Lee, 2011). Education policymakers view value-added assessment as a way to meet the need for a fair and complete evaluation system to account for both achievement and growth by determining the teacher's direct influence on learning (Battelle for Kids, 2011a).

In regard to teacher evaluation, value-added results could provide teachers with a clearer understanding of their strengths and challenges and allow for more concentrated improvement efforts through targeted professional development (Hovanetz & Hellman, 2012; Oklahoma State Department of Education, 2013). Additionally, proponents of value-added models believe that value-added data has the potential to influence teaching in the following ways:

- provide educational leaders with a clearer understanding of a teacher's assets and weaknesses allowing them to focus on improvement efforts;

- reveal a more in-depth evaluation of where curriculum and programs are more or less effective;
- present an assessment of professional development needs of teachers;
- and offer a chance to improve placement of teachers and students to maximize the impact of the highly effective educators. (Hovanetz & Hellman, 2012)

In theory, value-added measures identify highly successful teachers, and this knowledge should create opportunities from colleagues to learn from those teachers (Goe, 2008). The data also has the potential to reveal whether teachers who have participated in specific professional development experiences and implemented certain instructional programs have a positive impact on student achievement (Goe, 2008).

However, many educational researchers are wary of the limitations of value-added data and what it does not reveal about teacher effectiveness, especially in relation to teaching practices linked to improved student achievement (Amrein-Beardsley & Collins, 2012; Betebenner et al., 2012; Goe, 2008). Those who question the benefit of using value-added data in teacher evaluation programs argue that there is very little that educators and administrators can learn simply from seeing a teacher's value-added score because it gives no indication of what the teacher might have done that can be specifically tied to student learning (Goe, 2008). Amrein-Beardsley and Collins (2012) noted that teachers within Houston's teacher evaluation system learned little about how to improve their instruction or what they did effectively after receiving value-added results. In a recent Institute of Education Sciences meeting, the nations' leading researchers of measuring teacher effectiveness reported that value-added results were not useful in providing teachers with feedback to improve their practice; teachers do not understand the link between their value-

added score and instructional actions; and value-added data provides no information on what teachers need to do in order to improve student outcomes (Betebenner et al., 2012). Although numerous empirical studies have analyzed the value-added impact of teachers on gains in student achievement, few have addressed the matter of determining exactly what effective teachers do differently in their classrooms (Stronge et al., 2011).

Statement of the Problem

Value-added assessment is designed to measure teacher contributions to student achievement in order to promote effective teaching (Battelle for Kids, 2011b; Darling-Hammond et al., 2012; Lee, 2011). According to Berliner (2013), the logic behind value-added measures of teacher effectiveness is quite simple. Students should show growth in their knowledge and skill from year to year. Students with similar test scores on a standardized test one year should exhibit comparable growth patterns the following year. If students' standardized test scores reveal more growth than the average peer, then those students must have been taught by more effective teachers. Those students who show less growth from year to year must have been taught by less effective teachers (Berliner, 2013). By analyzing value-added data, teachers and administrators can assess the impact of their instruction and curriculum on student achievement (Battelle for Kids, 2011b).

When value-added assessment is used, research indicates that in some cases effective teaching is promoted, but in other cases it is not (Amrein-Beardsley & Collins, 2012; Betebenner et al., 2012; Darling-Hammond et al., 2012; McCaffrey & Hamilton, 2007; Quattrochi & Chapman, 2010). The American Statistical Association (2014) claimed that teachers and schools can target new teaching techniques and focus on specific professional development opportunities when value-added data is used appropriately. Quattrochi and

Chapman (2010) confirmed that Ohio teachers used value-added data to more closely monitor student progress through formative benchmarks and differentiate instruction by providing appropriate interventions and enrichments to maximize student learning. Unfortunately, very few descriptions of how value-added data is actually used to promote effective teaching exist in educational research. Lee (2011) reported that teachers' negative attitudes towards the value-added process can potentially hinder positive use of results. After their value-added results were published in the *Los Angeles Times*, Los Angeles teachers perceived negative implications of the value-added process including: increased competition among teachers, unfair branding of teachers, narrowing of the curriculum to teach to the tests, enticement of cheating as a way to game the system, and lack of recognition for their efforts and dedication to the profession (Lee, 2011).

One reason that effective teaching may not be promoted through the use of value-added data is value-added models are not specifically designed to be diagnostic tools of effective teaching (Betebenner et al., 2012; Darling-Hammond et al., 2012; Goe, 2008; RAND Corporation, 2004). In a study of Houston's teacher evaluation system, Amrein-Beardsley and Collins (2012) noted the uncertainty of teachers using value-added data to modify and improve instructional practice. Their investigation revealed that Houston teachers were not able to pinpoint or describe what they did differently in terms of instruction from year to year to receive fluctuation in value-added scores (Amrein-Beardsley & Collins, 2012). Educational researchers from across the nation reported at an Institute of Education Sciences meeting that value-added measures did not provide teachers with feedback to improve their practice (Betebenner et al., 2012). The American Statistical Association (2014) also acknowledged that value-added scores do not provide information for teachers on how

to become more effective. In Oklahoma, a recent report from the Southern Regional Education Board (2014) about the state's evaluation system exposed that teachers did not understand how value-added scores were calculated and had no support in interpreting the value-added scores and reports.

Oklahoma's Teacher/Leader Effectiveness evaluation system, also referred to as the "Tulsa Model," was adopted for statewide use in 2010 and provides a unique and unexplored context to examine potential relationships between a teacher's value-added score and specific teaching practices associated with improved student achievement. According to Tulsa Public Schools (n.d.-a), over 90% of Oklahoma public school districts used the Tulsa Model for Observation and Evaluation as of 2012. The Tulsa Model is an "evidence-based process of educator evaluation anchored in specific domains, dimensions, and indicators reflecting national best practices and current research regarding effective instruction" (Tulsa Public Schools, n.d.-b). The overarching domains of the Tulsa Model that reflect best practices of teaching are classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership. Within each domain are specific dimensions of teaching that are defined within a rubric that classifies different levels of teaching performance as ineffective, in need of improvement, effective, highly effective, and superior (Tulsa Public Schools, 2015). The research behind the Tulsa Model is based on findings from the Northwest Regional Education Lab, Harvard professor Thomas Kane and colleagues, Charlotte Danielson, and Robert Marzano (Tulsa Public Schools, 2015).

Purpose of the Study

The purpose of this study is to examine the relationship between the value-added scores for elementary, junior high, and high school teachers of English/language arts,

reading, and/or math in a large suburban Oklahoma public school district and administrators' ratings of their specific teaching practices as measured by the Tulsa Model for Observation and Evaluation. Understanding if there is a relationship between value-added data and administrator perceptions of effective teaching is important because little understanding exists regarding the implications of value-added data on specific teaching practices. Additionally, even though value-added models are designed to control for contextual and learning variables that are out of the teacher's control, value-added data have been criticized as being influenced by those variables resulting in unfair evaluation practices. This study can provide insight into the use of value-added data as a measure of teaching quality by examining the relationship between administrator perception of effective teaching and gains or losses in student outcomes. By associating specific teaching practices with improved student learning as measured by value-added scores, this study will contribute much needed data to the field and potentially provide a narrower focus for teacher enhancement and improvement that will have the greatest impact on student learning. Both educators in the field as well as pre-service teachers could benefit from this information.

Epistemological Perspective

According to Crotty (1998), epistemology is the theory of knowledge embedded within the theoretical perspective and methodology of the research. Objectivism is the epistemological view that meaning and reality exists apart from the operation of consciousness and that scientific research can attain that objective truth and meaning. An objectivist researcher seeks to represent facts and not be influenced by personal feelings or opinions. The objectivist viewpoint is the basis of quantitative research (Huglin, 2003). Creswell (2014) stated that quantitative research serves as a way of testing objective theories

by examining the relationship among variables. These variables can then be measured so numbered data can be analyzed using statistical procedures. Observations are expected to be neutral descriptions and contain objective reports of the relationships among variables (Huglin, 2003).

In this study, the value-added scores provide an objective measure of a teacher's unique contributions to student learning as measured on impartial standardized tests. The value-added process controls for factors unrelated to teacher performance and is designed to isolate a teacher's value from other factors that might affect a student's test score that are out of the teacher's control (Oklahoma State Department of Education, 2014). A teacher's ratings on the Tulsa Model rubric should be objective measures of teaching practices based on the descriptive language within the rubric's levels of effectiveness. Administrators observing the teachers must participate in and pass calibration tests as part of their training before being allowed to evaluate a teacher (Empirical Education, n.d.). The Tulsa Model itself has been independently validated by the Measures of Effective Teaching (MET) Validation Project and University of Wisconsin's Value-added Research Center and confirmed to measure teacher practices that track student achievement growth (Tulsa Public Schools, n.d.-c)

Theoretical Framework

Total Quality Management (TQM) is a continuous improvement philosophy for organizing the whole in order to achieve excellence as well a set of guidelines for ongoing improvements for the services/products offered to customers (Farooq, Akhtar, Ullah, & Memon, 2007). The mindset behind TQM is that employees want to provide their best work and that it is the management's job to inspire this excellence in the workplace through a combination of human resources, quality assurance methods, and collaborative efforts aimed

at satisfying the customers' needs (Farooq et al., 2007; Sallis, 2002). In the educational setting, TQM has three main components: meeting the expectations of both external and internal stakeholders (customers), emphasizing quality assurance to specific standards, and highlighting the means for measurement based on academic performance and feedback (Manaf & Seng, 2010).

According to Sallis (2002), there are four quality imperatives of TQM related to the field of education – moral, professional, competitive, and accountability. For this study, the moral imperative is that every student receives the best possible education. The professional imperative centers on a commitment to students' needs and an obligation to meeting those needs by implementing appropriate pedagogical practices, which is in line with developing more effective teachers. The competitive imperative focuses on the needs of the clients (parents and other community stakeholders) and developing strategies to differentiate public schools from their competitors (other public schools/districts, charter schools, and private schools). The accountability imperative involves objective and measurable outcomes of the educational process and provides mechanisms for quality improvement, such as student achievement results linked to value-added scores within the state teacher evaluation system (Sallis, 2002).

Proponents of TQM believe performance measurement and quality monitoring in the form of teacher evaluation and student achievement are essential to detecting and eliminating variability in the teaching and learning process. These theorists favor using statistical processes to measure and eliminate variability in the manufacturing process, which is similar to the notion of using value-added models to measure teacher effectiveness (Sallis, 2002). Sallis (2002) purported that using a value-added approach allows teachers and administrators

to set realistic performance targets for students based on past data, chart student progress, and take action if a student is under-achieving. Value-added models are useful in improving educational quality for students because the difference between a students' performance upon entry and exit of a course are measured and teachers/schools can still do well on value-added measures regardless of the students' abilities and background (Sallis, 2002).

Research Questions and Hypotheses

Context

In Oklahoma under the statewide Teacher/Leader Effectiveness evaluation system, school districts measure teacher effectiveness using both value-added scores related to student standardized test scores and ratings of teaching practices from administrator observations of classroom performance (Glisson & White, 2014). In most Oklahoma school districts including the large suburban district in this study, the Tulsa Model for Observation and Evaluation is used to rate teachers under five varying levels of effectiveness within five overarching domains that include 20 specific indicators of teaching practices known to improve student achievement. Those domains include classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership (Tulsa Public Schools, 2015). Appendix A provides a complete list of teaching practice indicators within each domain of the Tulsa Model for Observation and Evaluation.

Overarching Research Question

Is there a relationship between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores?

Sub-questions

1. Is there a relationship between teachers' ratings on the classroom management

domain of the Tulsa Model for Observation and Evaluation and their value-added scores?

2. Is there a relationship between teachers' ratings on the instructional effectiveness domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
3. Is there a relationship between teachers' ratings on the professional growth and continuous improvement domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
4. Is there a relationship between teachers' ratings on the interpersonal skills domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
5. Is there a relationship between teachers' ratings on the leadership domain of the Tulsa Model for Observation and Evaluation and their value-added scores?

Hypotheses

- H_1. There will be no relationship between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores
- H_2. There will be no relationship between teachers' ratings on the classroom management domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_3. There will be no relationship between teachers' ratings on the instructional effectiveness domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_4. There will be no relationship between teachers' ratings on the professional

growth and continuous improvement domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

- H_5. There will be no relationship between teachers' ratings on the interpersonal skills domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_6. There will be no relationship between teachers' ratings on the leadership domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

Methodology

This quantitative correlational study is designed to examine the relationship between teachers' value-added scores and administrator's ratings of their teaching practices on the Tulsa Model for Observation and Evaluation. With approximately 41,000 total teachers in Oklahoma, about 20% of these teachers receive value-added scores. The population of this study is one large, suburban school district in Oklahoma that utilizes value-added data in teacher evaluation. Data was provided by a large suburban school district in Oklahoma in the form of spreadsheets with value-added scores and Tulsa Model for Observation and Evaluation scores for the 2013-2014 and 2014-2015 school years. The sample included 439 teachers of the following courses who receive value-added scores from the state - fourth through eighth grades Reading, English/Language Arts, and Math; Algebra I; Algebra II; Geometry; and English III. (For the state of Oklahoma, these courses were the only ones to have a value-added score calculated for the teacher.) To investigate the overarching question, the independent (predictor) variable is a teacher's overall effectiveness score of teaching practices on the Tulsa Model for Observation and Evaluation rubric, and the dependent

(criterion) variable is the teacher's overall value-added score as measured by the students' performance on specific state assessments. To explore the sub-questions, a relationship between the predictor variable of a teacher's domain score on the Tulsa Model rubric and the criterion variable of the teacher's overall value-added score was examined.

Using SPSS, a bivariate correlation coefficient in the form of a Pearson Product Moment Correlation was calculated as a way to express the size and strength of the relationships among variables. Additional correlation coefficients were calculated for each indicator of teaching practices within that domain. For example, within the classroom management domain of the Tulsa Model rubric, additional calculations were conducted on specific practices within that domain – preparation, discipline, building-wide climate, lesson plans, assessment practices, and student relations – to further examine relationships among the variables. Due to the correlational design of this study, it is important to note that any relationships between value-added scores and rating of teacher practices cannot be taken as causal evidence.

Significance of the Study

Significance to Theory

Because the philosophy underpinning Total Quality Management is the need for continuous improvement to achieve organizational excellence (Farooq et al., 2007), it is crucial that specific teaching practices be identified that have a positive relationship with increased student achievement. In order to meet the three components of TQM in an educational setting (Manaf & Seng, 2010), stakeholders expect student knowledge and skills to increase from year to year, the quality of the teaching force based on observed indicators of teaching practices linked to improved student learning such as those on the Tulsa Model

for Observation and Evaluation must be assured, and value-added models can serve as a means for measurement based on academic performance. Educational researchers must investigate the potential relationship between value-added measures of teacher effectiveness with observed teaching practices so that current and pre-service educators will have information that may lead to continuous pedagogical improvement and increased student achievement.

Significance to Research

Educational researchers have stated that value-added models focus exclusively on student learning. Value-added data is limited on what it can tell educators and administrators about teacher quality and effectiveness and should be supplemented by other sources of evidence. Empirical evidence about which specific teaching practices improve student learning is lacking (Betebenner et al., 2012; Goe, 2008; RAND Corporation, 2004; Stronge et al., 2011). This study seeks to identify higher effectiveness ratings of specific teaching practices as measured by the Tulsa Model rubric that are correlated with higher value-added scores. Higher value-added scores for teachers imply improved student learning. By attempting to associate specific teaching practices with improved student learning, this study will hopefully contribute much needed data to the field that is currently lacking.

Significance to Practice

The Oklahoma State Department of Education (2013) developed an Educator Effectiveness Theory of Action that purports that the single most important factor of student academic achievement is teacher effectiveness. Oklahoma's teacher evaluation system is based on the notion that every child deserves an effective teacher in each subject during each school year. This effectiveness can be measured via the qualitative evaluation of teacher

practices using the Tulsa Model for Observation and Evaluation and quantitative data yielding value-added scores of a teacher's contributions to student learning. Although value-added models have proven useful in determining the effectiveness of teachers in relation to student test scores, the data is not useful in providing teachers with feedback to improve their practice (Amrein-Beardsley & Collins, 2012; Betebenner et al., 2012; Darling-Hammond et al., 2012). Upon the identification of specific teaching practices of effective teachers that are linked to positive student achievement, other educators and students in pre-service programs can work on further developing those practices within their pedagogical repertoire to improve student learning. Administrators can also facilitate professional development and peer observation opportunities focused on specific knowledge and skills that will improve student learning.

Definition of Terms

The following definitions are included to facilitate familiarity and understanding with the terms commonly used throughout this study.

- Value-added models/assessment: a measure of a teacher's contribution to students' academic growth; isolates a teacher's contribution from factors not within the teacher's control such as gender, race/ethnicity, attendance, mobility, special education classification, English language learner status, and free and reduced lunch status; compares the achievement of teacher's students to an estimate how those same students might have achieved with an average teacher (Walsh, Liu, & Dotter, 2015).
- Value-added score: calculated by averaging each student's actual performance against his/her expected growth predictions on standardized test scores. Students scoring above the prediction count as positive credits to the value added by the teacher;

whereas students scoring below the prediction are negative debits. Students scoring as predicted maintain the teacher's value-added score as zero. (Lee, 2011; MET, 2013b).

- Value-added report: provides the teacher with a value-added score, information about their performance relative to other teachers of the same content area, comparison data between students' scores on the state test and the value-added result, and value-added results for specific subgroups of students; refer to Appendix B for a sample value-added report (Oklahoma State Department of Education, 2014).
- Value-added teachers/educators: teachers of the following courses are considered value-added teachers/educators in Oklahoma and receive a value-added score/report - fourth through eighth grades Reading, English/Language Arts, and Math; Algebra I; Algebra II; Geometry; and English III (Walsh et al., 2015).
- Tulsa Model for Observation and Evaluation/Tulsa Model rubric: In Oklahoma, districts may adopt this rubric as a way for administrators to qualitatively measure teaching effectiveness within five overarching domains (classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership) that encompass 20 specific indicators of teaching practices known to improve student achievement as indicated on independent validation studies (Tulsa Public Schools, n.d.-b). Refer to Appendix A for a detailed list of the indicators of effective teaching included in each domain.
- Teacher effectiveness: In research, teacher effectiveness typically refers to a teacher's instructional expertise, classroom learning environment, and student achievement (Stronge et al., 2007). For the Tulsa Model for Observation and Evaluation and Oklahoma value-added reports, a teacher is considered effective with a score of 3.0 or

higher on a scale of 1.0 through 5.0.

- Teacher/Leader Effectiveness (TLE): name of the Oklahoma teacher evaluation system. Value-added is one of the measures used to obtain quantitative data about a teacher's effectiveness in order to make personnel decisions about the retention and exiting of an educator in the profession. The goals of the system are to assess educator's strengths and areas needing improvement, offer individualized high-quality professional development opportunities, and provide meaningful, actionable feedback to teachers in order to promote continuous learning within the profession (Oklahoma State Department of Education, 2013).

Summary

Chapter I introduced the concept of value-added models as a measure of teacher effectiveness and established how this study will contribute to the void of empirical evidence regarding specific teaching practices that are positively related to improved student achievement. The statement of the problem and purpose of the study were provided. A brief overview of Total Quality Management was explained as the theoretical underpinning of this study. The research questions and hypotheses were advanced and a brief overview of the methodology was provided. The significance of the study as related to theory, research, and practice as well as the definition of terms were also included.

Chapter II of this study provides a review of the literature on value-added measures of teacher effectiveness. A more developed explanation of Total Quality Management is also included in this chapter. Chapter III details the research design, describes the participants and instrumentation, and explains the procedures and how data will be analyzed. Assumptions and limitations of the study are also addressed. Chapter IV presents findings from the

descriptive and correlational analysis. Chapter V discusses findings through the lens of Total Quality Management. It concludes with implications for practice and further research.

.

CHAPTER II

REVIEW OF LITERATURE

This first part of literature review highlights the educational research on value-added models of teacher effectiveness. It details how the national focus on educator effectiveness came to exist over time, defines what value-added measures of teacher effectiveness are, explains validity and reliability concerns about value-added models in general, describes how value-added results can promote effective teaching, and addresses the drawbacks of value-added models as diagnostic tools of effective teaching.

The second part of this chapter focuses on the theory of total quality management (TQM), which is a continuous improvement philosophy for organizing the whole in order to achieve excellence as well as a set of guidelines for ongoing improvements for the services/products offered to customers (Farooq, et al., 2007). It explores the origins of TQM, defines TQM, relates TQM to the field of education, discusses TQM in regard to educational research, and includes an explanation of how TQM is linked to this study.

Value-Added Measures and Educator Effectiveness

What Constitutes Teacher Effectiveness?

Although there is consensus among educational researchers that teacher quality matters, there is little agreement on which aspects of teacher quality matter most and in defining exactly what aspects of teachers and/or teaching constitute educator effectiveness

(Goe, 2007; Munoz, Scoskie, & French, 2013). According to Stronge, Ward, & Grant (2011), "Effectiveness is an elusive concept to define when we consider the complex task of teaching and the multitude of contexts in which teachers work" (p. 304).

Goe (2007) purported that there are four distinct ways of examining teacher quality: teacher qualifications (education, certification, credentials, and experience), teacher characteristics (attitudes, attributes, beliefs, race, gender), teacher practices (planning, instructional delivery, classroom management, and interactions with students), and teacher effectiveness linked to student learning outcomes in the form of standardized tests. Other researchers incorporate teacher practices in their definition of effectiveness. Stronge, Ward, & Grant (2011) identified four dimensions of teacher effectiveness: instructional effectiveness, use of assessment for student learning, learning environment, and personal qualities of the teacher. Darling-Hammond acknowledged that effective teachers create a collaborative classroom, understand subject matter, capitalize on prior knowledge, scaffold and support students, and continuously assess student learning (Betebenner et al., 2012). In addition to teacher practices, researchers believe psychological characteristics of teachers in the form of motivation, emotions, self-regulation, and personality are believed to be aspects of effective teaching, but research on their effects on student outcomes is lacking (Bardach, Klassen, & Perry, 2021).

Ultimately, education researchers agree that differences in teacher effectiveness do exist, but are unable to collectively decide which teacher qualifications, practices, and/or characteristics contribute to those differences in teacher effectiveness (Goe, 2007). "There is considerable debate as to whether we should judge teacher effectiveness based on teacher inputs…the teaching process…the product of teaching…or a composite of these elements"

(Stronge, Ward, & Grant, 2011).

A National Focus on Teacher Effectiveness

In 1965, President Lyndon B. Johnson signed the Elementary and Secondary Education Act (ESEA) into federal law (Brenchley, 2015; Hunt Institute, 2016). The law primarily focused on providing federal government funding through Title I to local education agencies for the education of students from disadvantaged families (Hunt Institute, 2016). Less than a year after the Civil Rights Act of 1964, ESEA demonstrated a major commitment by the federal government to provide both quality and equity in educating all of the nation's school-aged children (Brenchley, 2015). Over the next fifteen years, ESEA was amended four times to provide more clarification and guidance regarding the use of Title I funds to ensure money was truly being spent for the benefit of low-income students (Hunt Institute, 2016).

In 1983, the National Commission on Excellence in Education (NCEE) published a report titled "A Nation at Risk" which chastised our country's educational system and declared, "For the first time in the history of our country, the educational skills of one generation will not surpass, will not equal, and will not even approach, those of their parents" (p. 12). The report indicated declining levels of student achievement across America's public school system and noted that average scores on the Scholastic Aptitude Test had steadily dropped by 50 points in the verbal category and 40 points in mathematics area over the past two decades (NCEE, 1983). The report claimed that 23 million Americans were functionally illiterate and that 13% of the country's 17-year-olds were functionally illiterate (NCEE, 1983). In addition, the quality of the nation's teaching force was questioned as too many teachers were from the lowest quarter of graduating

high school and college students; half of new teachers were not qualified to teach math, science, or English; and severe teacher shortages existed in the areas of math, science, foreign languages, gifted/talented, and handicapped (NCEE, 1983; Smith, 2008). In order to strengthen the teaching profession, the report called for raising standards for the training, entry, and professional development of teachers with an emphasis that prospective teachers should focus more on subject-area classes rather than pedagogical classes (Smith, 2008).

From 1981 to 1988, what became known as the "excellence agenda" was pushed by business leaders and civil rights organizations across the nation and was backed by the notion that increased academic rigor would bolster student achievement and ultimately boost the country's economy. The reauthorization of ESEA in 1988 called for states to identify schools that were not making considerable progress toward raising student achievement (Hunt Institute, 2016). From 1989 to 2000, federal focus shifted towards standards-based reform and the establishment of national standards, national tests, and school accountability. President Clinton's reauthorization of ESEA in 2000 granted states more flexibility with their Title I fund in lieu of adopting curriculum standards and installing school accountability measures (Hunt Institute, 2016).

In 2002, federal legislation to reauthorize ESEA titled No Child Left Behind (NCLB) was signed into law and required states to adopt academic standards for mathematics, reading, and science and align assessments with those standards. Students in school districts that received federal funds had to be annually assessed from third to eighth grades and once again in high school in reading and mathematics. Science testing was also required once in elementary, middle school, and high school (Hunt Institute,

2016). Student achievement data had to be reported by student subgroups such as disadvantaged students, students with disabilities, and students from major racial/ethnic groups. By the end of the 2013-2014 school year, all students were expected to meet academic proficiency in reading and mathematics, but states had flexibility to define proficiency within their educational system (Hunt Institute, 2016; Klein, 2015). In terms of teacher quality, NCLB also specified that every teacher of a core academic subject must be "highly qualified" by the 2005-2006 school year, which meant that a teacher was certified and demonstrated proficiency in his/her subject matter. Highly qualified teachers must also be evenly distributed across poor and wealthy schools (Klein, 2015). This part of the legislation emphasized the importance of teacher quality on student success (Smith, 2008). Overall, the NCLB approach was aimed at closing the achievement gap and assumed that achievement data would provide relevant info to gauge how effective schools and teachers were at educating all students (Amrein-Beardsley, 2008; Doran & Fleischman, 2005; McCaffrey & Hamilton, 2007). Although NCLB did provide definitions of a highly-qualified educator, it did not predict or ensure that teachers meeting these criteria would be successful at increasing student achievement (Lee, 2011).

In 2009, a report from the New Teacher Project called "The Widget Effect" was published and expanded upon the nation's failure to differentiate and acknowledge teacher effectiveness (Weisberg et al., 2009). The report declared that "a teacher's effectiveness – the most important factor for schools in improving student achievement – is not measured, recorded, or used to inform decision-making in any meaningful way" (Weisberg et al., 2009, p. 2). The Widget Effect referred to the tendency of school leaders to assume that teacher effectiveness is the same from teacher to teacher and ignore

individual teachers' strengths and weaknesses to ultimately improve instructional effectiveness. In terms of teacher evaluation in the states that were studied, the report concluded that all teachers were rated as effective, instructional excellence was unrecognized, meaningful professional development was lacking, novice teachers were ignored, and poor performance was not addressed. To alter the Widget Effect, the report suggested that evaluation systems should be based on the teacher's effectiveness in promoting student achievement and combined with personnel policies that include the compensation, retention, and dismissal of teachers (Weisberg et al, 2009).

Later this same year, President Obama announced a Race to the Top initiative with $4.4 billion in competitive federal grant funds available to states if student achievement data was linked to teacher evaluations (Duffrin, 2011; Fryer, 2011). The goal was to improve teacher effectiveness by linking teacher evaluation scores to student growth measures and tying personnel decisions about pay, raises, and tenure to achievement data (Fryer, 2011). In its Race to the Top grant application, the state of Oklahoma's plan to improve student achievement emphasized the goal of having an effective teacher in every classroom. Strategies to achieve this goal included establishment of a uniform statewide evaluation system in development by Tulsa Public Schools and the Bill and Melinda Gates Foundation, use of the new evaluation system to make staffing decisions and to revise compensation structures, removal of ineffective teachers by their third year in the profession, and competitive grants for school districts willing to implement pay-for-performance models linking the new statewide evaluation system to student achievement data (Office of the Governor, 2010). It is important to note that the Bill and Melinda Gates Foundation promoted that teacher pay increases should

not be based on seniority or a master's degree, but instead on measures of teacher effectiveness that could be estimated by tracking student progress on standardized tests (Duffrin, 2011). In fact, the Bill and Melinda Gates Foundation provided funding for the Measures of Effective Teaching (MET) Project, an educational partnership committed to investigating better ways to identify and develop effective teaching (MET, 2013b). The MET project released the results of a three-year study on 3,000 teacher volunteers throughout the United States which revealed that students of more effective teachers outperformed students of teachers with low effective estimates. The magnitude of students' actual gains largely corresponded with gains predicted by their performance as measured on standardized tests the previous year (MET, 2013a). Because of results like this and the ability to evaluate one teacher's effectiveness on student learning in comparison to other teachers, value-added assessment measures that accounted for teacher effectiveness on student achievement became an integral component of educator evaluation (Doran & Fleischman, 2005; Hellman & Hovanetz, 2012).

In 2010, the United States Department of Education offered $1.2 billion as part of the Teacher Incentive Fund for states who tied individual and group performance-based incentives for teachers based on student growth and achievement (Chait & Miller, 2009; Fryer, 2011). In 2011, President Obama offered to waive the NCLB requirement of all students being proficient in math and reading by 2014 if states agreed to three tenets: adopting rigorous standards for college and career readiness, focusing improvement efforts on the bottom 15% of low performing schools, and tying teacher evaluations to student performance (McNeil & Klein, 2011). When coupled with the national interest in determining educator effectiveness along with Race to the Top and Teacher Incentive

Fund grants, many states like Oklahoma passed legislation mandating that student growth measures such as value-added models be factored into teacher evaluation systems (Battelle for Kids, 2011a, 2011b; Glisson & White, 2014; Lee, 2011; McCaffrey & Hamilton, 2007; Rothstein, 2004; Yeh, 2012). Oklahoma state law (70 O.S. § 6-101.16) established a statewide evaluation system called the Oklahoma Teacher and Leader Effectiveness (TLE) Evaluation System made up of multiple measures of teacher effectiveness in 2010 – 50% qualitative measures of observable characteristics of teacher performance that are correlated with student achievement (Tulsa Model rubric), 35% quantitative measures of student academic growth based on multiple years of standardized test data (value-added model), and 15% quantitative measures of other academic factors (Oklahoma State Department of Education, n.d.). Despite the proposed and adopted changes to the statewide evaluation system, Oklahoma did not receive a Race to the Top federal grant (United States Department of Education, 2016).

In 2015, President Obama signed the Every Student Succeeds Act (ESSA) into law, which was a reauthorization of the Elementary and Secondary Education Act and replaced No Child Left Behind. ESSA required all students to be taught using high academic standards that will prepare them for success in college and careers (U.S. Department of Education, (n.d.). It allows states to set their own accountability goals to address testing proficiency, English language learner proficiency, and graduation rates while submitting accountability plans to the U.S. Education Department. ESSA eliminated the "highly qualified teacher" definition of NCLB and the requirement of including student outcomes (such as value-added measures) in teacher evaluation systems. Instead of the Teacher Incentive Fund, ESSA promotes the Teacher and School

Leader Innovation Program, which provides grants to school districts that want to incorporate performance pay and other teacher quality improvement measures (Education Week, 2016).

Value-added Measures of Teacher Effectiveness

William Sanders, an agricultural statistician, designed the first value-added assessment model for education purposes in 1982 (Duffrin, 2011). Similar to determining the influence of environmental factors on crop yields, Sanders devised a way to separate the contributions of teachers on student achievement from that of a child's socioeconomic status, race, health, and other family and home factors (Rothstein, 2004). The purpose of value-added assessment is to use statistical regression models to follow students over time through different classrooms, teachers, schools, and districts in order to examine and assess learning trajectories of progress on standardized tests (American Statistical Association, 2014; Amrein-Beardsley, 2008; Battelle for Kids, 2011b; Doran & Fleischman, 2005; McCaffrey & Hamilton, 2007; Rubin et al., 2004; Yeh, 2012). Value-added measures assume that standardized tests are aligned to curriculum and valid and reliable as indicators of student learning (Battelle for Kids, 2011a; Darling-Hammond et al., 2012). Standardized tests also have common procedures for administration which make them objective measures of student achievement and growth (American Statistical Association, 2014). Value-added measures should "level the playing field" as the statistical procedures allow for direct comparisons between schools and teachers – even those with varying populations of students (Doran & Fleishman, 2005). By using the student's prior test score as the regressor, it controls for the student's initial achievement and other background factors that might influence student achievement so the

contribution of schools and teachers on student growth is based on the residual differences in post-test scores (Ballou, Sanders, & Wright, 2004). Analyzation of value-added data allows school leaders and teachers to assess the impact of their curriculum, instruction, and other educational programs on student progress and achievement (Battelle for Kids, 2011b).

Simple value-added models compare predicted achievement of groups of students to their actual level of achievement to measure teacher effectiveness (Battelle for Kids, 2011a)). A teacher's value-added score is calculated by averaging each student's actual performance against his/her expected growth predictions on standardized test scores (Duffrin, 2011; MET, 2013b). Students scoring above the prediction count as positive credits to the value added by the teacher; whereas students scoring below the prediction are negative debits. Students scoring as predicted maintain the teacher's value-added score as zero (Lee, 2011; MET, 2013b). Value-added assessment models measure teaching effectiveness and reveal quantifiable differences among teachers yielding comparative results of classrooms that perform above, at, or below expected predictions and patterns of growth (Battelle for Kids, 2011a; Darling-Hammond et al., 2012; Lee, 2011). Although simple value-added models can be internally gathered by states and school districts, these models do not account for additional factors that might influence student achievement and are substantially affected by missing student test data from year to year (Battelle for Kids, 2011a).

Advanced value-added models are not typically run internally, use multiple prior test scores for students, and specifically account for other student characteristic variables that affect student achievement such as poverty level, English proficiency, special

education status, class size, and even race/ethnicity (Baker, Oluwole, & Green, 2013; Battelle for Kids, 2011a; Duffrin, 2011). The goal is to accurately identify the value-added that should be attributed to a teacher for a student or group of students as opposed to factors outside of the teacher's control (American Statistical Association, 2014; Baker, Oluwole, & Green, 2013). Advanced value-added models typically include a roster verification process in which teachers mark the percentage of each student's instruction that they should be held accountable for which details time in their classroom if not a full academic year to account for student mobility and shared instruction in co-teaching situations (Amrein-Beardsley & Collins, 2012; Battelle for Kids, 2011a). Advanced value-added models allow researchers to measure changes in student test scores over multiple data points while considering student characteristics and other factors that might influence achievement so that the student gains in achievement will be attributed to the specific teacher in order to determine instructional effectiveness (Darling-Hammond et al., 2012).

Value-added assessment became a popular education reform effort because it can be used to improve student achievement by allowing for the identification, selection, and retention of high-performance teachers (Yeh & Ritter, 2009). According to Berliner (2013), the logic is simplistic:

> From one testing occasion to another, students should show growth in their knowledge and skill. Similar types of students should show similar patterns of growth. Those students that show more growth than the average must have better teachers, while those that show less growth than average must have poorer teachers. (p. 235)

Because of the claim of objectively distinguishing between effective and ineffective teachers, value-added measures have been used in high stakes personnel decisions such as hiring/dismissing, compensation and pay-for-performance bonuses, awarding tenure, ranking teachers, and even the closing of schools (American Statistical Association, 2014).

Validity and Reliability Concerns Related to Value-added Results

Many educational researchers have expressed concerns over the validity and reliability of value-added measures. The first issue of validity relates to the standardized tests themselves and whether they are truly aligned with standards, curriculum, and materials used in the classrooms and vertically scaled to consistently measure growth over time (Doran & Fleishman, 2005; Goe, 2007). Berliner (2013) postulated issues with the choice of test items being constantly altered from year to year in order to generate a normal curve distribution in test results. He also discussed a lack of instructional sensitivity of test items used, and that no evidence exists that indicates an effective teacher will increase the pass rate of specific test items. The American Statistical Association (2014) reported that standardized tests do not fully measure student achievement in terms of the total number of curriculum objectives and content standards adopted by the state. The second issue of validity relates to the elimination of all other factors that might contribute to a student's test score except for that of the teacher. The validity of value-added scores depends on how well the regression model adjusts for other factors that might affect or bias a teacher's score. For example, a teacher of gifted students or certain special education students may exhibit only small achievement gains if the regression model does not account for the students' status (American Statistical

Association, 2014). Goe (2007) described possible classroom effects such as peer relations, availability of books and materials, school climate, class size, and other factors out of the teacher's control that cannot be separated and might affect a teacher's value-added score.

In terms of reliability, the value-added ratings of teachers differ substantially from class to class, year to year, model to model, test to test, and between different regression models (American Statistical Association, 2014; Darling-Hammond et al., 2012). Fuller (2012) found that only 19 out of thousands of value-added teachers ranked highly effective every year for four years. In Houston, 46% of teachers who moved either across grade levels or to a different grade level reported switching value-added ranks from ineffective to effective or vice versa (Amrein-Beardsley & Collins, 2012). Value-added models tend to be unstable from year to year with very large error ranges when applied to specific teachers making distinctions between effective and ineffective teachers difficult to determine (Baker, Oluwole, & Green, 2013).

Battelle for Kids (2011b) acknowledges that although value-added measures are not perfect, the data provide an important productivity measure that was previously unavailable to educators. The American Statistical Association (2014) warns educators to remember that value-added models measure correlation and not causation when interpreting results.

Value-added Results and the Promotion of Effective Teaching

There is universal consensus that teacher quality is the most important factor in terms of improving student achievement (Fryer, 2011; Goe, 2007; Oklahoma State Department of Education, 2013; Weisberg et al., 2009). A one standard deviation

increase in teacher quality has been shown to raise reading achievement by 0.15 to 0.20 standard deviations and math achievement by 0.15 to 0.24 standard deviations per year (Fryer, 2011). However, Goe (2007) pointed out that "teacher quality" (certification, teacher test scores, college degrees) differs from "teaching quality" (actual task of teaching and the student learning that teachers foster). Teacher effectiveness is the outcome of both teacher quality and teaching quality and is defined by using value-added measures to rank teachers by how much their students gained compared to how much those students were predicted to grow in achievement (Goe, 2007). Every classroom needs an effective teacher focused on continuous improvement of their craft (Battelle for Kids, 2011b). Effective teaching has also been shown to erase achievement deficits of low-income students (Duffrin, 2011).

Value-added results should provide teachers with a clearer understanding of where one's strengths and challenges are and allow for more concentrated improvement efforts through targeted professional development (Hovanetz & Hellman, 2012; Oklahoma State Department of Education, 2013). Value-added measures identify highly successful teachers, which should create opportunities from colleagues to learn from those teachers. It also has the potential to reveal whether teachers who have participated in specific professional development experiences and implemented certain instructional programs have a positive impact on student achievement (Goe, 2008). Quattrochi and Chapman (2010) confirmed that Ohio teachers used value-added data to provide appropriate interventions/enrichments and/or modify instruction to maximize student learning. These teachers also improved their effectiveness through the value-added results by more closely monitoring student progress through formative benchmarks and

differentiating instruction to address the educational needs of all students. McCaffrey & Hamilton (2007) reported that Pennsylvania teachers who received value-added data were more engaged in data use and test preparation.

Value-added measures provide quantitative information that is relevant to improving education processes. Evidence suggests that when value-added measures are combined with other types of data, it can be used to personalize professional development and coaching opportunities for teachers and guide future curricular and instructional decisions (Battelle for Kids, 2011b; Duffrin, 2011). School districts can determine how vigorously principals are evaluating their teachers on classroom observations in comparison to their value-added data (Duffrin, 2011). Positive correlations exist between value-added data and students' future academic performance (American Statistical Association, 2014). Schools can examine value-added results for specific subgroups of students and determine the best fit between student abilities and available courses as well as future college and career choices (Battelle for Kids, 2011b). Value-added models can also be used to evaluate the effects of teacher training programs by comparing average VAM scores of teachers from different programs (American Statistical Association, 2014).

Unfortunately, very few descriptions of how value-added data are actually used to promote effective teaching exist in educational research. Some researchers believe that providing educators with assistance in understanding and using value-added data and results is the primary challenge in the adoption of a value-added system (McCaffrey & Hamilton, 2007).

Value-added Models Not Designed as Diagnostic Tools

Value-added measures are specifically designed to evaluate student test score gains from one year to the next by using statistical methods and controlling for student characteristics and other factors found to influence student achievement. Value-added results do not provide any diagnostic data for a teacher to know what to do in order to improve his/her teaching practices and become more effective. There is very little evidence that teachers and administrators can learn simply from seeing a teacher's value-added score because it gives no indication of what the teacher might have done that can be specifically tied to student learning (Goe 2007; Goe, 2008). In Houston, teachers could not identify a relationship between teaching practices and value-added ratings. One puzzled educator stated:

> I do what I do every year. I teach the way I teach every year. [My] first year got me pats on the back; [my] second year got me kicked in the backside. And for year three, my scores were off the charts. I got a huge bonus, and now I am in the top quartile of all the English teachers. What did I do differently? I have no clue. (Amrein-Beardsley & Collins, 2012, p. 15)

Andrew Ho of the Harvard Graduate School of Business postulated that it is time to advance from listing symptoms to providing diagnosis and treatment by determining how value-added results can improve teaching. Damian Betebenner of the National Center for the Improvement of Educational Assessment has called for a detailed theory of action on how to use value-added data to transform education (Betebenner et al., 2012). The RAND Corporation (2004) suggested that the selection of effective teachers and/or training to improve teacher effectiveness will have significant impacts in education if value-added

results can be linked to specific teaching practices. Although numerous empirical studies have analyzed the value-added impact of teachers on gains in student achievement, few have addressed the matter of determining exactly what effective teachers do differently in their classrooms (Stronge et al, 2011).

Many researchers claim that value-added measures alone are insufficient to determine the impact of specific teaching practices on student achievement. They call for multiple measures of teacher performance including observations supplemented with value-added measures and student outcomes (Betebenner et al., 2012; Darling-Hammond et al., 2012; Goe, 2008). Battelle for Kids (2011b), a non-profit organization dedicated to the advancement of practices for improving educator effectiveness, suggested that it takes multiple measures to capture the complexities associated with teaching and learning such as value-added measures coupled with principal evaluations. It is very hard to support effective teaching without quality information about actual teaching practices and their link to positive student outcomes (MET, 2013b; Minnici, 2014). In order to improve education, the American Statistical Association (2014) advised that value-added scores must provide meaningful information about a teacher's ability to improve student learning. Goe (2007) purported that value-added scores tell us nothing about why teachers vary in quality or which classroom practices might be used to grow student achievement. There is an overall lack of strong evidence in educational research to show significant positive correlations between teaching practice and student achievement (Goe, 2007).

Total Quality Management in Education

Origins of the Quality Movement

From the 1790s through the 1830s, the United States began transforming from an agricultural society to an industrial economy (Independence Hall Association, 2014). The Industrial Revolution lasted for more than a century and saw the production of goods shift from home-based businesses to machine-aided factories. With the advent of industrialization, the manufacturing process resulted in the narrowing and repetitive-nature of production tasks on the factory line for millions of Americans who had previously handcrafted raw materials into products from start to finish (Independence Hall Association, 2014; Library of Congress, n.d.; Sallis, 2002). Workers no longer had the opportunity to self-check and reflect upon their quality of work. Labor guilds who had previously regulated the quality of workmanship and developed apprenticeship programs for training were no longer needed (Sallis, 2002). Nevertheless, this transformation from hand-made to machine-made products ushered into the United States a much higher standard of living than had ever been known prior to the Industrial Revolution (Independence Hall Association, 2014).

In the early 1900s, Frederick W. Taylor's theory of scientific management focused on quality control and the need to ensure that products conform to specific standards (Sallis, 2002). Taylor's methods broke down the manufacturing process into the smallest possible units and relied on scientific study, not human judgment, of the completion of tasks to determine the optimal amount of work that could be accomplished within a certain timeframe (Dininni, 2011). He asserted that the organization should identify the most efficient methods for completing each task, training workers to

complete in task in a specified manner, and rewarding workers for increased productivity (Dininni, 2011; Juran Institute, 1994). As a result of employing Taylor's methods, industries saw improvements in quality control, productivity, employee incentive, personnel practices, and cooperation between management and workers (Dininni, 2011). However, Sallis (2002) noted that the quality control process highlighted by Taylor was often wasteful and expensive because it was designed to detect defective products at the end of the manufacturing process.

Two future pioneers of Total Quality Management, W. Edwards Deming and Joseph Juran, worked together at Western Electric's Hawthorne plant in Chicago during the late 1920s and early 1930s (Landesberg, 1999; Sallis, 2002). Elton Mayo conducted his famous Hawthorne experiments there to study the effects of lighting levels along with the length of the work day and rest periods on productivity. The results of these experiments purported that management leadership style, group cohesiveness, and attention to workers' needs led to increased productivity. The Hawthorne studies indicated that workers' participation in the decision-making process improved productivity more than monetary incentives or the physical working environment (Sallis, 2002). As a result, focus in the manufacturing process switched from quality control to quality assurance as workers' responsibility for quality during the production process was emphasized (Sallis, 2002). Soldiers returning from the World War II in the late 1940s were more educated, especially in technology, and the production incentives of the quality control method were not as motivational for them. In addition, labor unions rose to power and required that decisions affecting productivity be made through the collective bargaining process (Juran Institute, 1994).

In the 1950s, both Deming and Juran were called to Japan by the Union of Japanese Scientists and Engineers to help them reconstruct their war-torn industry (Landesberg, 1999: Martinez-Lorente, Dewhurst, & Dale, 1998; Sallis, 2002). Deming did not believe in focusing on quality control, but in quality assurance methods and procedures that produced products to the highest level of customer satisfaction (Sallis, 2002). Deming proposed a systematic view of the organization that focused on the interrelationships between consumer research, product design, suppliers, materials, production/assembly, inspection, and distribution (Landesberg, 1999). Juran offered advice on management practices to promote improvements in product and service through the quality functions of market research, product design and development, production, inspection, and sales (Landesberg, 1994). Both Deming and Juran are credited with changing the culture of Japanese industries to one that integrated quality processes into their daily practices (Juran, 2019). Deming focused his teachings on top management statistical methods and a production system that included suppliers and consumers (Landesberg, 1999). Juran went across the country meeting with senior and middle managers to explain how to incorporate quality control practices (Juran, 2019). Both men were awarded the Order of the Sacred Treasure medal by the Japanese emperor for their efforts, which helped reconstruct Japan's postwar economy (Juran, 2019; Lanesberg, 1999). Japan successfully embedded the total quality philosophy into their manufacturing systems and captured a prominent share of the world's markets in automobiles and electronics. In 1980, a documentary titled *If Japan Can Do It, Why Can't We?* aired on national television and highlighted the dominance of Japanese industry while also spurring the total quality movement in the United States (Sallis,

2002). Like many other United States companies at the time, the United States Navy adopted the Japanese total quality philosophies and coined the term "total quality management" in 1985 (Martinez-Lorente et al., 1998).

Total Quality Management

Total Quality Management (TQM) is a continuous improvement philosophy for organizing the whole in order to achieve excellence as well as a set of guidelines for ongoing improvements for the services/products offered to customers (Farooq et al., 2007). The mindset behind TQM is that employees want to provide their best work and that it is the management's job to inspire this excellence in the workplace through a combination of human resources, quality assurance methods, and collaborative efforts aimed at satisfying the customers' needs (Farooq et al., 2007; Sallis, 2002). Sallis (2002) offered two concepts of quality – procedural and transformational. The procedural concept of quality involves ensuring conformity to pre-determined specifications and asking whether the good/service does what is asked/expected. The transformational notion of quality encompassed within TQM focuses on continuous improvement and emphasizes a vision of organizational transformation dedicated to achieving excellence and enriching the entire system (Farooq et al., 2007, Sallis, 2002).

Within the TQM philosophy, the word "total" refers to every employee in the organization being dedicated to the enterprise of continuous improvement. The term "management" implies that every employee in the institution no matter their role or status is a manager of their own responsibilities. TQM extends the notion of quality assurance to the creation of a quality culture where the goal of every staff member is to provide customers with what they want: when and how they want it. In order to create a culture of

continuous improvement, managers must trust the staff and delegate decisions to the appropriate level in order to provide employees with the responsibility to deliver quality within their sphere of the organization. For this culture to be created, the staff must be provided with a suitable work environment with the appropriate and necessary tools of the trade, simple procedures and systems in place that aid in their production of the good or service, and encouragement and recognition of achievements that empowers them as employees to continually improve and seek excellence (Sallis, 2002). Figure 1 provides a graphic depiction of the essential components of TQM.

Total Quality Management in Education

"For education as for industry, quality improvement is no longer an option, it is a necessity" (Sallis, 2002, p. 4). In education, the sources of quality reside in effective teachers, increased student achievement, parental and community support, abundant resources, technologic advances, challenging curriculum, and purposeful leadership (Sallis, 2002). By using the principles and methodology of TQM, education institutions can access a set of continuous improvement tools to manage change and deal with external accountability pressures. However, the most important role of management and school administration is to support and empower teaching and learning (Sallis, 2002). Figure 2 depicts the upside-down organizational structure of TQM in the educational setting where students are the focus. School administrators must inspire teachers to engage in effective pedagogical practices that will meet the academic needs of each and every student.

The three main components of TQM in the education setting include meeting the expectations of external and internal stakeholders, emphasizing quality assurance to

specific standards, and highlighting the means for measurement of academic performance (Manaf & Seng, 2010). Sallis (2002) defines three groups of external stakeholders – primary (students), secondary (parents and local community), and tertiary (future employers of students, legislators, and society as a whole). Internal stakeholders represent the employees within the organization (Sallis, 2002). Examples of quality assurance to specific standards include academic standards and rubrics designed to assess various levels of teaching effectiveness. High-stakes testing and value-added measures within teacher evaluation illustrate various types of measures of academic performance.

According to Sallis (2002), there are four quality imperatives of TQM related to the field of education – moral, professional, competitive, and accountability. The moral imperative is that every student receives the best possible education. The professional imperative centers on a commitment to students' needs and an obligation to meeting those needs by implementing appropriate pedagogical practices, a strategy that is in line with developing more effective teachers. The competitive imperative focuses on the needs of the clients (parents and other community stakeholders) and developing strategies to differentiate public schools from their competitors (other public schools/districts, charter schools, private schools, etc.). The accountability imperative involves objective and measurable outcomes of the educational process and provides mechanisms for quality improvement, such as student achievement results linked to value-added scores within the state teacher evaluation system (Sallis, 2002).

Total Quality Management in Educational Research

In 1987, the United States government initiated its own prize based on TQM, titled the Malcolm Baldrige National Quality Award, to encourage American companies

and organizations to strive for excellence and engage in continuous improvement transformations (Baldrige Performance Excellence Program, 2011). In the field of education, school districts like Iredell-Statesville Schools in North Carolina and Jenks Public Schools in Oklahoma have won the Baldrige Award based on seven criteria for performance excellence: leadership, strategic planning, student and market focus, information and analysis, faculty and staff satisfaction, process management in the classroom and business office, and results (Baldrige Performance Excellence Program, 2011; Maurer & Pederson, 2004). Iredell-Statesville Schools boasted the following student achievement gains during their six year continuous improvement plan: 55th to ninth place in student achievement in the state of North Carolina, over 90% proficiency rate in reading, closing the African American achievement gap in reading from 23% to 12%, closing the special education gap in reading from 42% to 21%, improving from 61% to 81% in high school graduation rate, decreasing the dropout rate, and ranking 7th in the state in average SAT scores ((Baldrige Performance Excellence Program, 2011). Other effects included a decrease in teacher turnover and improved faculty and staff satisfaction. Over their six-year continuous improvement transformation process, Jenks Public Schools increased students' proficiency in eighth grade math, enrollment in Advanced Placement courses, passing rates on Advanced Placement exams, the district's graduation rate, and employee enthusiasm for the district (Baldrige Performance Excellence Program, 2011).

In the 1990s, many higher education institutions in the United States adopted TQM philosophies as a "do more with less" response to limited fiscal resources (Mohammad & Aspinwall, 1997). In a case study of 14 higher education institutions,

TQM strategies focused on student satisfaction, employee involvement and teamwork, strategic planning, and analysis of information. Results of these continuous improvement initiatives included increased enrollment, placement rates, employee satisfaction, and student services as well as fewer dropouts, reduced costs, and savings in time and materials (Mohammad & Aspinwall, 1997). However, Koch (2003) was critical of these quality initiatives of higher education organizations in the United States because of their focus on non-academic activities such as registration, physical plant, bursar payments, and purchasing. He argued that academic areas such as faculty tenure and curriculum have been ignored at the higher education level.

Total Quality Management Link to this Study

Proponents of TQM believe performance measurement and quality monitoring in the form of teacher evaluation and student achievement are essential to detecting and eliminating variability in the teaching and learning process. In Oklahoma under the state-wide Teacher/Leader Effectiveness evaluation system, there is a quality assurance piece within the qualitative part and a quality control piece within the quantitative portion. For the majority of the state, including the large suburban school district involved in this study, the qualitative portion involves use of the Tulsa Model of Observation and Evaluation. Administrators use detailed rubrics of teaching effectiveness to observe and evaluate teachers multiple times throughout the school year. Career teachers with over four years of experience are observed twice and evaluated once, whereas probationary teachers with less than four years of experience are observed four times and evaluated twice throughout the school year. In terms of TQM, the goals of the qualitative piece are to emphasize quality assurance to specific levels of teaching, ensure pedagogical

practices are operating at the highest possible standards, allow for feedback on teaching practices, provide some control of the measurement tools for teachers who can supply artifacts demonstrating their teaching effectiveness in particular areas. If teachers are not meeting a certain level of effectiveness, then they are provided with instructional coaching and support in the form of goal setting plans or personal development plans to improve their pedagogy.

On the quantitative side, value-added measures serve as a quality control element to ensure objective and measurable outcomes of student achievement. TQM theorists favor using statistical processes to measure and eliminate variability in the manufacturing process, which is similar to the notion of using value-added models to measure teacher effectiveness (Sallis, 2002). Sallis (2002) purported that using a value-added approach allows teachers and administrators to set realistic performance targets for students based on past data, chart student progress, and take action if a student is under-achieving. He believed value-added models are useful in improving educational quality for students because the difference between a students' performance upon entry and exit of a course are measured and teachers/schools can still do well on value-added measures regardless of the students' abilities and background. "It is not student demographics that make a difference in student achievement, it is the quality and teaching ability of the teacher" (Conyers & Ewy, 2004, p. 103).

Deming, a pioneer in the TQM movement, did not "buy into" the idea of mass inspection with the "aim of finding the bad ones and throwing them out" (Farooq et al., 2007, p. 2). Similar to a mass inspection system, the Tulsa Model for Observation and Evaluation as well as value-added measures are used for high-stakes purposes in

Oklahoma's teacher evaluation system. After an "ineffective" rating in the quantitative portion (value-added score) and qualitative portion (Tulsa Model Rubric) of the evaluation for two consecutive years, a teacher must be dismissed by the school district. In the event that a teacher has a rating of "needs improvement" on the quantitative portion or qualitative portion of the evaluation for three consecutive years, a teacher may be dismissed by the school district (Thompson & Miller, 2015). However, Deming claimed that the managers (school administrators) are responsible for 85% of the defects in products/services (pedagogical practices) (Farooq et al., 2007).

At this time, there is no systematic correlation of the qualitative and quantitative components of the Oklahoma Teacher/Leader evaluation system at the state, district, or school site level. Although a value-added score may identify that a teacher has a significant impact on student achievement, the teacher does not know what specific teaching practices are associated with the student learning gains. In terms of TQM, this study correlates the quality control side of value-added measures with the quality assurance side of effectiveness levels of specific teaching practices on the Tulsa Model of Observation and Evaluation so that continuous improvement efforts can be focused throughout the state of Oklahoma in public school districts and university teacher preparation programs to transform pedagogy and ensure positive influence on student achievement.

Summary

This chapter was a literature review highlighting the educational research on value-added models of teacher effectiveness. It also focused on the theory of total quality management (TQM) as it relates to the field of education and this study in particular.

Chapter III details the research design, describes the participants and instrumentation, and explains the procedures and how data will be analyzed. Assumptions and limitations of the study are also addressed. Chapter IV presents findings from the descriptive and correlational analysis. Chapter V discusses findings through the lens of Total Quality Management. It concludes with implications for practice and further research.

CHAPTER III

METHODOLOGY

This chapter reviews the purpose of the study and the research questions. It also details the research design and procedures as well as the assumptions and limitations of the study.

Purpose of the Study

Oklahoma's newly adopted Teacher/Leader Effectiveness evaluation system provides a unique and unexplored context to examine potential relationships between a teacher's value-added score and specific teaching practices associated with improved student achievement. According to Tulsa Public Schools (n.d.-a), over 90% of Oklahoma public school districts used the Tulsa Model for Observation and Evaluation as of 2012. The Tulsa Model is an "evidence-based process of educator evaluation anchored in specific domains, dimensions, and indicators reflecting national best practices and current research regarding effective instruction" (Tulsa Public Schools, n.d.-b). The overarching domains of the Tulsa Model that reflect best practices of teaching are classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership. Within each domain are specific dimensions of teaching that are defined within a rubric that classifies different levels of

teaching performance as ineffective, in need of improvement, effective, highly effective, and superior (Tulsa Public Schools, 2015). The research behind the Tulsa Model is based on findings from the Northwest Regional Education Lab, Harvard professor Kane and colleagues, Danielson, and Marzano (Tulsa Public Schools, 2015). The purpose of this study is to examine the relationship between the value-added scores for elementary, junior high, and high school teachers of English/language arts, reading, and/or math in a large suburban Oklahoma public school district and administrators' ratings of their specific teaching practices as measured by the Tulsa Model for Observation and Evaluation. By associating specific teaching practices with improved student learning as measured by value-added scores, this study will contribute much needed data to the field and potentially provide a narrower focus for teacher enhancement and improvement that will have the greatest impact on student learning. Both educators in the field as well as pre-service teachers could benefit from this information.

Research Questions

Context

In Oklahoma under the statewide Teacher/Leader Effectiveness evaluation system, school districts measure teacher effectiveness using both value-added scores related to student standardized test scores and ratings of teaching practices from administrator observations of classroom performance (Glisson & White, 2014). In most Oklahoma school districts including the large suburban district in this study, the Tulsa Model for Observation and Evaluation is used to rate teachers under five varying levels of effectiveness within five overarching domains that include 20 specific indicators of teaching practices known to improve student achievement. Those domains include

classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership (Tulsa Public Schools, 2015). Appendix A provides a complete list of teaching practice indicators within each domain of the Tulsa Model for Observation and Evaluation.

Overarching Question

- Q1. Is there a relationship between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores?

Sub-questions

- Q2. Is there a relationship between teachers' ratings on the classroom management domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
- Q3. Is there a relationship between teachers' ratings on the instructional effectiveness domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
- Q4. Is there a relationship between teachers' ratings on the professional growth and continuous improvement domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
- Q5. Is there a relationship between teachers' ratings on the interpersonal skills domain of the Tulsa Model for Observation and Evaluation and their value-added scores?
- Q6. Is there a relationship between teachers' ratings on the leadership domain of the Tulsa Model for Observation and Evaluation and their value-added scores?

Hypotheses

- H_1. There will be no relationship between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores
- H_2. There will be no relationship between teachers' ratings on the classroom management domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_3. There will be no relationship between teachers' ratings on the instructional effectiveness domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_4. There will be no relationship between teachers' ratings on the professional growth and continuous improvement domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_5. There will be no relationship between teachers' ratings on the interpersonal skills domain of the Tulsa Model for Observation and Evaluation and their value-added scores.
- H_6. There will be no relationship between teachers' ratings on the leadership domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

Research Design

This quantitative correlational study was designed to examine the relationship between teachers' value-added scores and ratings of their teaching practices on the Tulsa Model for Observation and Evaluation. Data was provided by a large suburban school

district in the form of spreadsheets complied to fulfill state requirements for reporting Tulsa Model for Observation and Evaluation scores and in the form of individual value-added reports for teachers for the same school years. Any relationship between value-added scores and rating of teacher practices cannot be taken as causal evidence as the data are from a correlational design.

Population

The population are educators from a large suburban school district located in Oklahoma that encompasses 126 square miles and includes 24 elementary schools, five junior highs, and three high schools. Over 1,500 certified teachers attend to more than 23,000 students. The district's ethnicity breakdown is as follows: 69% Caucasian, 14% Hispanic, 6% Black, 6% Native American, and 5% Asian. According to the Office of Educational Quality and Accountability information for this school district, approximately 43% of students qualify for free and reduced lunch, 15% are special education students, 18% are gifted/talented, and 2% are English language learners. Although all certified teachers in the district are evaluated using the Tulsa Model Observation and Evaluation, only teachers of the following courses received value-added scores from the Oklahoma State Department of Education fourth through eighth grades Reading, English/Language Arts, and Math; Algebra I; Algebra II; Geometry; and English III.

Participants

The total sample included 439 educators of the following courses who received value-added scores from the Oklahoma State Department of Education from the 2013-2014 and/or 2014-2015 school years. All teachers who had both value-added scores and

Tulsa Model Observation and Evaluation ratings for one or both school years were included in the study. The statistical analyses for this study were examined by school year to determine if there were statistically significant results for two consecutive years when determining whether the null hypotheses would be accepted or rejected. For the 2013-2014 school year, there were 185 teachers who had both scores – 83 were elementary teachers (fourth through sixth grades) and 102 were secondary teachers (seventh through 12th grades). For the 2014-2015 school year, there were 254 educators who had both scores – 131 were at the elementary level (fourth through sixth grades) and 123 were at the secondary level (seventh through 12th grades).

Instrumentation

The district provided data for both the Tulsa Model for Observation and Evaluation and value-added measures during the 2013-2014 and 2014-2015 school years. These are the only two school years of data that the school district has that include ratings for both the Tulsa Model and value-added scores, which is the reason for inclusion of these two specific school years in this study. The Tulsa Model for Observation and Evaluation ratings was provided in a spreadsheet for each school site as reported to the Oklahoma State Department of Education. Teachers were identified by state identification number only and not by name. Teachers received an overall rating for the entire rubric; five individual domain ratings of classroom management, instructional effectiveness, professional growth and continuous improvement, interpersonal skills, and leadership; and 20 ratings of indicators of specific teaching practices within each domain (Appendix A). Both the value-added scores and Tulsa Model for Observation and Evaluation scores ranged from 1.0 to 5.0 with a score of 3.0 or higher deemed effective.

The higher the score, then the more effective the teacher is determined to be on these measures.

The Oklahoma State Department of Education outsourced to a vendor named Mathematica to compile value-added data for its teachers. Value-added scores were only provided to teachers of reading and math in grades fourth through eighth based on student test results from the Oklahoma Core Curriculum Tests and teachers of algebra I in eighth and ninth grades, geometry in ninth through eleventh grades, algebra II for ninth through twelfth grades, and English III in eleventh grade based on student test results from the End of Instruction state assessments. Mathematica took four steps to estimate a teacher's value-added score: used a multiple regression model accounting for student characteristics of poverty, gender, race/ethnicity, existence of an individualized education plan, limited English proficiency, and student mobility; accounted for measurement error in the pre-test by employing a statistical technique that uses data on the reliability of the state tests provided by the test developers; compared students across grades at the elementary level; and accounted for imprecisely estimated measures based on too few students by reporting value-added estimates for teachers with at least 10 students and using a statistical technique called shrinking that adjusts estimates of teachers with fewer students more towards to overall average (Walsh et al., 2015). The value-added information was provided by the district in the form of individual value-added score reports for teachers. Teachers received an overall value-added score and subject area value-added scores for reading/English/language arts and/or math from the Oklahoma State Department of Education. In the event that a teacher only taught one of the subjects, then the subject area score and overall score were the same. For this study, only the

overall value-added scores for teachers were used.

Research Procedures

Data provided by the district was compiled into a spreadsheet. Teachers were only identified by their state identification number and not by name. To investigate the overarching question, the independent (predictor) variable was a teacher's overall effectiveness score of teaching practices on the Tulsa Model for Observation and Evaluation rubric and the dependent (criterion) variable was the teacher's overall value-added score as measured by the students' performance on specific state assessments. To explore the sub-questions, a correlation analysis between the predictor variable, the teacher's domain score on the Tulsa Model rubric, and the criterion variable, the teacher's overall value-added score, was tested.

In order to determine whether and to what degree a relationship between ratings of teacher practices and value-added scores are correlated, a correlation coefficient was computed. The Pearson Product-Moment Correlation coefficient is a common statistical method used to examine the relationship between quantitative variables. Using SPSS, a bivariate correlation coefficient in the form of a Pearson Product-Moment Correlation was calculated as a way to understand if a relationship existed and further test strength and direction of the relationship. Due to the correlational design of this study, it is important to note that any relationships between value-added scores and rating of teacher practices cannot be taken as causal evidence.

Data Analysis

The Pearson correlation coefficient was used to test a linear relationship between variables. Values range from -1.00, representing a strong negative relationship, to +1.00,

representing a strong positive relationship. Statistical significance was also calculated at an alpha of .05. The relationship between each of the specific domains of the Tulsa Model rubric and the value-added score were tested. Using SPSS, a bivariate correlation coefficient in the form of a Pearson Product Moment Correlation was calculated as a way to express the size and strength of the relationships among variables. Additional correlation coefficients were calculated for each indicator of teaching practices within that domain. For example, within the classroom management domain of the Tulsa Model rubric, additional calculations were conducted on specific practices within that domain – preparation, discipline, building-wide climate, lesson plans, assessment practices, and student relations – to further examine relationships among the variables.

Assumptions of the Study

The following assumptions were made regarding this study:

- Each participant had a score for both variables.
- Both variables were continuous.
- Values for variables across cases were not related.
- Value-added scores were collected, measured, and reported without error.
- Administrators subjectively rated teaching practices according to the rubric language of the Tulsa Model for Observation and Evaluation. Observations across these ratings were reported without error.
- A linear relationship and normal distribution existed among both variables.
- Residuals were uncorrelated and had constant variance.
- An absence of outliers existed.

Limitations of the Study

Several limitations existed in this study. The first involved the generalizability of results. Because data were collected from one suburban school district, results should only be generalized to teachers in this district or other suburban districts with similar characteristics. The second limitation is that this study was correlational and causality cannot be inferred. The third limitation was the subjective nature of the teacher rating scores given by the evaluating administrator on the Tulsa Model rubric.

Summary

This chapter described the research design and procedures as well as the assumptions and limitations of the study. Chapter IV presents findings from the descriptive and correlational analysis. Chapter V discusses findings through the lens of Total Quality Management and concludes with implications for practice and further research.

CHAPTER IV

RESEARCH FINDINGS

This chapter presents findings from the descriptive and correlational analysis for each null hypothesis. Oklahoma's Teacher/Leader Effectiveness evaluation system provides a unique and unexplored context to examine potential relationships between a teacher's value-added score and observed teaching practices rated on the Tulsa Model rubric.

In order to determine whether and to what degree a relationship between ratings of teacher practices and value-added scores are correlated, a bivariate correlation coefficient in the form of a Pearson Product Moment Correlation was calculated using SPSS to understand if a relationship exists and further test strength and direction of the relationship. Statistical significance was calculated at an alpha of .05. Cohen's effect size for Pearson correlation coefficients (r) were used with $0.1 < | r | < 0.3$ considered a small/weak effect size, $0.3 < | r | < 0.5$ indicating a medium/moderate effect size, and $| r | > 0.5$ reflecting a large/strong effect size (Cohen, 1988). The coefficient of determination (r^2) indicates the percentage of variance in the variables. Prior to analysis, SPSS skewness ("measure of the asymmetry of a distribution") and kurtosis ("measure of the extent to which observations cluster around a central point") were conducted on all variables and found to be within the range of normality, which is required for variables included in Pearson's correlation coefficient tests (IBM, 2016b, Distribution section).

Descriptive Statistics and Correlational Analyses

Null Hypothesis 1 for Overall Evaluation Score on Tulsa Model Rubric

H_1. There will be no relationship between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores

For the 2013-2014 school year, there was a total sample of 185 teachers who had overall value-added scores and overall ratings of teaching practice on the Tulsa Model rubric. A Pearson product-moment correlation coefficient was computed to assess the relationship between overall value-added scores and overall ratings of teaching effectiveness on the Tulsa Model rubric. Although there was a statistically significant, positive correlation between the two variables at $r(183) = .213$, $p = .004$, it was a weak effect size with higher overall ratings of teacher practices explaining 4.5% of the variation in higher overall value-added scores. For the 2014-2015 school year, there was a total sample of 254 educators with scores for both variables. There was also a statistically significant, small positive correlation, $r(252) = .160$, $p = .011$, with higher overall ratings of effective teaching practices accounting for 2.5% of the variation in overall value-added scores. The correlation analyses presented in Table 1 suggests that the null hypothesis may be rejected and that small, positive correlations that are statistically significant exist for both years of data for overall value-added scores and overall evaluation scores on the Tulsa Model rubric. Higher overall ratings on the Tulsa Model rubric are associated with higher value-added scores.

Table 1

Pearson Product-Moment Correlation Coefficient Results for Null Hypothesis 1

		2013-2014 Overall Evaluation Scores	2014-2015 Overall Evaluation Scores
Overall Value-Added Scores by Year	Pearson Correlation	.213**	.160*
	Sig. (2 tailed)	.004	.011
	N	185	254

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

These relationships were examined even further to determine if a correlation existed between the same variables for elementary and secondary teachers for these school years. There was a statistically significant, small positive correlation between overall value-added scores and overall evaluation scores for elementary teachers at $r(81) = .291$, $p = .008$ for the 2013-2014 school year. However, for the 2014-2015 school year, there was a weak positive correlation that was not statistically significant for elementary teachers at $r(129) = .152$, $p = .083$. For secondary teachers, there were negligible, positive correlations with no statistical significance for both school years with 2013-2014 at $r(100) = .077$, $p = .444$ and 2014-2015 at $r(121) = .031$, $p = .730$.

Null Hypothesis 2 for Classroom Management Domain

H_2. There will be no relationship between teachers' ratings on the classroom management domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

The first domain within the Tulsa Model rubric is the classroom management area of the evaluation, and this domain includes preparation for teaching, lesson planning, discipline, student relations, assessment practices, and contribution to building-wide

climate (Tulsa Public Schools, 2015). Sample sizes for this hypothesis testing were the same as testing for the first null hypothesis for both school years and included educators who had an overall value-added score and a score for the classroom management domain on the Tulsa Model rubric. For both school years, there was a statistically significant, small positive correlation between the classroom management domain score and the overall value-added score for all teachers with $r(183) = .223$, $p = .002$ for the 2013-2014 school year and $r(252) = .183$, $p = .003$ for the 2014-2015 school year. Higher ratings of teaching practices on the classroom management domain of the Tulsa Model accounted for 4.9% of the variability in 2013-2014 and 3.3% of the variability in 2014-2015 in overall value-added scores. The correlation analyses presented in Table 2 suggests that the null hypothesis may be rejected and that small, positive correlations that are statistically significant exist for both years of data for overall value-added scores and classroom management domain scores on the Tulsa Model rubric. Higher ratings on the classroom management domain of the Tulsa Model rubric are associated with higher value-added scores.

Table 2

Pearson Product-Moment Correlation Coefficient Results for Null Hypothesis 2

		2013-2014 Evaluation Scores for Classroom Management	2014-2015 Evaluation Scores for Classroom Management
Overall Value-Added Scores by Year	Pearson Correlation	.223**	.183**
	Sig. (2 tailed)	.002	.003
	N	185	254

**. Correlation is significant at the 0.01 level (2-tailed)

These correlations were further examined between the same variables for elementary and secondary teachers for these school years. For the 2013-2014 school year,

a statistically significant, small positive correlation between overall value-added scores and classroom management domain scores for elementary teachers was found at $r(81) = .248$, $p = .024$. However, for the 2014-2015 school year, there was a negligible positive correlation that was not statistically significant for elementary teachers at $r(129) = .170$, $p = .052$. For secondary teachers, there was no statistical significance between variables for both school years with 2013-2014 having a small positive correlation at $r(100) = .150$, $p = .132$ and 2014-2015 having a negligible positive correlation at $r(121) = .065$, $p = .477$.

Indicator 1 – Preparation

Within the classroom management domain, the Tulsa Model Rubric has six indicators of effective teaching. The first one is titled Preparation and is centered upon the notion of teachers executing instructional strategies that promote performance, critical thinking, and problem-solving skills; planning for both short- and long-term objectives; having lesson plans in alignment with state standards, curriculum maps, and pacing guides; differentiating instruction to address student diversity; maximizing learning time; and having necessary materials and equipment ready for lessons (Tulsa Public Schools, 2015). For the 2013-2014 school year, a statistically significant, small positive correlation existed between overall value-added scores and ratings on the Preparation indicator of the Tulsa Model rubric for all teachers with $r(183) = .251$, $p = .001$. For elementary teachers that same year, the relationship between these variables was moderate and positive at $r(81) = .308$ and statistically significant with $p = .005$. For secondary teachers, the correlation was weak and positive with no statistical significance at $r(100) = .157$, $p = .116$. During the 2014-2015 school year, the relationship between overall value-added scores and preparation indicator scores for all teachers was statistically significant at $p =$

.018 and indicated a small, positive correlation at r(252) = .148. There were small, positive correlations that were not statistically significant between these variables at the elementary level with r(129) = .134, p = .127 and the secondary level with r(121) = .100, p = .270 for this same school year. Table 3 summarizes the Pearson correlation coefficient data for the Preparation Indicator.

Table 3
Pearson Correlation Coefficient for Preparation Indicator under Classroom Management

		2013-2014 Overall Value- Added Scores	2014-2015 Overall Value- Added Scores
Preparation Indicator Scores for All Teachers by Year	Pearson Correlation	.251**	.148*
	Sig. (2 tailed)	.001	.018
	N	185	254
Preparation Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.308**	.134
	Sig. (2 tailed)	.005	.127
	N	83	131
Preparation Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.157	.100
	Sig. (2 tailed)	.116	.270
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 2 – Discipline

The second indicator under the classroom management domain is titled Discipline and is centered upon the teacher implementing appropriate standards of conduct, monitoring the behavior of all students in a variety of school settings, and stopping misbehavior promptly and consistently as well as students clearly understanding and following the classroom expectations (Tulsa Model, 2015). For the 2013-2014 school year, a statistically significant, small positive correlation existed between overall value-

added scores and ratings on the Discipline indicator of the Tulsa Model rubric with $r(183) = .237$, $p = .001$. For elementary teachers that same year, the relationship between these variables was small and positive at $r(81) = .284$ and statistically significant with $p = .009$. For secondary teachers, the correlation was weak and positive with no statistical significance at $r(100) = .138$, $p = .165$. During the 2014-2015 school year, the relationship between overall value-added scores and preparation indicator scores for all teachers was statistically significant at $p = .002$ and indicated a small, positive correlation at $r(252) = .190$. There was a small, positive correlation that was statistically significant between these variables at the elementary level with $r(129) = .181$, $p = .038$. However, the secondary level indicated a negligible, positive relationship with no statistical significance with $r(121) = .057$, $p = .534$ for this same school year. Table 4 summarizes the Pearson correlation coefficient data for the Discipline Indicator.

Table 4

Pearson Correlation Coefficient for Discipline Indicator under Classroom Management

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Discipline Indicator Scores for All Teachers by Year	Pearson Correlation	.237**	.190**
	Sig. (2 tailed)	.001	.002
	N	185	254
Discipline Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.284**	.181*
	Sig. (2 tailed)	.009	.038
	N	83	131
Discipline Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.138	.057
	Sig. (2 tailed)	.165	.534
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 3 – Building-wide Climate Responsibilities

Within the classroom management domain, the third indicator is titled Building-wide Climate Responsibilities and includes the teacher regularly participating in school initiatives that promote orderly behavior of students and following procedures and guidelines to keep students healthy and safe (Tulsa Model, 2015). For the 2013-2014 school year, a statistically significant, small positive correlation existed between overall value-added scores and ratings on the Building-wide Climate indicator scores of the Tulsa Model rubric for all teachers with $r(183) = .147$, $p = .046$. The relationship between these same variables for elementary teachers was not statistically significant with $p = .212$ but was small and positive at $r(81) = .138$. For secondary teachers, the correlation was negligible yet positive with no statistical significance at $r(100) = .138$, $p = .165$. During the 2014-2015 school year, the relationship between overall value-added scores

and Building-wide Climate indicator scores for all teachers was not statistically significant at p = .093 and indicated a small, positive correlation at r(252) = .106. There was a small, positive correlation that was not statistically significant between these variables at the elementary level with r(129) = .115, p = .191. At the secondary level, there was a negative, negligible relationship with r(121) = -.037 that was not statistically significant at p = .688 for this same school year. Table 5 summarizes the Pearson correlation coefficient data for the Building-wide Climate Indicator.

Table 5
Pearson Correlation Coefficient for Building-wide Climate Indicator under Classroom Management

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Building-wide Climate Indicator Scores for All Teachers by Year	Pearson Correlation	.147*	.106
	Sig. (2 tailed)	.046	.093
	N	185	254
Building-wide Climate Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.138	.115
	Sig. (2 tailed)	.212	.191
	N	83	131
Building-wide Climate Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.099	-.037
	Sig. (2 tailed)	.322	.688
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

Indicator 4 – Lesson Plans

The fourth indicator under the Classroom Management domain is titled Lesson Plans and is focused on the teacher planning with other members of the grade-level/content-area team, consistently developing lesson plans based upon an analysis of

student data, and providing adequate substitute plans (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was not statistically significant existed between overall value-added scores and ratings on the Lesson Plans indicator of the Tulsa Model rubric for all teachers with $r(183) = .102$, $p = .169$. For that same year, the relationship between these variables was negligible, yet positive with no statistical significance for both elementary teachers at $r(81) = .097$, $p = .383$ and secondary teachers at $r(100) = .093$, $p = .354$. During the 2014-2015 school year, the relationship between overall value-added scores and the Lesson Plans indicator scores for all teachers was not statistically significant at $p = .061$ and indicated a small, positive correlation at $r(252) = .118$. There was a weak, positive correlation that was not statistically significant between these variables at the elementary level with $r(129) = .115$, $p = .189$. The secondary level indicated a negligible, positive relationship with no statistical significance with $r(121) = .067$, $p = .462$ for this same school year. Table 6 summarizes the Pearson correlation coefficient data for the Lesson Plans Indicator.

Table 6
Pearson Correlation Coefficient for Lesson Plans Indicator under Classroom Management

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Lesson Plans Indicator Scores for All Teachers by Year	Pearson Correlation	.102	.118
	Sig. (2 tailed)	.169	.061
	N	185	254
Lesson Plans Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.097	.115
	Sig. (2 tailed)	.383	.189
	N	83	131
Lesson Plans Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.093	.067
	Sig. (2 tailed)	.354	.462
	N	102	123

Indicator 5 – Assessment Practices

Within the classroom management domain, the fifth indicator is titled Assessment Practices and includes the teacher using assessments to evaluate learning and support differentiated instruction, having fair and transparent grading practices consistent with district policies, providing students with pertinent and immediate feedback, and acknowledging student progress at significant intervals to encourage success (Tulsa Model, 2015). For the 2013-2014 school year, a small positive correlation with no statistical significance existed between overall value-added scores and ratings on the Assessment Practices indicator scores of the Tulsa Model rubric for all teachers with $r(183) = .076$, $p = .302$. The relationship between these same variables for elementary teachers was not statistically significant with $p = .331$ but was small and positive at $r(81) = .108$. For secondary teachers, the correlation was negligible yet positive with no statistical significance at $r(100) = .004$, $p = .971$. During the 2014-2015 school year, the

relationship between overall value-added scores and Assessment Practices indicator scores for all teachers was statistically significant at p = .005 and indicated a small, positive correlation at r(252) = .177. There was a small, positive correlation that was not statistically significant between these variables at the elementary level with r(129) = .128, p = .144. At the secondary level, there was a small, positive relationship with r(121) = .132 of no statistical significance at p = .145 for this same school year. Table 7 summarizes the Pearson correlation coefficient data for the Assessment Practices Indicator.

Table 7
Pearson Correlation Coefficient for Assessment Practices Indicator under Classroom Management

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Assessment Practices Indicator Scores for All Teachers by Year	Pearson Correlation	.076	.177**
	Sig. (2 tailed)	.302	.005
	N	185	254
Assessment Practices Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.108	.128
	Sig. (2 tailed)	.331	.144
	N	83	131
Assessment Practices Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.004	.132
	Sig. (2 tailed)	.971	.145
	N	102	123

**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 6 – Student Relations

The sixth indicator under the classroom management domain is titled Student Relations and is centered upon the teacher engaging in considerate and respectful

communications with students, demonstrating high expectations for students, and conveying a positive view of learning (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was not statistically significant existed between overall value-added scores and ratings on the Student Relations indicator of the Tulsa Model rubric for all teachers with $r(183) = .134$, $p = .068$. For that same year, the relationship between these variables was small and positive with no statistical significance for both elementary teachers at $r(81) = .145$, $p = .191$ and secondary teachers at $r(100) = .121$, $p = .226$. During the 2014-2015 school year, the relationship between overall value-added scores and the Student Relations indicator scores for all teachers was not statistically significant at $p = .079$ and indicated a small, positive correlation at $r(252) = .110$. There was a weak, positive correlation that was not statistically significant between these variables at the elementary level with $r(129) = .122$, $p = .164$. The secondary level indicated a negligible, negative relationship with no statistical significance with $r(121) = -.008$, $p = .928$ for this same school year. Table 8 summarizes the Pearson correlation coefficient data for the Student Relations Indicator.

Table 8
Pearson Correlation Coefficient for Student Relations Indicator under Classroom Management

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Student Relations Indicator Scores for All Teachers by Year	Pearson Correlation	.134	.110
	Sig. (2 tailed)	.068	.079
	N	185	254
Student Relations Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.145	.122
	Sig. (2 tailed)	.191	.164
	N	83	131
Student Relations Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.121	-.008
	Sig. (2 tailed)	.226	.928
	N	102	123

Null Hypothesis 3 for Instructional Effectiveness Domain

H_3. There will be no relationship between teachers' ratings on the instructional effectiveness domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

On the Tulsa Model rubric, the second domain is the instructional effectiveness area of the evaluation and includes embedding literacy components, incorporating current state standards, involving all learners, explaining content, giving clear instructions and directions, modeling, monitoring, making adjustments based on monitoring, establishing closure, and acknowledging student achievement (Tulsa Public Schools, 2015). Sample sizes for this hypothesis testing were the same as testing for the first two null hypothesis and included educators who had an overall value-added score and a score for the classroom management domain on the Tulsa Model rubric. For both school years, there

was a statistically significant, small positive correlation between the instructional effectiveness domain score and the overall value-added score for all teachers with r(183) = .195, p = .008 for the 2013-2014 school year and r(252) = .145, p = .021 for the 2014-2015 school year. Higher ratings of teaching practices on the instructional effectiveness domain of the Tulsa Model accounted for 3.8% of the variability in 2013-2014 and 2.1% of the variability in 2014-2015 in overall value-added scores. The correlation analyses presented in Table 9 suggests that the null hypothesis may be rejected and that small, positive correlations that are statistically significant exist for both years of data for overall value-added scores and the instructional effectiveness domain scores on the Tulsa Model rubric. Higher ratings on the instructional effectiveness domain of the Tulsa Model rubric are associated with higher value-added scores.

Table 9

Pearson Product-Moment Correlation Coefficient Results for Null Hypothesis 3

		2013-2014 Evaluation Scores for Instructional Effectiveness	2014-2015 Evaluation Scores for Instructional Effectiveness
Overall Value-Added Scores by Year	Pearson Correlation	.195**	.145*
	Sig. (2 tailed)	.008	.021
	N	185	254

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

These relationships were examined even further between the same variables for elementary and secondary teachers for these school years. There was a statistically significant, small positive correlation between overall value-added scores and overall instructional effectiveness domain scores for elementary teachers at r(81) = .299, p = .006 for the 2013-2014 school year. However, for the 2014-2015 school year, there was a

weak positive correlation that was not statistically significant for elementary teachers at r(129) = .126, p = .310. For secondary teachers, there were negligible, positive correlations with no statistical significance for both school years with 2013-2014 at r(100) = .030, p = .765 and 2014-2015 at r(121) = .042, p = .643.

Indicator 7 – Literacy

Within the instructional effectiveness domain, the Tulsa Model Rubric has ten indicators of effective teaching. The Literacy indicator is centered on embedding the components of literacy into instruction in all content areas (reading, writing, developing vocabulary, spelling, and/or listening/speaking) and providing instruction through text (Tulsa Public Schools, 2015). For the 2013-2014 school year, a negligible, negative correlation that was not statistically significant existed between overall value-added scores and ratings on the Literacy indicator of the Tulsa Model rubric for all teachers with r(183) = -.017, p = .815. For elementary teachers that same year, the relationship between these variables was weak and positive at r(81) = .190 and not statistically significant with p = .085. For secondary teachers, the correlation was small and negative with statistical significance at r(100) = -.235, p = .018. During the 2014-2015 school year, the relationship between overall value-added scores and Literacy indicator scores for all teachers was not statistically significant at p = .827 and indicated a negligible, positive correlation at r(251) = .014. For the elementary level, there was a small, positive correlations that was not statistically significant between these variables at the with r(128) = .144, p = .102. At the secondary level, a small, negative relationship with statistical significance existed with r(121) = -.272, p = .002 for this same school year. Table 10 summarizes the Pearson correlation coefficient data for the Literacy Indicator.

Table 10

Pearson Correlation Coefficient for Literacy Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Literacy Indicator Scores for All Teachers by Year	Pearson Correlation	-.017	.014
	Sig. (2 tailed)	.815	.827
	N	185	253
Literacy Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.190	.144
	Sig. (2 tailed)	.085	.102
	N	83	130
Literacy Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-.235*	-.272**
	Sig. (2 tailed)	.018	.002
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 8 – Current State Standards

The Current State Standards indicator under the instructional effectiveness domain focuses on the teacher understanding the current state standards and implementing alternate instructional strategies to align content with those state standards (Tulsa Model, 2015). For the 2013-2014 school year, a negligible, positive correlation that was not statistically significant existed between overall value-added scores and ratings on the Current State Standards indicator of the Tulsa Model rubric for all teachers with $r(183) = .080$, $p = .279$. For that same year, the relationship between these variables was small and positive with statistical significance for elementary teachers at $r(81) = .217$, $p = .049$, but the correlation for secondary teachers was not statistically significant and negligible yet positive at $r(100) = .011$, $p = .911$. During the 2014-2015 school year, the relationship between overall value-added scores and the Current State Standards

indicator scores for all teachers was not statistically significant at p = .068 and indicated a small, positive correlation at r(251) = .115. There was a negligible, positive correlation that was not statistically significant between these variables at the elementary level with r(129) = .037, p = .679. The secondary level indicated a small, positive relationship with no statistical significance with r(120) = .148, p = .105 for this same school year. Table 11 summarizes the Pearson correlation coefficient data for the Current State Standards Indicator.

Table 11
Pearson Correlation Coefficient for Current State Standards Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Current State Standards Indicator Scores for All Teachers by Year	Pearson Correlation	.080	.115
	Sig. (2 tailed)	.279	.068
	N	185	253
Current State Standards Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.217*	.037
	Sig. (2 tailed)	.049	.679
	N	83	131
Current State Standards Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.011	.148
	Sig. (2 tailed)	.911	.105
	N	102	122

*. Correlation is significant at the 0.05 level (2-tailed).

Indicator 9 – Involves All Learners

The Involves All Learners indicator within the instructional effectiveness domain is centered on the educator using strategies to engage all students in active learning at least 80% of the class time, using questioning techniques that scaffold to the middle

levels of Bloom's Taxonomy, providing adequate wait time for student response, and incorporating students' skills and interests into lessons (Tulsa Public Schools, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed between overall value-added scores and ratings on the Involves All Learners indicator of the Tulsa Model rubric for all teachers with $r(183) = .226$, $p = .002$. For elementary teachers that same year, the relationship between these variables was moderate and positive at $r(81) = .358$ and statistically significant with $p = .001$. For secondary teachers, the correlation was small and positive with no statistical significance at $r(100) = .048$, $p = .633$. During the 2014-2015 school year, the relationship between overall value-added scores and Involves All Learners indicator scores for all teachers was not statistically significant at $p = .151$ and indicated a negligible, positive correlation at $r(252) = .090$. For this same year, there was a negligible, positive relationship that was not statistically significant between these variables with $r(129) = .052$, $p = .553$ for elementary educators and $r(121) = .046$, $p = .611$ for secondary teachers. Table 12 summarizes the Pearson correlation coefficient data for the Involves All Learners Indicator.

Table 12
Pearson Correlation Coefficient for Involves All Learners Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Involves All Learners Indicator Scores for All Teachers by Year	Pearson Correlation	.226**	.090
	Sig. (2 tailed)	.002	.151
	N	185	254
Involves All Learners Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.358**	.052
	Sig. (2 tailed)	.001	.553
	N	83	131
Involves All Learners Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.048	.046
	Sig. (2 tailed)	.633	.611
	N	102	123

**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 10 – Explains Content

The Explains Content indicator under the instructional effectiveness domain focuses on the teacher implementing a variety of activities to support outcomes and meet students' needs while also incorporating technology as an instructional tool on a consistent basis (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed between overall value-added scores and ratings on the Explains Content indicator of the Tulsa Model rubric for all teachers with $r(183) = .159$, $p = .030$. For that same year, the relationship between these variables was small and positive with no statistical significance for elementary teachers at $r(81) = .192$, $p = .081$; and the correlation for secondary teachers was not statistically significant and negligible yet positive at $r(100) = .083$, $p = .409$. During the 2014-2015 school year,

the relationship between overall value-added scores and the Explains Content indicator scores for all teachers was not statistically significant at p = .076 and indicated a small, positive correlation at r(252) = .112. There were negligible, positive correlations that were not statistically significant between these variables at the elementary level with r(129) = .080, p = .366 and secondary level with r(121) = .047, p = .607 for this same school year. Table 13 summarizes the Pearson correlation coefficient data for the Explains Content Indicator.

Table 13
Pearson Correlation Coefficient for Explains Content Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Explains Content Indicator Scores for All Teachers by Year	Pearson Correlation	.159*	.112
	Sig. (2 tailed)	.030	.076
	N	185	254
Explains Content Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.192	.080
	Sig. (2 tailed)	.081	.366
	N	83	131
Explains Content Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.083	.047
	Sig. (2 tailed)	.409	.607
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

Indicator 11 – Clear Instruction and Directions

The Clear Instruction and Directions indicator within the instructional effectiveness domain is centered on the educator providing instructions in a variety of modes, giving students directions for transitions while optimizing academic learning

time, and using appropriate spoken and written language for students' ages and interests (Tulsa Public Schools, 2015). For the 2013-2014 school year, a small, positive correlation that was not statistically significant existed for all teachers between overall value-added scores and ratings on the Clear Instruction and Directions indicator of the Tulsa Model rubric with $r(183) = .135$, $p = .066$. For this same year, the relationship between these variables was weak, positive and not statistically significant at $r(81) = .213$, $p = .053$ for the elementary level and $r(100) = .012$, $p = .904$ for the secondary level. During the 2014-2015 school year, the relationship between overall value-added scores and Clear Instruction and Directions indicator scores for all teachers was not statistically significant at $p = .055$ and indicated a small, positive correlation at $r(252) = .120$. For this same year, there was a negligible, positive relationship that was not statistically significant between these variables with $r(129) = .079$, $p = .373$ for elementary educators and $r(121) = .066$, $p = .467$ for secondary teachers. Table 14 summarizes the Pearson correlation coefficient data for the Clear Instruction and Directions Indicator.

Table 14
Pearson Correlation Coefficient for Clear Instruction and Directions Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Clear Instruction and Directions Indicator Scores for All Teachers by Year	Pearson Correlation	.135	.120
	Sig. (2 tailed)	.066	.055
	N	185	254
Clear Instruction and Directions Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.213	.079
	Sig. (2 tailed)	.053	.373
	N	83	131
Clear Instruction and Directions Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.012	.066
	Sig. (2 tailed)	.904	.467
	N	102	123

Indicator 12 – Models

The Models indicator under the instructional effectiveness domain focuses on the teacher providing demonstrations and modeling of the desired skill or process that are clear (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed for all teachers between overall value-added scores and ratings on the Models indicator of the Tulsa Model rubric with $r(183) = .251$, $p = .001$. For that same year, the relationship between these variables was moderate and positive with statistical significance for elementary teachers at $r(81) = .329$, $p = .002$. However, the correlation for secondary educators was not statistically significant and weak yet positive at $r(100) = .167$, $p = .093$. During the 2014-2015 school year, the relationship between overall value-added scores and the Models indicator scores for all

teachers was statistically significant at p = .005 and indicated a small, positive correlation at r(252) = .177. There was a weak, positive correlations that was not statistically significant between these variables at the elementary level with r(129) = .133, p = .130. The secondary level had a statistically significant, small, positive relationship with r(121) = .207, p = .022 for this same school year. Table 15 summarizes the Pearson correlation coefficient data for the Models Indicator.

Table 15
Pearson Correlation Coefficient for Models Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Models Indicator Scores for All Teachers by Year	Pearson Correlation	.251**	.177**
	Sig. (2 tailed)	.001	.005
	N	185	254
Models Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.329**	.133
	Sig. (2 tailed)	.002	.130
	N	83	131
Models Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.167	.207*
	Sig. (2 tailed)	.093	.022
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 13 – Monitors

The Monitors indicator within the instructional effectiveness domain is centered on the educator moving throughout the classroom while students are working to reinforce progress, utilizing various types of response techniques and student feedback to check for understanding, using appropriate wait time after asking a question, and providing opportunities to develop more thorough responses when presented with additional

information (Tulsa Public Schools, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed for all teachers between overall value-added scores and ratings on the Monitors indicator of the Tulsa Model rubric with $r(183) = .188$, $p = .010$. For this same year, the relationship between these variables was weak, positive, and statistically significant at $r(81) = .234$, $p = .033$ for elementary teachers. It was small and positive for secondary teachers, but not statistically significant at $r(100) = .103$, $p = .305$. During the 2014-2015 school year, the relationship between overall value-added scores and Monitors indicator scores for all teachers was statistically significant at $p = .027$ and indicated a small, positive correlation at $r(252) = .138$. For this same year, there was a negligible, positive relationship that was not statistically significant between these variables with $r(129) = .094$, $p = .288$ for elementary educators; and a weak, positive correlation that was not statistically significant $r(121) = .103$, $p = .259$ for secondary teachers. Table 16 summarizes the Pearson correlation coefficient data for the Monitors Indicator.

Table 16

Pearson Correlation Coefficient for Monitors Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Monitors Indicator Scores for All Teachers by Year	Pearson Correlation	.188*	.138*
	Sig. (2 tailed)	.010	.027
	N	185	254
Monitors Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.234*	.094
	Sig. (2 tailed)	.033	.288
	N	83	131
Monitors Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.103	.103
	Sig. (2 tailed)	.305	.259
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

Indicator 14 – Adjusts Based upon Monitoring

The Adjusts Based upon Monitoring indicator under the instructional effectiveness domain focuses on the teacher adjusting education plans to engage more students, assessing mastery of new learning to see if reteaching is necessary, and reviewing student data to adapt instruction and guide intervention (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed for all teachers between overall value-added scores and ratings on the Adjusts Based upon Monitoring indicator of the Tulsa Model rubric with $r(182) = .173$, $p = .019$. For that same year, the relationship between these variables was small and positive with statistical significance for elementary teachers at $r(80) = .295$, $p = .007$. However, the correlation for secondary teachers was not statistically significant and weak yet positive at $r(100) = .026$, $p = .798$. During the 2014-2015 school year, the relationship between overall value-added scores and the Adjusts Based upon Monitoring indicator

scores for all teachers was statistically significant at p = .017 and indicated a small, positive correlation at r(252) = .150. There was a weak, positive correlations that was not statistically significant between these variables at the elementary level with r(129) = .143, p = .104. The secondary level had a negligible, positive relationship of no statistical significance with r(121) = .053, p = .558 for this same school year. Table 17 summarizes the Pearson correlation coefficient data for the Adjusts Based upon Monitoring Indicator.

Table 17
Pearson Correlation Coefficient for Adjusts Based upon Monitoring Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Adjusts Based upon Monitoring Indicator Scores for All Teachers by Year	Pearson Correlation	.173*	.150*
	Sig. (2 tailed)	.019	.017
	N	184	254
Adjusts Based upon Monitoring Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.295**	.143
	Sig. (2 tailed)	.007	.104
	N	82	131
Adjusts Based upon Monitoring Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.026	.053
	Sig. (2 tailed)	.798	.558
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Indicator 15 – Establishes Closure

The Establishes Closure indicator within the instructional effectiveness domain is centered on the educator using closure strategies to help students reflect in a variety of ways and relate current instruction to prior and future learning (Tulsa Public Schools,

2015). For the 2013-2014 school year, a small, positive correlation that was not statistically significant existed for all teachers between overall value-added scores and ratings on the Establishes Closure indicator of the Tulsa Model rubric with $r(183) = .099$, $p = .182$. For that same year, the relationship between these variables was weak, positive, and not statistically significant at $r(81) = .158$, $p = .154$ for elementary teachers. It was negligible and negative for secondary teachers but not statistically significant at $r(100) = -.024$, $p = .810$. During the 2014-2015 school year, the relationship between overall value-added scores and Establishes Closure indicator scores for all teachers was statistically significant at $p = .022$ and indicated a small, positive correlation at $r(250) = .145$. For this same year, there was a weak, positive relationship that was not statistically significant between these variables with $r(129) = .105$, $p = .231$ for elementary educators and $r(119) = .109$, $p = .234$ for secondary teachers. Table 18 summarizes the Pearson correlation coefficient data for the Establishes Closure Indicator.

Table 18
Pearson Correlation Coefficient for Establishes Closure Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Establishes Closure Indicator Scores for All Teachers by Year	Pearson Correlation	.099	.145*
	Sig. (2 tailed)	.182	.022
	N	185	252
Establishes Closure Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.158	.105
	Sig. (2 tailed)	.154	.231
	N	83	131
Establishes Closure Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-0.24	.109
	Sig. (2 tailed)	.810	.234
	N	102	121

*. Correlation is significant at the 0.05 level (2-tailed).

Indicator 16 – Student Achievement

The Student Achievement indicator under the Instructional Effectiveness domain focuses on the teacher being responsible for the success of all students, modifying assessments for special education students, providing feedback to students and parents, and assuring all students have access to district curriculum and state standards (Tulsa Model, 2015). For the 2013-2014 school year, a negligible, positive correlation that was not statistically significant existed for all teachers between overall value-added scores and ratings on the Student Achievement indicator of the Tulsa Model rubric with $r(183) = .047$, $p = .527$. For that same year, the relationship between these variables was negligible and negative without statistical significance for elementary teachers at $r(81) = -.017$, $p = .877$. The correlation for secondary teachers was not statistically significant and negligible yet positive at $r(100) = .055$, $p = .583$. During the 2014-2015 school year, the

relationship between overall value-added scores and the Student Achievement indicator scores for all teachers was not statistically significant at p = .850 and indicated a negligible, positive correlation at r(252) = .012. There was a negligible, positive correlation that was not statistically significant between these variables at the elementary level with r(129) = .070, p = .430. The secondary level had a negligible, negative relationship of no statistical significance with r(121) = -.128, p = .159 for this same school year. Table 19 summarizes the Pearson correlation coefficient data for the Student Achievement Indicator.

Table 19
Pearson Correlation Coefficient for Student Achievement Indicator under Instructional Effectiveness

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Student Achievement Indicator Scores for All Teachers by Year	Pearson Correlation	.047	.012
	Sig. (2 tailed)	.527	.850
	N	185	254
Student Achievement Indicator Scores for Elem. Teachers by Year	Pearson Correlation	-.017	.070
	Sig. (2 tailed)	.877	.430
	N	83	131
Student Achievement Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.055	-.128
	Sig. (2 tailed)	.583	.159
	N	102	123

Null Hypothesis 4 for Professional Growth Domain

H_4. There will be no relationship between teachers' ratings on the professional growth and continuous improvement domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

On the Tulsa Model rubric, the third domain is the professional growth and continuous improvement area of the evaluation and includes professional development and accountability (Tulsa Public Schools, 2015). Sample sizes for this hypothesis testing were the same as testing for the other null hypothesis and included educators who had an overall value-added score and a score for the professional growth domain on the Tulsa Model rubric. There was a statistically significant, small positive correlation for all teachers between the professional growth domain score and the overall value-added score with $r(183) = .157$, $p = .032$ for the 2013-2014 school year. The 2014-2015 school year data revealed a negligible, positive relationship that was not statistically significant at $r(252) = .083$, $p = .189$. The correlation analyses presented in Table 20 suggests the null hypothesis cannot be rejected since there is insufficient evidence from one of the two school years to conclude that a linear relationship exists between overall value-added scores and the professional growth domain scores on the Tulsa Model rubric for both data sets.

Table 20

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 4

		2013-2014 Evaluation Scores for Professional Growth	2014-2015 Evaluation Scores for Professional Growth
Overall Value-Added Scores by Year	Pearson Correlation	.157*	.083
	Sig. (2 tailed)	.032	.189
	N	185	254

*. Correlation is significant at the 0.05 level (2-tailed).

These relationships were examined even further between the same variables for elementary and secondary teachers for these school years. There were small but positive

correlations that were not statistically significant between overall value-added scores and overall instructional effectiveness domain scores for elementary teachers at $r(81) = .162$, $p = .143$ for the 2013-2014 school year and at $r(129) = .089$, $p = .310$ for the 2014-2015 school year. For secondary teachers, there was a negligible, positive relationship with no statistical significance for the 2013-2014 school year at $r(100) = .087$, $p = .386$. For the 2014-2015 school year, there was a negligible, negative correlation with no statistical significance for secondary teachers at $r(121) = -.038$, $p = .673$.

Indicator 17 – Professional Development

Within the professional growth and continuous improvement domain, the Tulsa Model Rubric has two indicators of effective teaching. The Professional Development indicator is centered on the educator's participation in professional development to update his/her content and pedagogical knowledge and his/her contribution to the profession as an instructional coach, peer mentor, and trainer (Tulsa Public Schools, 2015). For all teachers during the 2013-2014 school year, a negligible, positive correlation that was not statistically significant existed between overall value-added scores and ratings on the Professional Development indicator of the Tulsa Model rubric with $r(183) = .038$, $p = .605$. For elementary teachers that same year, the relationship between these variables was negligible and negative at $r(81) = -.015$ and not statistically significant with $p = .890$. For secondary teachers, the correlation was negligible yet positive with no statistical significance at $r(100) = .032$, $p = .750$. During the 2014-2015 school year, the relationship between overall value-added scores and Professional Development indicator scores for all teachers was not statistically significant at $p = .331$ and indicated a negligible, positive correlation at $r(252) = .331$. For the elementary level,

there was a small, positive correlations that was not statistically significant between these variables with r(129) = .126, p = .153. At the secondary level, a negligible, negative relationship without statistical significance existed with r(121) = -.077, p = .400 for this same school year. Table 21 summarizes the Pearson correlation coefficient data for the Professional Development Indicator.

Table 21
Pearson Correlation Coefficient for Professional Development Indicator under Professional Growth

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Professional Development Indicator Scores for All Teachers by Year	Pearson Correlation	.038	.061
	Sig. (2 tailed)	.605	.331
	N	185	254
Professional Development Indicator Scores for Elem. Teachers by Year	Pearson Correlation	-.015	.126
	Sig. (2 tailed)	.890	.153
	N	83	131
Professional Development Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.032	-.077
	Sig. (2 tailed)	.750	.400
	N	102	123

Indicator 18 – Professional Accountability

The Professional Accountability indicator under the professional growth and continuous improvement domain focuses on the teacher being punctual, dependable, and complying with reporting timelines (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was statistically significant existed for all teachers between overall value-added scores and ratings on the Professional Accountability

indicator of the Tulsa Model rubric with r(183) = .220, p = .003. For that same year, the relationship between these variables was small and positive with statistical significance for elementary teachers at r(81) = .280, p = .010 and secondary teachers at r(100) = .106, p = .287. During the 2014-2015 school year, the relationship between overall value-added scores and the Professional Accountability indicator scores for all teachers was not statistically significant at p = .224 and indicated a negligible, positive correlation at r(251) = .077. There were also negligible, positive correlations that were not statistically significant between these variables at the elementary level with r(128) = .018, p = .843 and at secondary level with r(121) = .009, p = .925 for this same school year. Table 22 summarizes the Pearson correlation coefficient data for the Professional Accountability Indicator.

Table 22
Pearson Correlation Coefficient for Professional Accountability Indicator under Professional Growth

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Professional Accountability Indicator Scores for All Teachers by Year	Pearson Correlation	.220**	.077
	Sig. (2 tailed)	.003	.224
	N	185	253
Professional Accountability Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.280**	.018
	Sig. (2 tailed)	.010	.843
	N	83	130
Professional Accountability Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.106	.009
	Sig. (2 tailed)	.287	.925
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Null Hypothesis 5 for Interpersonal Skills Domain

H_5. There will be no relationship between teachers' ratings on the interpersonal skills domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

On the Tulsa Model rubric, the fourth domain is the interpersonal skills area of the evaluation and includes effective interpersonal skills in communicating with students' families and colleagues (Tulsa Public Schools, 2015). Sample sizes for this hypothesis testing were the same as testing for the other null hypothesis and included educators who had an overall value-added score and a score for the interpersonal skills domain on the Tulsa Model rubric. There was a negligible, positive correlation that was not statistically significant for all teachers between the interpersonal skills domain score and the overall value-added score with $r(183) = .079$, $p = .283$ for the 2013-2014 school year. The 2014-2015 school year data revealed a small, positive relationship that was statistically significant at $r(252) = .131$, $p = .037$. The correlation analyses presented in Table 23 suggests the null hypothesis cannot be rejected since there is insufficient evidence from one of the two school years to conclude that a linear relationship exists between overall value-added scores and the interpersonal skills domain scores on the Tulsa Model rubric for both data sets.

Table 23

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 5

		2013-2014 Evaluation Scores for Interpersonal Skills	2014-2015 Evaluation Scores for Interpersonal Skills
Overall Value-Added Scores by Year	Pearson Correlation	.079	.131*
	Sig. (2 tailed)	.283	.037
	N	185	254

*. Correlation is significant at the 0.05 level (2-tailed).

These relationships were examined even further between the same variables for elementary and secondary teachers for these school years. There were small but positive correlations that were not statistically significant between overall value-added scores and overall interpersonal skills domain scores for elementary teachers at $r(81) = .167$, $p = .131$ for the 2013-2014 school year and $r(129) = .125$, $p = .155$ for the 2014-2015 school year. For secondary teachers, there was a negligible, negative relationship with no statistical significance for the 2013-2014 school year at $r(100) = -.059$, $p = .554$. For the 2014-2015 school year, there was a negligible, positive correlation with no statistical significance for secondary teachers at $r(121) = .020$, $p = .827$.

Indicator 19 – Effective Interpersonal Skills

Within the interpersonal skills domain, the Tulsa Model Rubric has one indicator of effective teaching. The Effective Interpersonal Skills indicator is focused on the educators' interactions with families and colleagues, compliance with school procedures for communicating with families and peers, and collaborating with families and colleagues to make professional decisions (Tulsa Public Schools, 2015). For all teachers during the 2013-2014 school year, a negligible, positive correlation that was not

statistically significant existed between overall value-added scores and ratings on the Interpersonal Skills indicator of the Tulsa Model rubric with $r(183) = .079$, $p = .283$. For elementary teachers that same year, the relationship between these variables was small and positive at $r(81) = .167$ and not statistically significant with $p = .131$. For secondary teachers, the correlation was negligible and negative without statistical significance at $r(100) = -.059$, $p = .554$. During the 2014-2015 school year, the relationship between overall value-added scores and Interpersonal Skills indicator scores for all teachers was statistically significant at $p = .037$ and indicated a weak, positive correlation at $r(252) = .131$. For the elementary level, there was a small, positive correlations that was not statistically significant between these variables with $r(129) = .125$, $p = .155$. At the secondary level, a negligible, positive relationship without statistical significance existed with $r(121) = .020$, $p = .827$ for this same school year. Table 24 summarizes the Pearson correlation coefficient data for the Interpersonal Indicator.

Table 24
Pearson Correlation Coefficient for Interpersonal Skills Indicator under Interpersonal Skills

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Interpersonal Skills Indicator Scores for All Teachers by Year	Pearson Correlation	.079	.131*
	Sig. (2 tailed)	.283	.037
	N	185	254
Interpersonal Skills Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.167	.125
	Sig. (2 tailed)	.131	.155
	N	83	131
Interpersonal Skills Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-.059	.020
	Sig. (2 tailed)	.554	.827
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

Null Hypothesis 6 for Leadership Domain

H_6. There will be no relationship between teachers' ratings on the leadership domain of the Tulsa Model for Observation and Evaluation and their value-added scores.

On the Tulsa Model rubric, the fifth domain is the leadership area of the evaluation and includes professional involvement and leadership (Tulsa Public Schools, 2015). Sample sizes for this hypothesis testing were the same as testing for the other null hypothesis and included educators who had an overall value-added score and a score for the Leadership domain on the Tulsa Model rubric. There was a small, positive correlation that was not statistically significant for all teachers between the leadership domain score and the overall value-added score with $r(183) = .105$, $p = .154$ for the 2013-2014 school year. The 2014-2015 school year data revealed a negligible, positive relationship that was

not statistically significant at r(252) = .082, p = .194. The correlation analyses presented in Table 25 suggests the null hypothesis can be accepted since there are no statistically significant relationships between overall value-added scores and the leadership domain scores on the Tulsa Model rubric from the data sets for the two school years.

Table 25
Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 6

		2013-2014 Evaluation Scores for Leadership	2014-2015 Evaluation Scores for Leadership
Overall Value-Added Scores by Year	Pearson Correlation	.105	.082
	Sig. (2 tailed)	.154	.194
	N	185	254

These relationships were examined even further through the correlation between the same variables for elementary and secondary teachers for these school years. There was a small but positive correlation that was not statistically significant between overall value-added scores and overall leadership domain scores for elementary teachers at r(81) = .212, p = .054 for the 2013-2014 school year and r(129) = .151, p = .085 for the 2014-2015 school year. For secondary teachers, there was a negligible, positive relationship with no statistical significance for the 2013-2014 school year at r(100) = .018, p = .860. For the 2014-2015 school year, there was a negligible, negative correlation with no statistical significance for secondary teachers at r(121) = -.073, p = .423.

Indicator 20 – Professional Involvement

The Professional Involvement indicator under the leadership domain focuses on the teachers' participation in school/district events, contributions to the profession, and proactive role in addressing students' needs (Tulsa Model, 2015). For the 2013-2014 school year, a small, positive correlation that was not statistically significant existed for

all teachers between overall value-added scores and ratings on the Professional Involvement indicator of the Tulsa Model rubric with $r(183) = .111$, $p = .133$. For that same year, the relationship between these variables was small and positive without statistical significance for elementary teachers at $r(81) = .210$, $p = .057$. For secondary teachers, the correlation was negligible yet positive without statistical significance at $r(100) = .029$, $p = .771$. During the 2014-2015 school year, the relationship between overall value-added scores and the Professional Involvement indicator scores for all teachers was not statistically significant at $p = .194$ and indicated a negligible, positive correlation at $r(252) = .082$. There was a small, positive correlation that was not statistically significant between these variables at the elementary level with $r(129) = .151$, $p = .085$. The secondary level had a negligible, negative relationship of no statistical significance with $r(121) = -.073$, $p = .423$ for this same school year. Table 26 summarizes the Pearson correlation coefficient data for the Student Achievement Indicator.

Table 26
Pearson Correlation Coefficient for Professional Involvement Indicator under Leadership

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Professional Involvement Indicator Scores for All Teachers by Year	Pearson Correlation	.111	.082
	Sig. (2 tailed)	.133	.194
	N	185	254
Professional Involvement Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.210	.151
	Sig. (2 tailed)	.057	.085
	N	83	131
Professional Involvement Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.029	-.073
	Sig. (2 tailed)	.771	.423
	N	102	123

Summary

This chapter presented findings from two consecutive years of data from the descriptive and correlational analysis of value-added scores and various ratings of observed teaching practices from the Tulsa Model rubric. There are two additional tables at the end of this dissertation that completely summarize the findings from this chapter. Table 27 has all findings for the null hypotheses in one table and includes the further correlation analysis of elementary and secondary teachers for both school years. Table 28 contains all findings for the correlation analyses for each Tulsa Model indicator with overall value-added scores for all teachers, elementary teachers, and secondary teachers for both school years.

Chapter V summarizes findings; offers implications for practice, theory, and

future research; and concludes with the study's significance.

CHAPTER V

SUMMARY, IMPLICATIONS, AND CONCLUSIONS

This chapter summarizes findings; offers implications for practice, in relation to the theory of Total Quality Management, and future research; and concludes with the study's significance.

Summary of Findings

Many researchers claim that value-added measures alone are insufficient to determine the impact of specific teaching practices on student achievement. They call for multiple measures of teacher performance including observations supplemented with value-added measures and student outcomes (Betebenner et al., 2012; Darling-Hammond et al., 2012; Goe, 2008). Thus, the purpose of this study was to examine the relationship between value-added scores for elementary and secondary teachers of English/language arts, reading, and/or math in a large suburban Oklahoma public school district with their administrators' ratings of their specific teaching practices as measured by the Tulsa Model for Observation and Evaluation.

To answer this study's overarching research question of determining whether a relationship exists between educators' overall ratings of teaching practices on the Tulsa Model for Observation and Evaluation and their value-added scores, there were small, positive correlations that were statistically significant for all teachers between overall

value-added scores and overall evaluation scores on the Tulsa Model rubric for both school years. This result indicates that higher ratings of effective teaching were slightly associated with higher overall value-added scores and increased student achievement levels for this sample. However, this finding alone does not link any specific teaching practices with the increased student achievement levels. It is very hard to support effective teaching without quality information about actual teaching practices and their link to positive student outcomes (MET, 2013b; Minnici, 2014). The additional research questions aimed to determine the size and strength of any relationships between value-added scores and more specific measures of teaching effectiveness of the Tulsa Model rubric.

For the classroom management domain of the Tulsa Model rubric, both school years revealed small, positive correlations that were statistically significant for all teachers between the scores on that domain and their value-added scores with $r(183) = .223$, $p = .002$ for the 2013-2014 school year and $r(252) = .183$, $p = .003$ for the 2014-2015 school year. This indicates that higher ratings of effective teaching in the area of classroom management were slightly associated with higher overall value-added scores for this sample. When we examined the indicators of effective teaching under the classroom management domain, both the Preparation and Discipline indicators had weak, positive correlation with value-added scores for all teachers for both school years. This indicates that educators who are ranked higher by their evaluating administrators at planning for instructional strategies that encourage critical thinking, problem solving, and performance skills; implementing appropriate standards of conduct; monitoring the behavior of all students in a variety of school settings; and stopping misbehavior

promptly and consistently are somewhat linked with higher value-added scores. None of the other indicators within the classroom management domain had statistically significant relationships with value-added scores for both school years.

For the instructional effectiveness domain, both school years revealed weak, positive correlations that were statistically significant for all teachers between the scores on that domain and their value-added scores with $r(183) = .195$, $p = .008$ for the 2013-2014 school year and $r(252) = .145$, $p = .021$ for the 2014-2015 school year. Within this domain, the scores for all teachers on the Models, Monitors, and Adjusts Based upon Monitoring indicators all had small, positive relationships that were statistically significant with their overall value-added scores. This indicates that educators who are ranked higher by their evaluator on modeling the desired skill or process, checking to determine if students are progressing towards stated objectives, and changing instruction based on the results of monitoring are somewhat linked to higher value-added scores. None of the other indicators within the instructional effectiveness domain had statistically significant relationships with value-added scores for both school years.

There are three other domains within the Tulsa Model rubric – professional growth, interpersonal skills, and leadership. None of these domains or any of their specific indicators of teaching effectiveness had statistically significant relationships with value-added scores. The primary explanation for these results is that they are not student-focused domains like classroom management and instructional effectiveness. They are centered on professional development, communicating with families and colleagues, and being a leader in the profession. It makes sense that the classroom management and instructional effectiveness domains, which are dedicated to the pedagogy and content of

teaching and based on student-teacher interactions, have statistically significant relationships to value-added scores in this study. However, it is important to note that only five out of twenty indicator scores on the Tulsa Model rubric had statistically significant relationships with value-added scores for both school years. The lack of statistically significant relationships between the variables could be contributed to the possibility of principals interpreting the performance of teachers differently despite having the same rubric language for all teachers and the principals all having the same training in order to be qualified to conduct evaluations using the Tulsa Model.

In addition to these main research findings that included all teachers for each school year, additional correlation analyses were tested to determine if there were statistically significant relationships for elementary teachers or secondary educators with the Tulsa Model rubric domains/indicators and value-added scores. For elementary educators, the only statistically significant relationship was the Discipline indicator, which had a small, positive relationship with the teachers' value-added score for both school years. Elementary teachers who are ranked higher by their evaluating administrators on this classroom management indicator are slightly linked with higher value-added scores. For secondary teachers, the Literacy indicator (predictor variable) under the instructional effectiveness domain had a small, but negative relationship with value-added scores that was statistically significant for both school years. This would indicate that junior high and high school teachers that were ranked higher on embedding the components of literacy into instruction and providing instruction through text were associated with lower value-added scores. There were no other statistically significant relationships for both school years between domains or indicators on the Tulsa Model

rubric and value-added scores for elementary or secondary teachers.

Implications for Practice

Overall, the results of this study as a whole do not provide convincing conclusions about which teacher practices could be attributed to differences of teacher effectiveness as measured by value-added scores. One reason for this could be that the Tulsa Model rubric might not be appropriate for discerning differences in teacher practices. As Goe (2007) purported, "When the spread of the teacher's scores on [an] instrument is so constrained, it is very difficult to correlate the scores with student achievement and find meaningful, statistically significant effects" (p. 43).

In line with the findings of the Widget Effect, it might be that the majority of teachers received indicator ratings of "effective" and "highly effective" on the Tulsa Model rubric, which makes it more difficult to discern noticeable differences of effectiveness. Another issue is that statewide standardized tests might not be ideal for measuring the effects of changes to instructional practice. When students are only in a classroom for a little more than half a year before taking the test, important changes a teacher has made to instructional practices for that school year might not be revealed in test scores (Goe, 2007). A third reason is that correlating specific teaching practices to student learning through quantitative models like value-added is impossible when the system cannot completely control for other factors that also affect students' learning such as poverty, peer relationships, school facilities, class sizes, parental involvement, and even the effects of other teachers among other examples (Goe, 2007). A final reason is that a teacher's psychological characteristics in the form of motivation, emotions, self-regulation, and personality, which are believed to be aspects of effective teaching, are not accounted for

on the Tulsa Model rubric or value-added scores ((Bardach, Klassen, & Perry, 2021).

A 2014 report on focus group findings on Oklahoma's teacher evaluation system noted that most districts and schools in the state had not aligned the Tulsa Model rubric with "systemic" professional development (Sothern Regional Education Board, 2014). As this is the only known study to correlative ratings on the Tulsa Model rubric with value-added scores in the state, it would be practical to consider professional development efforts for this district aligned with those indicators that revealed small but statistically significant relationships with improved student achievement on the classroom management and instructional effectiveness domains. The focus should center on the rubric language differentiations for the highly effective and superior categories for the indicators of Preparation, Discipline, Models, Monitors, and Adjusts Based upon Monitoring. Regarding the Preparation indicator, professional development efforts could center on implementing instructional strategies that promote critical thinking and problem solving, developing short- and long-term instructional plans related to state standards and approved curriculum maps, maximizing learning time, and using educational materials and manipulatives that enhance learning (Tulsa Public Schools, 2015). For the Discipline indicator, this involves instilling a sense of self-discipline in students, having standards of conduct that extend beyond the classroom walls, and maintaining the dignity of the student when addressing inappropriate behavior (Tulsa Public Schools, 2015). Regarding the Models indicator, educators should focus on taking preventative action to avoid student misunderstanding and the majority of students should be able to demonstrate the skill/process after it is modeled (Tulsa Public Schools, 2015). For the Monitors indicator, higher levels of teaching effectiveness involve moving to all areas of the room while

students are working to promote progress, using student response techniques to check for understanding, reteaching when a problem is observed, and rephrasing questions to search for detailed student understanding (Tulsa Public Schools, 2015). Concerning the Adjusts Based upon Monitoring indicator, teachers should work with individual students or small groups to reteach, use peer tutoring to accelerate mastery, make adjustments to lesson to engage more students and accommodate their interests, and use ongoing assessments to continually modify instruction (Tulsa Public Schools, 2015). It would also be appropriate to compare other research findings regarding these identified teaching practices in relation to Hattie and Marzano to determine which of these Tulsa Model rubric indicators, if any, might give a teacher more "bang for their buck" in terms of their purported effect on student achievement.

Implications for the Theory of Total Quality Management

The philosophy underpinning Total Quality Management is the need for continuous improvement to achieve organizational excellence (Farooq et al., 2007). The key word here is "organizational." Value-added measures in this study were only provided for certain teachers of specific courses in reading/English/language arts and math due to statewide evaluation system specifications. Out of a population size of 1,500 teachers per year in the selected school district, the sample included 185 teachers for the 2013-2014 school year (roughly 12%) and 254 teachers for the 2014-2015 school year (almost 17%). An organization cannot achieve excellence when the data it is using to improve the organization is only relevant to few in the population. In addition, because value-added measures in Oklahoma are centered on subject areas with state testing components, many content areas like the arts, sciences, and social studies are excluded.

As an organization, this school district does not have any valid data to support continuous improvement initiatives for teachers of these particular subject areas that are not tested.

In terms of TQM, this study correlated the quality control side of value-added measures with the quality assurance side of effectiveness levels of specific teaching practices on the Tulsa Model of Observation and Evaluation. Out of twenty indicators of effective teaching on the Tulsa Model rubric, only 5 of those (25%) were associated with increased value-added scores for all teachers for both school years. However, upon examining the data from an elementary versus secondary lens, discipline was the only one of those 20 indicators to have a statistically significant, positive correlation to overall value-added scores for elementary teachers. This does coincide with findings from Stronge at al. (2011) that top quartile teachers at the elementary level had fewer classroom disruptions, better classroom management, and better relationships with students. At the elementary level in this district, students may have one to three teachers responsible for their instruction per year, whereas at the secondary level, students switch teachers for every instructional period. Classroom management might play a bigger role in creating a climate conducive to learning at the elementary level than the secondary level. Literacy was the only indicator for secondary teachers and it had a statistically significant, negative correlation to overall value-added scores for both school years. This is not supported by any other research findings. Unfortunately, there were not enough significant findings from this study that could help with continuous improvement efforts and lead to organizational excellence for this school district.

Just a few years after adoption in Oklahoma, state legislation reversed the demand for the use of value-added measures in the statewide evaluation system due to the

enormous costs and made it an optional component for school districts (Oklahoma House of Representatives, 2016). This may have been a prudent decision on the part of Oklahoma's lawmakers as this study revealed that correlating teacher effectiveness scores on the Tulsa Model rubric with overall value-added scores did not yield sufficient data to systemically improve the entire teaching force in this school district. In this district and across the state, the Tulsa Model is also a "one size fits all" rubric for the teaching profession and is used for beginning and veteran teachers, regular education, special education, all content areas whether tested or not tested, teachers of all grades PreK through twelfth grade, and even for teachers who have non-instructional responsibilities such those who monitor students who have in school suspension at the secondary level. Continuous improvement efforts for all teachers cannot happen without a differentiated evaluation system that specifically addresses the teaching positions being evaluated with appropriate levels of performance that relate to their unique job descriptions.

Implications for Current and Future Research

Educational researchers have stated that value-added models focus exclusively on student learning. Value-added data is limited on what it can tell educators and administrators about teacher quality and effectiveness and should be supplemented by other sources of evidence. Empirical evidence about which specific teaching practices improve student learning is lacking (Betebenner et al., 2012; Goe, 2008; RAND Corporation, 2004; Stronge et al., 2011). This study did identify higher effectiveness ratings of specific teaching practices as measured by the Tulsa Model rubric that are correlated with higher value-added scores, which will add to the empirical evidence base. It is one of the few studies (if not the only study) that correlates this data in a pre-existing

setting where either the value-added model, teaching effectiveness rubric, or both are not artificially imposed by the researcher. However, the problem still lies in the fact that educational researchers cannot agree on a shared definition of teacher effectiveness and that multiple studies on teacher effectiveness do not yield the same results. As Stronge et al. (2011) stated, "Effective teaching involves a dynamic interplay among content, pedagogical methods, characteristics of learners, and the contexts in which the learning will occur" (p. 349).

There are three suggestions for future research. This study focused on the overall value-added scores for the sample and their relationship to ratings of effective teaching on the Tulsa Model rubric. However, some of the teachers in this sample also have subject area value-added scores for reading and/or math and/or specific courses (Algebra I, Algebra II, Geometry, and English III at the secondary level). It would be interesting to examine correlations for those particular value-added scores with the domains and indicators of the Tulsa Model rubric if the sample size was large enough. Another area of future research involves analyzing the data through a socioeconomic lens. Although these schools are in the same school district, the socioeconomic level varies from school to school. Running these same statistical analyses for Title I schools versus non-Title I schools might produce noteworthy results on specific aspects of teaching and their relation to improved value-added scores for teachers of students from varying socioeconomic levels. A third recommendation for future research would be to look at this data for regular education versus special education teachers. Analyzing the data through a regular education versus special education lens might indicate quantifiable

differences in teaching practices associated with better support for students with special needs.

Conclusions

The Oklahoma State Department of Education (2013) developed an Educator Effectiveness Theory of Action that purports that the single most important factor of student academic achievement is teacher effectiveness. Oklahoma's teacher evaluation system is based on the notion that every child deserves an effective teacher in each subject during each school year. This effectiveness was measured via the qualitative evaluation of teacher practices using the Tulsa Model for Observation and Evaluation and quantitative data yielding value-added scores of a teacher's contributions to student learning. This study identified specific teaching practices of effective teachers in the areas of Preparation, Discipline, Modeling, Monitoring, and Adjusts Based Upon Monitoring that are slightly linked to increased value-added scores for certain teachers in one Oklahoma school district. However, it is with caution that these results are interpreted as the current body of educational research is not able to arrive at a shared definition of teacher effectiveness nor convincing evidence about which teacher qualifications, practices, or characteristics impact differences in effectiveness (Goe, 2007).

REFERENCES

American Statistical Association. (2014). *ASA statement on using value-added models for educational assessment*. Retrieved from http://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*(2), 65-75. https://doi.org/10.3102/0013189X08316420

Amrein-Beardsley, A. & Collins, C. (2012). The SAS Education Value-Added System (SAS® EVAAS®) in the Houston Independent School District (HISD): Intended and unintended consequences. *Education Policy Analysis Archives, 20*(12), 1-31. Retrieved from http://files.eric.ed.gov/fulltext/EJ971428.pdf

Baker, B. D., Oluwole, J. O., Green III, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race-to-the-Top era. *Education Policy Analysis Archives, 21*(5). https://doi.org/10.14507/epaa.v21n5.2013

Baldrige Performance Excellence Program. (2011). *Baldrige 20/20: An executive's guide in the criteria for performance excellence.* Gaithersburg, Maryland: National Institute of Standards and Technology.

.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*(1), 37–65. https://doi.org/10.3102/10769986029001037

Bardach, L., Klassen, R. M., Perry, N. E. (2022). Teachers' psychological characteristics: Do they matter for teacher effectiveness, teachers' well-being, retention, and interpersonal relations? An integrative review. *Educational Psychology Review, 34*, 259-300. https://doi.org/10.1007/s10648-021-09614-9

Battelle for Kids. (2011a). *Selecting growth measures: A guide for education leaders.* Retrieved from http://static.battelleforkids.org/documents/edgrowth/Selecting_Growth_measures.pdf

Battelle for Kids. (2011b). *Value-added analysis: Informing instruction, improving teacher effectiveness and increasing opportunities for student success.* Retrieved from http://static.battelleforkids.org/images/BFK/BFK_ValueAdded_Position_Paper_Nov_2011.pdf

Berliner, D. C. (2013). Problems with value-added evaluations of teachers? Let me count the ways! *The Teacher Educator, 48*(4), 235-243. https://doi.org/10.1080/08878730.2013.827496

Betebenner, D., Braun, H., Corcoran, S., Darling-Hammond, L., Friedman, J., Goldhaber, D., … Rothstein, J. (2012). *Learning from recent advances in measuring teacher effectiveness.* Retrieved from ERIC Database: http://files.eric.ed.gov/fulltext/ED539040.pdf

Brenchley, C. (2015, April 8). *What is ESEA?* [Blog post from the United States Department of Education]. Retrieved from https://blog.ed.gov/2015/04/what-is-esea/

Chait, R. & Miller, R. (2009, July). *Getting the facts straight on the Teacher Incentive Fund* [A brief from the Center for American Progress]. Retrieved from https://cdn.americanprogress.org/wpcontent/uploads/issues/2009/07/pdf/tif_myth_fact.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Conyers, J. G., & Ewy, R. (2004). *Charting your course: Lessons learned during the journey toward performance excellence.* Milwaukee, Wisconsin: ASQ Quality Press.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, California: SAGE Publications.

Crotty, M. J. (1998). *The foundations of social research: Meaning and perspective in the research process.* Thousand Oaks, California: SAGE Publications.

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

Dininni, J. (2011). *Management theory of Frederick Taylor: Can the management theory of Frederick Taylor make your business more productive?* Retrieved from http://www.business.com/management/management-theory-of-frederick-taylor/

Doran, H. C., & Fleischman, S. (2005). Challenges of value-added assessment.

Educational Leadership, 63(3), 85-87.

Duffrin, E. (2011) What's the value in value-added? *Education Digest, 77*(2), 46-49.

Empirical Education (n.d.). *Case in point: The Tulsa Model.* Retrieved from http://www.empiricaleducation.com/pdfs/OEcase_study.pdf

Education Week. (2016). *The Every Student Succeeds Act: Explained.* Retrieved from https://www.edweek.org/ew/articles/2015/12/07/the-every-student-succeeds-act-explained.html?print=1

Farooq, M. S., Akhtar, M. S., Ullah, S. Z., & Memom, R. A. (2007). Application of Total Quality Management in education. *Journal of Quality and Technology Management, 3*(2), 87-97. Retrieved from http://files.eric.ed.gov/fulltext/ED506778.pdf

Fryer, R. G. (2011). Teacher incentives and student achievement: Evidence from New York City Public Schools. *Journal of Labor Economics, 31*(2), 373-407.

Glisson, J., & White, K. (2014) *TLE: Nuts and bolts presentation.* Retrieved from Oklahoma State Department of Education website: http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/TLE%20Nuts%20and%20Bolts_2014.pdf

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Retrieved from Center on Great Teachers and Leaders website: https://gtlcenter.org/sites/default/files/docs/LinkBetweenTQandStudentOutcomes.pdf

Goe, L. (2008). *Key issue: Using value-added models to identify and support highly effective teachers.* Retrieved from Center on Great Teachers and Leaders website:

http://www.gtlcenter.org/sites/default/files/docs/UsingValueAddedModels.pdf

Hovanetz, C., & Hellman, T. (2012). *Value-added models presentation.* Retrieved from Oklahoma State Department of Education website: http://ok.gov/sde/tle quantitative-components

Huglin, L. M. (2003). *The relationship between personal epistemology and learning style in adult learners* (Doctoral dissertation). Available from ProQuest Dissertations and Theses Database. (UMI No. 3085727)

Hunt Institute. (2016, August). *Evolution of Elementary and Secondary Education Act 1965-2015*. Retrieved from http://www.hunt-institute.org/wp-content/uploads/2016/09/Development-of-the-Elementary-and-Secondary-Education-Act-August-2016.pdf

IBM Corporation. (2016a). *SPSS: Bivariate correlations.* Retrieved from https://www.ibm.com/docs/en/spss-statistics/24.0.0?topic=option-bivariate-correlations

IBM Corporation. (2016b). *SPSS: Descriptives options*. Retrieved from https://www.ibm.com/docs/en/spss-statistics/24.0.0?topic=descriptives-options

Independence Hall Association. (2014). *Social change and national development: 22a. Economic growth and the Industrial Revolution.* Retrieved from http://www.ushistory.org/us/22a.asp

Juran Institute. (1994). *The Taylor system and quality control.* Retrieved from http://www.juran.com/elifeline/elifefiles/2009/11/Taylor-System-and-Quality-Control_JMJuran-94.pdf

Juran Institute. (2019). *The History of Quality.* Retrieved from https://www.juran.com/blog/the-history-of-quality/

Koch, J. V. (2003). TQM: Why is its impact in higher education so small? *The TQM Magazine, 15*(5), 325-333. Retrieved from http://www.jamesvkoch.com/uploads/TQM_impact_in_higher_education.pdf

Klein, A. (2015, April 10). *No Child Left Behind: An overview.* Retrieved from Education Week website: https://www.edweek.org/ew/section/multimedia/no-child-left-behind-overview-definition-summary.html#timeline

Landesberg, P. (1999). *In the beginning, there were Deming and Juran.* Retrieved from http://www.virginia.edu/processsimplification/resources/DemingJuran.pdf

Lee, L. (2011). What did the teachers think? Teachers' responses to the use of value-added modeling as a tool for evaluating teacher effectiveness. *Journal of Urban Learning, 7*, 97-103. Retrieved from http://eric.ed.gov/?id=EJ952063

Library of Congress. (n.d.) *Teacher's guide primary source set: The Industrial Revolution in the United States.* Retrieved from http://www.loc.gov/teachers/classroommaterials/primarysourcesets/industrial-revolution/pdf/teacher_guide.pdf

Manaf, Z. A., & Seng, L. S. (2010). The construction and validation of a school quality management scale for quality improvement in school management. *Journal of Islam and Contemporary Society, 3*, 3-18. Retrieved from https://journal.unisza.edu.my/jimk/index.php/jimk/article/view/36

Martinez-Lorente, A. R., Dewhurst, F., & Dale, B. G. (1998). Total Quality Management: Origins and evolution of the term. *The TQM Magazine, 10*(5), 378-386.

https://doi.org/10.1108/09544789810231261

Maurer, R. E., & Pederson, S. C. (2004). *Malcolm and me: How to use the Baldrige process to improve your school.* Lenham, Maryland: Rowman & Littlefield Publishers.

McCaffrey, D. F., & Hamilton, L. S. (2007). *Value-added assessment in practice: Lessons from the Pennsylvania value-added assessment system pilot project.* Santa Monica, CA: RAND Corporation. Retrieved from RAND Corporation website: http://www.rand.org/pubs/technical_reports/TR506.html

McNeil, M. & Klein, A. (2011, September 27). *Obama offers waivers from key provision of NCLB: Plan waives cornerstone provisions of law.* Retrieved from Education Week website: https://www.edweek.org/ew/articles/2011/09/28/05waiver_ep.h31.html

Measures of Effective Teaching Project. (2013a). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project's Three-Year Study*. Retrieved from MET Project website: https://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf

Measures of Effective Teaching Project. (2013b). *Feedback for better teaching: Nine principles for using measures of effective teaching.* Retrieved from MET Project website: http://www.metproject.org/downloads/MET_Feedback%20for%20Better%20Teaching_Principles%20Paper.pdf

Minnici, A. (2014). The mind shift in teacher evaluation: Where we stand and where we need to go. *American Educator, 38*(1), 22-26. Retrieved from

https://files.eric.ed.gov/fulltext/EJ1023882.pdf

Mohammad, S. O., & Aspinwall, E. M. (1997). TQM in higher education – a review. *International Journal of Quality & Reliability Management, 14*(5), 527-543. https://doi.org/10.1108/02656719710170747

Munoz, M. A., Scoskie, J. R., & French, D. L. (2013). Investigating the "black box" of effective teaching: The relationship between teachers' perception and student achievement in a large urban school district. *Educational Assessment, Evaluation, and Accountability, 25*, 205-230. https://doi.org/10.1007/s11092-013-9167-9

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Retrieved from https://www.edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

Office of the Governor of the State of Oklahoma. (2010, January 29). *Race to the Top: Application for initial funding CFDA number 84.395A*. Retrieved from https://www.ok.gov/recovery/documents/OklahomaRaceToTheTopApplication.pdf

Oklahoma House of Representatives. (2016). *Measure would give schools flexibility in evaluating teachers.* Retrieved from https://www.okhouse.gov/media/News_Story.aspx?NewsID=5115

Oklahoma State Department of Education. (2013). *Educator effectiveness theory of action.* Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/TLE_Theory_of_Action_Updated_nov2013.pdf

Oklahoma State Department of Education. (n.d.) *Oklahoma Teacher and Leader Effectiveness Evaluation System (TLE) report.* Retrieved from https://sde.ok.gov/sites/ok.gov.sde/files/TLE-ReportLeg.pdf

Oklahoma State Department of Education. (2014). *Value-added measures: Frequently asked questions*. Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/documents/files/TLE_Value_Added%20Measures_FAQs_3-6-14.pdf

Quattrochi, D. P., & Chapman, P. E. (2010). Giving a structural framework for Ohio's value-added model: What all educators should know? *Southeastern Regional Association of Teacher Educators Journal, 19*(2), 7-14.

RAND Corporation. (2004). *The promise and peril of using value-added modeling to measure teacher effectiveness.* Retrieved from http://www.rand.org/content/dam/rand/pubs/research_briefs/2005/RAND_RB9050.pdf

Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the black-white achievement gap.* Washington DC: Teacher's College/Economic Policy Institute.

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Sciences, 29*(1), 103-116. Retrieved from http://www.jstor.org/stable/3701308

Sallis, E. (2002). *Total Quality Management in education* (3rd ed.). New York: Taylor & Francis.

Smith, E. (2008). Raising standards in American schools? Problems with improving

teacher quality. *Teaching and Teacher Education, 24*, 610-622. https://doi.org/10.1016/j.tate.2007.09.013

Southern Regional Education Board. (2014). *Oklahoma teacher and leader evaluation systems: Focus group findings.* Retrieved from https://sde.ok.gov/sde/sites/ok.gov.sde/files/documents/files/SREB%20Focus%20Group%20Report.pdf

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education, 62*(4), 339-355. https://doi.org/10.1177%2F0022487111404241

Thompson, C., & Miller, R. (2015). *Teacher Leader Effectiveness (TLE) update.* Retrieved from http://engage.ok.gov/wp-content/uploads/2015/07/SB-706-TLE-Update.pdf

Tulsa Public Schools. (n.d.-a) *The Tulsa Model for Observation and Evaluation: Excellence and high expectations with a commitment to all.* Retrieved from http://portal.battelleforkids.org/TulsaModel/home

Tulsa Public Schools. (n.d.-b). *The Tulsa Model for Observation and Evaluation: What the process entails.* Retrieved from http://portal.battelleforkids.org/TulsaModel/overview

Tulsa Public Schools. (n.d.-c). *Tulsa Public Schools' teacher observation and evaluation system: Its research base and validation studies.* Retrieved from http://www.tulsaschools.org/4_About_District/_documents/TLE/Teacher_Eval_System_Research_Brief.pdf

Tulsa Public Schools. (2015). *TLE observation and evaluation rubric: Teachers.* Retrieved from http://www.tulsaschools.org/4_about_District/_documents/TLE/Observation_Evaluation_Rubric_Teachers.pdf

U.S. Department of Education. (n.d.). *Every Student Succeeds Act (ESSA).* Retrieved from https://www.ed.gov/esea

U.S. Department of Education. (2016). *Race to the Top fund: Awards.* Retrieved from https://www2.ed.gov/programs/racetothetop/awards.html

Walsh, E., Liu, A. Y., & Dotter, D. (2015). *Measuring teacher and school value added in Oklahoma, 2013-2014 school year.* Retrieved from http://www.ok.gov/sde/sites/ok.gov.sde/files/2013-2014%20Measuring%20VAM%20in%20Oklahoma.pdf

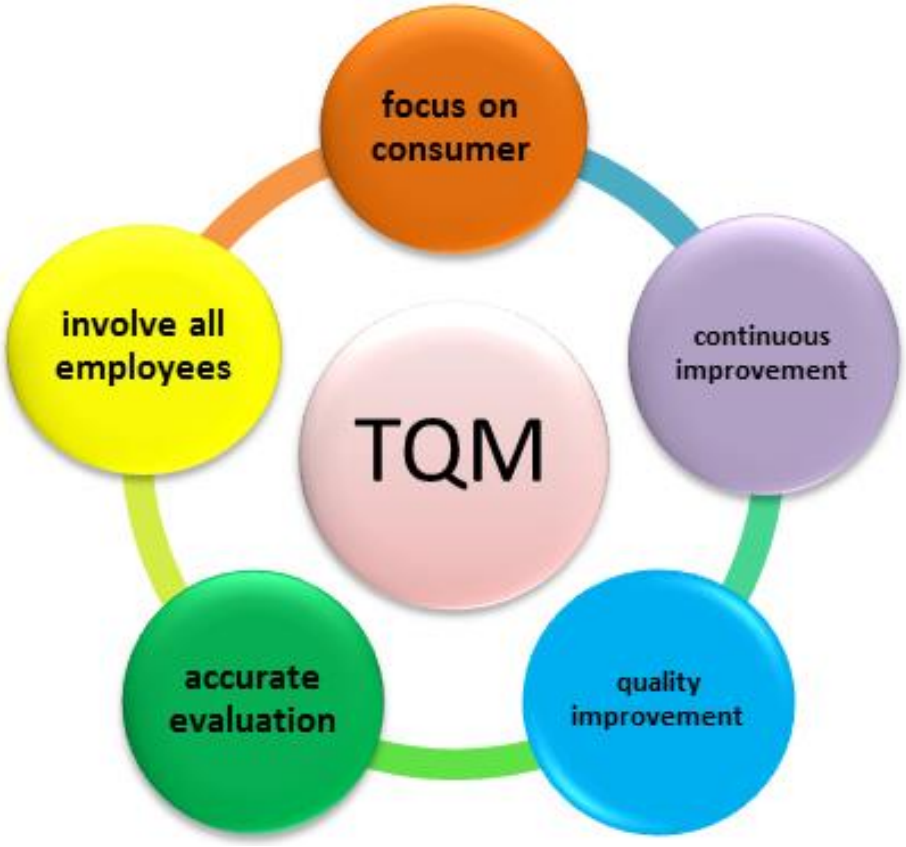
Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *Widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* Retrieved from The New Teacher Project website: http://widgeteffect.org/downloads/TheWidgetEffect.pdf

Yeh, S. (2012). The reliability, impact, and cost-effectiveness of value-added teacher assessment methods. *Journal of Education Finance, 37*(4), 374-399. https://www.jstor.org/stable/23255493

Yeh, S. S. & Ritter J. (2009). The cost-effectiveness of replacing the bottom quartile of novice teachers through value-added teacher assessment. *Journal of Education Finance, 34*(4), 426-451. Retrieved from www.jstor.org/stable/40704368

Figure 1

Essential Components of TQM



© EduKart, 2013. http://www.edukart.com/blog/total-quality-management/

Figure 2

Upside-down Organizational Structure of TQM in the Educational Setting

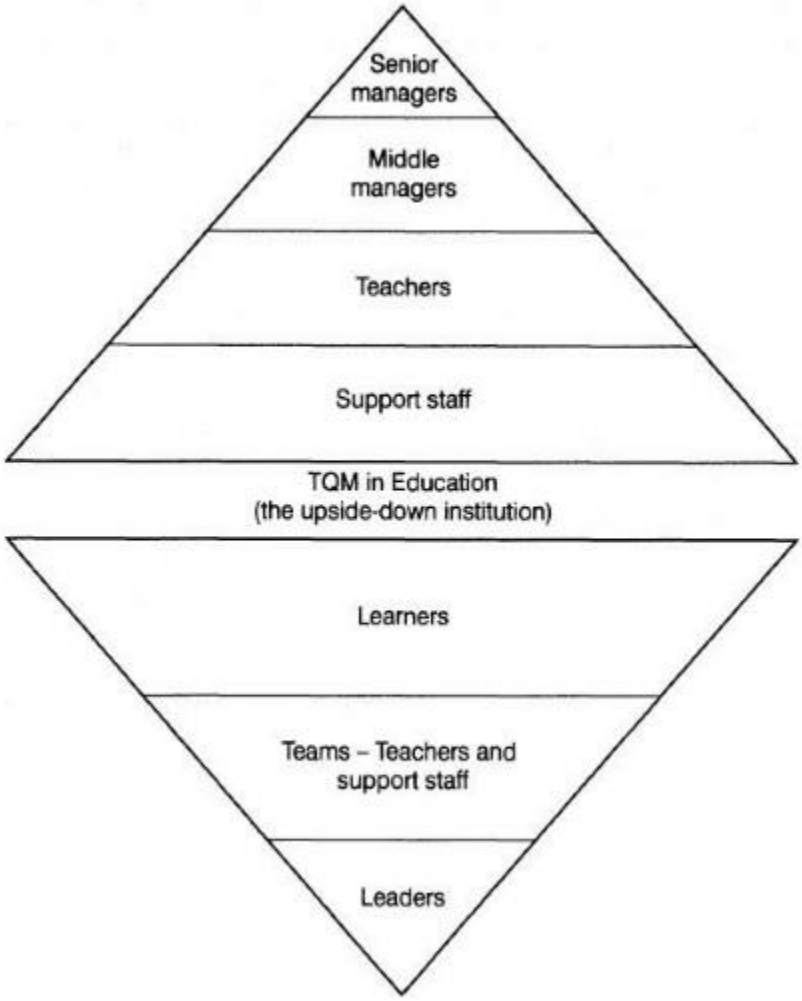


Table 27 - *Summary Table for All Correlation Analyses for Null Hypotheses*

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 1

		2013-2014 Overall Evaluation Scores	2014-2015 Overall Evaluation Scores
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.213**	.160*
	Sig. (2 tailed)	.004	.011
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.291**	.152
	Sig. (2 tailed)	.008	.083
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	.077	.031
	Sig. (2 tailed)	.444	.730
	N	102	123

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 2

		2013-2014 Evaluation Scores for Classroom Management	2014-2015 Evaluation Scores for Classroom Management
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.223**	.183**
	Sig. (2 tailed)	.002	.003
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.248*	.170
	Sig. (2 tailed)	.024	.052
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	.150	.065
	Sig. (2 tailed)	.132	.477
	N	102	123

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 3

		2013-2014 Evaluation Scores for Instructional Effectiveness	2014-2015 Evaluation Scores for Instructional Effectiveness
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.195**	.145*

	Sig. (2 tailed)	.008	.021
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.299**	.126
	Sig. (2 tailed)	.006	.152
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	.030	.042
	Sig. (2 tailed)	.765	.643
	N	102	123

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 4

		2013-2014 Evaluation Scores for Professional Growth	2014-2015 Evaluation Scores for Professional Growth
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.157*	.083
	Sig. (2 tailed)	.032	.189
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.162	.089
	Sig. (2 tailed)	.143	.310
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	.087	-.038
	Sig. (2 tailed)	.386	.673
	N	102	123

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 5

		2013-2014 Evaluation Scores for Interpersonal Skills	2014-2015 Evaluation Scores for Interpersonal Skills
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.079	.131*
	Sig. (2 tailed)	.283	.037
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.167	.125
	Sig. (2 tailed)	.131	.155
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	-.059	.020
	Sig. (2 tailed)	.554	.827
	N	102	123

Pearson Product-Moment Correlation Coefficient results for Null Hypothesis 6

		2013-2014 Evaluation Scores for Leadership	2014-2015 Evaluation Scores for Leadership
Overall Value-Added Scores by Year for All Teachers	Pearson Correlation	.105	.082
	Sig. (2 tailed)	.154	.194
	N	185	254
Overall Value-Added Scores by Year for Elem. Teachers	Pearson Correlation	.212	.151
	Sig. (2 tailed)	.054	.085
	N	83	131
Overall Value-Added Scores by Year for Sec. Teachers	Pearson Correlation	.018	-.073
	Sig. (2 tailed)	.860	.423
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 28 - *Summary Table for All Correlation Analyses for Tulsa Model Rubric Indicator Scores with Overall Value-added Scores*

Pearson Product-Moment Correlation Coefficient Results for Tulsa Model Rubric Indicators

		2013-2014 Overall Value-Added Scores	2014-2015 Overall Value-Added Scores
Preparation Indicator Scores for All Teachers by Year	Pearson Correlation	.251**	.148*
	Sig. (2 tailed)	.001	.018
	N	185	254
Preparation Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.308**	.134
	Sig. (2 tailed)	.005	.127
	N	83	131
Preparation Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.157	.100
	Sig. (2 tailed)	.116	.270
	N	102	123
Discipline Indicator Scores for All Teachers by Year	Pearson Correlation	.237**	.190**
	Sig. (2 tailed)	.001	.002
	N	185	254
Discipline Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.284**	.181*
	Sig. (2 tailed)	.009	.038
	N	83	131
Discipline Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.138	.057
	Sig. (2 tailed)	.165	.534
	N	102	123
Building-wide Climate Indicator Scores for All Teachers by Year	Pearson Correlation	.147*	.106
	Sig. (2 tailed)	.046	.093
	N	185	254
Building-wide Climate Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.138	.115
	Sig. (2 tailed)	.212	.191
	N	83	131

Building-wide Climate Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.099	-.037
	Sig. (2 tailed)	.322	.688
	N	102	123
Lesson Plans Indicator Scores for All Teachers by Year	Pearson Correlation	.102	.118
	Sig. (2 tailed)	.169	.061
	N	185	254
Lesson Plans Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.097	.115
	Sig. (2 tailed)	.383	.189
	N	83	131
Lesson Plans Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.093	.067
	Sig. (2 tailed)	.354	.462
	N	102	123
Assessment Practices Indicator Scores for All Teachers by Year	Pearson Correlation	.076	.177**
	Sig. (2 tailed)	.302	.005
	N	185	254
Assessment Practices Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.108	.128
	Sig. (2 tailed)	.331	.144
	N	83	131
Assessment Practices Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.004	.132
	Sig. (2 tailed)	.971	.145
	N	102	123
Student Relations Indicator Scores for All Teachers by Year	Pearson Correlation	.134	.110
	Sig. (2 tailed)	.068	.079
	N	185	254
Student Relations Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.145	.122
	Sig. (2 tailed)	.191	.164
	N	83	131

Student Relations Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.121	-.008
	Sig. (2 tailed)	.226	.928
	N	102	123
Literacy Indicator Scores for All Teachers by Year	Pearson Correlation	-.017	.014
	Sig. (2 tailed)	.815	.827
	N	185	253
Literacy Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.190	.144
	Sig. (2 tailed)	.085	.102
	N	83	130
Literacy Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-.235*	-.272**
	Sig. (2 tailed)	.018	.002
	N	102	123
Current State Standards Indicator Scores for All Teachers by Year	Pearson Correlation	.080	.115
	Sig. (2 tailed)	.279	.068
	N	185	253
Current State Standards Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.217*	.037
	Sig. (2 tailed)	.049	.679
	N	83	131
Current State Standards Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.011	.148
	Sig. (2 tailed)	.911	.105
	N	102	122
Involves All Learners Indicator Scores for All Teachers by Year	Pearson Correlation	.226**	.090
	Sig. (2 tailed)	.002	.151
	N	185	254
Involves All Learners Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.358**	.052
	Sig. (2 tailed)	.001	.553
	N	83	131

Involves All Learners Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.048	.046
	Sig. (2 tailed)	.633	.611
	N	102	123
Explains Content Indicator Scores for All Teachers by Year	Pearson Correlation	.159*	.112
	Sig. (2 tailed)	.030	.076
	N	185	254
Explains Content Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.192	.080
	Sig. (2 tailed)	.081	.366
	N	83	131
Explains Content Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.083	.047
	Sig. (2 tailed)	.409	.607
	N	102	123
Clear Instruction and Directions Indicator Scores for All Teachers by Year	Pearson Correlation	.135	.120
	Sig. (2 tailed)	.066	.055
	N	185	254
Clear Instruction and Directions Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.213	.079
	Sig. (2 tailed)	.053	.373
	N	83	131
Clear Instruction and Directions Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.012	.066
	Sig. (2 tailed)	.904	.467
	N	102	123
Models Indicator Scores for All Teachers by Year	Pearson Correlation	.251**	.177**
	Sig. (2 tailed)	.001	.005
	N	185	254
Models Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.329**	.133
	Sig. (2 tailed)	.002	.130
	N	83	131

Models Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.167	.207*
	Sig. (2 tailed)	.093	.022
	N	102	123
Monitors Indicator Scores for All Teachers by Year	Pearson Correlation	.188*	.138*
	Sig. (2 tailed)	.010	.027
	N	185	254
Monitors Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.234*	.094
	Sig. (2 tailed)	.033	.288
	N	83	131
Monitors Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.103	.103
	Sig. (2 tailed)	.305	.259
	N	102	123
Adjusts Based upon Monitoring Indicator Scores for All Teachers by Year	Pearson Correlation	.173*	.150*
	Sig. (2 tailed)	.019	.017
	N	184	254
Adjusts Based upon Monitoring Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.295**	.143
	Sig. (2 tailed)	.007	.104
	N	82	131
Adjusts Based upon Monitoring Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.026	.053
	Sig. (2 tailed)	.798	.558
	N	102	123
Establishes Closure Indicator Scores for All Teachers by Year	Pearson Correlation	.099	.145*
	Sig. (2 tailed)	.182	.022
	N	185	252
Establishes Closure Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.158	.105
	Sig. (2 tailed)	.154	.231
	N	83	131

Establishes Closure Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-0.24	.109
	Sig. (2 tailed)	.810	.234
	N	102	121
Student Achievement Indicator Scores for All Teachers by Year	Pearson Correlation	.047	.012
	Sig. (2 tailed)	.527	.850
	N	185	254
Student Achievement Indicator Scores for Elem. Teachers by Year	Pearson Correlation	-.017	.070
	Sig. (2 tailed)	.877	.430
	N	83	131
Student Achievement Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.055	-.128
	Sig. (2 tailed)	.583	.159
	N	102	123
Professional Development Indicator Scores for All Teachers by Year	Pearson Correlation	.038	.061
	Sig. (2 tailed)	.605	.331
	N	185	254
Professional Development Indicator Scores for Elem. Teachers by Year	Pearson Correlation	-.015	.126
	Sig. (2 tailed)	.890	.153
	N	83	131
Professional Development Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.032	-.077
	Sig. (2 tailed)	.750	.400
	N	102	123
Professional Accountability Indicator Scores for All Teachers by Year	Pearson Correlation	.220**	.077
	Sig. (2 tailed)	.003	.224
	N	185	253
Professional Accountability Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.280**	.018

	Sig. (2 tailed)	.010	.843
	N	83	130
Professional Accountability Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.106	.009
	Sig. (2 tailed)	.287	.925
	N	102	123
Interpersonal Skills Indicator Scores for All Teachers by Year	Pearson Correlation	.079	.131*
	Sig. (2 tailed)	.283	.037
	N	185	254
Interpersonal Skills Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.167	.125
	Sig. (2 tailed)	.131	.155
	N	83	131
Interpersonal Skills Indicator Scores for Sec. Teachers by Year	Pearson Correlation	-.059	.020
	Sig. (2 tailed)	.554	.827
	N	102	123
Professional Involvement Indicator Scores for All Teachers by Year	Pearson Correlation	.111	.082
	Sig. (2 tailed)	.133	.194
	N	185	254
Professional Involvement Indicator Scores for Elem. Teachers by Year	Pearson Correlation	.210	.151
	Sig. (2 tailed)	.057	.085
	N	83	131
Professional Involvement Indicator Scores for Sec. Teachers by Year	Pearson Correlation	.029	-.073
	Sig. (2 tailed)	.771	.423
	N	102	123

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

APPENDICES

Appendix A - *Tulsa Model Observation and Evaluation Rubric for Teachers*

TULSA

PUBLIC SCHOOLS

TLE Observation and Evaluation Rubric
Teachers
2013-2014

Domain/Relative Weight	*Dimension*
Classroom Management 30%	1. Preparation 2. Discipline 3. Building-Wide Climate Responsibility 4. Lesson Plans 5. Assessment Practices 6. Student Relations
Instructional Effectiveness 50%	7. Literacy 8. Common Core Standards 9. Involves All Learners 10. Explains Content 11. Explains Directions 12. Models 13. Monitors 14. Adjusts Based upon Monitoring 15. Establishes Closure 16. Student Achievement
Professional Growth & Continuous Improvement 10%	17. Uses Professional Growth as an Important Strategy 18. Exhibits Professional Behaviors and Efficiencies
Interpersonal Skills 5%	19. Effective Interactions/ Communications with Stakeholders
Leadership 5%	20. Leadership Involvements

.

Appendix A - Tulsa Model Observation and Evaluation Rubric for Teachers cont.

Indicator No.

1	Domain: **Classroom Management**	Dimension: **Preparation**

Teacher plans for delivery of the lesson relative to short-term and long-term objectives.

1 Ineffective	2 Needs Improvement	3 Effective	4 Highly Effective	5 Superior
Does not plan for instructional strategies that encourage the development of performance skills.	Occasionally plans for instructional strategies that encourage the development of performance skills.	Plans for instructional strategies that encourage the development of performance skills.	Plans for instructional strategies that encourage the development of critical thinking, problem solving and performance skills.	Plans for instructional strategies that encourage the development of critical thinking, problem solving and performance skills and consistently implements.
Materials and equipment are not ready at the start of the lesson or instructional activity.	Materials and equipment are usually not ready at the start of the lesson or instructional activity.	Ensures materials and equipment are ready at the start of the lesson or instructional activity (most of the time).	Materials and equipment are ready at the start of the lesson or instructional activity.	Materials and equipment are ready at the start of the lesson or instructional activity and learning environment is conducive to the activity.

2	Domain: **Classroom Management**	Dimension: **Discipline**

Teacher clearly defines expected behavior.

1 Ineffective	2 Needs Improvement	3 Effective	4 Highly Effective	5 Superior
Standards of conduct have not been established.	Standards of conduct have been established with inconsistent implementation.	Establishes and posts standards of conduct and implements with consistency.	Standards of conduct have been established and posted with consistent peer-based implementation.	Standards of conduct have been established and posted with consistent peer monitoring.
Students are disengaged and unclear about the expectations of the classroom.	Students are usually disengaged and unclear about the expectations of the classroom.	Ensures that students are engaged and clear as to the expectations of the classroom with few reminders given.	Students are engaged and clear about the expectations of the classroom with no need for reminders.	Students are engaged and are clear about the expectations of the classroom and are responsible for their own learning.
Does not monitor the behavior of students during whole class, small groups, seat work activities and transitions.	Rarely monitors the behavior of students during whole class, small groups, seat work activities and transitions.	Monitors the behavior of students during whole-class, small group and seat work activities and during transitions between instructional activities.	Monitors the behavior of all students during whole-class, small group and seat work activities and during transitions between instructional activities, lunch time, recess, assemblies, etc.	Monitors the behavior of all students at all times. Standards of conduct extend beyond the classroom.
Usually ignores inappropriate behavior and uses an inappropriate voice level / word choice when correction is attempted.	Most of the time ignores inappropriate behavior and / or uses an inappropriate voice level / word choice to attempt to bring correction.	Stops inappropriate behavior promptly and consistently with an appropriate voice level / word choice.	Stops inappropriate behavior promptly and consistently, with an appropriate voice level / word choice, while maintaining the dignity of the student.	Stops inappropriate behavior promptly and consistently, with an appropriate voice level / word choice, maintaining the dignity of the student and encouraging students to self-discipline.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

3	Domain: **Classroom Management**	Dimension: **Building-Wide Climate Responsibilities**		
	Teacher assures a contribution to building-wide positive climate responsibilities.			
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Is not involved in school projects and initiatives that contribute to promoting orderly behavior throughout the school.	Participates in school projects and initiatives that contribute to promoting orderly behavior throughout the school when specifically requested and only for specified time.	Regularly and routinely participates in school projects and initiatives that contribute to promoting orderly behavior throughout the school.	Participates actively in school projects and initiatives that promote orderly behavior throughout the school volunteering for extra assignments / time periods.	Makes substantial contribution to school projects and initiatives that promote orderly behavior throughout the school. Teacher assumes a leadership role in these projects and initiatives inspiring others to participate.
Ignores the procedures, practices and guidelines outlined by the school, district, state and federal laws intended to keep students healthy and safe.	Inconsistently follows the procedures, practices and guidelines outlined by the school, district, state and federal laws intended to keep students healthy and safe.	Follows the procedures, practices and guidelines outlined by the school, district, state and federal laws intended to keep students healthy and safe.	Follows the procedures, practices and guidelines outlined by the school, district, state and federal laws intended to keep students healthy and safe. Offers enhancements and suggestions to procedures and guidelines.	Always follows the procedures, practices and guidelines outlined by the school, district, state and federal laws intended to keep students healthy and safe. Is proactive in intervening on behalf of children and staff.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

4	Domain: **Classroom Management**		Dimension: **Lesson Plans**	
	Teacher develops daily lesson plans designed to achieve the identified objectives.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Only develops a brief outline of the daily schedule, which shows no alignment with state/common core standards and does not address student diversity and learning styles.	Develops instructional plans that are not in alignment with State / common core standards and does not address student's diversity and learning styles.	Develops instructional plans that are in alignment with State / common core standards including an amount of strategies that address student diversity and learning styles.	Develops instructional plans that are in alignment with State / common core standards and addresses student diversity and learning styles through differentiated instruction.	Has long and short-term instructional plans that are aligned with State / common core (CCSS) / district PASS standards and address student diversity and learning styles through differentiated instruction and other research-based learning strategies.
Plans are not completed.	Plans are rarely completed.	Plans are developed consistently and on time based upon an analysis of data.	Plans are developed consistently and on time, or in advance, based upon an analysis of data.	Plans are developed consistently and on time, or in advance, based upon an analysis of data, with inherent opportunity for continual revision.
Never plans with other members of the grade-level/school planning teams (when it is an expectation of the campus).	Rarely plans with other members of the grade-level/school planning teams (when it is an expectation of the campus).	Plans with other members of the grade-level / school planning teams (when it is an expectation of the campus).	Plans with other members of the grade-level/school planning teams (when it is an expectation of the campus). Revises plans according to student data analysis and shares same with fellow staff members to the benefit of the grade level, curricular area or building.	Plans with other members of the grade-level / school planning teams (when it is an expectation of the campus or based upon collegial decision-making). Revises plans according to student data and performance, sharing same with fellow staff members to the benefit of the grade level, curricular area or building.
Never provides substitute plans, classroom rosters, seating charts, behavior plans, emergency plans and identification of diverse learning groups.	Rarely provides substitute plans, classroom rosters, seating charts, behavior plans, emergency plans and identification of diverse learning groups.	Provides substitute plans, classroom rosters, seating charts, behavior plans, emergency plans and identification of diverse learning groups.	Provides in sequenced and organized fashion substitute plans, classroom rosters, seating charts, behavior plans, emergency plans and identification of diverse learning groups.	Can serve as a grade level, curricular area and/or building-wide model for substitute plans, classroom rosters, seating charts, behavior plans, emergency plans and identification of diverse learning groups.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

5	Domain: **Classroom Management**		Dimension: **Assessment Practices**	
Teacher acknowledges student progress and uses assessment practices that are fair and based on identified criteria.				
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Assessment is inconsistent and insufficient to determine student's overall progress and is not based on the district's grading policy.	Assessment is inconsistent and is not based on district's grading policy.	Formative and summative assessments are recorded consistently based on district's grading policy and are used to guide instruction.	Formative and summative assessments are recorded consistently based on district's grading policy and are used to develop and evaluate instruction.	Formative and summative assessments are recorded consistently based on district's grading policy and utilized to develop, refine and evaluate instruction.
Assessments provide delayed and inadequate feedback for students to assess themselves.	Assessments provide delayed and inadequate feedback for students to assess themselves.	Provides adequate and timely feedback from assessment results for students to reflect and set goals.	Assessments provide useful and immediate feedback that assists students in assessing themselves in meeting their learning goals.	Assessments provide useful and immediate feedback that assists students in assessing themselves to develop and evaluate their progress with their learning goals. Learning goals are not only designed by the teacher but the student has an opportunity to direct his/her own learning by contributing goals.
There is no evidence that the teacher recognizes student progress or achievement.	There is some evidence that students are recognized for their progress and achievement; however, recognition is sporadic.	Recognizes student progress and achievement at significant intervals and encourages behaviors that would result in student success.	Students are informed regularly regarding their progress and achievement and are provided opportunities to improve and achieve academic success.	Students are informed regularly regarding their progress and achievement and are provided opportunities to improve and achieve academic success. The teacher informs parents on a timely basis of their student's progress and achievement through systematic communication procedures.

6	Domain: **Classroom Management**			Dimension: **Student Relations**
Teacher optimizes the learning environment through respectful and appropriate interactions with students, conveying high expectations for students and an enthusiasm for the curriculum.				
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Oral, written and nonverbal communication with students is inconsiderate, as characterized by insensitivity, demeaning language and condescension.	Oral, written, and nonverbal communication may not be considerate or respectful.	Oral, written and nonverbal communications with students are considerate and respectful.	Oral, written, and nonverbal communications with students are considerate and positive, demonstrating genuine respect for individual students and the class as a whole.	Oral, written, and nonverbal communication with students is considerate and positive. There is abundant evidence of mutual respect and trust between teacher and student, as well as between students.
Does not consistently display an interest in the curriculum or high expectations for most students.	Does not consistently display an interest in the curriculum or high expectations for most students.	Consistently conveys a generally positive view of learning and of the curriculum, demonstrating high expectations for most students.	Consistently displays a genuine enthusiasm for the curriculum and high expectations for all students	Exudes a passion for the content and actively exploring the curriculum with students. Students appear to have internalized the value of the content as well as the teacher's high expectations for them.

7	Domain: **Instructional Effectiveness**			Dimension: **Literacy**
Teacher embeds the components of literacy into all instructional content.				
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Literacy, the practice of reading, writing, developing vocabulary, spelling, listening and speaking, is not embedded / woven into instructional lessons; rather, literacy is presented as a single, stand-alone skill.	Literacy, the practice of reading, writing, spelling, listening and speaking, is rarely embedded / woven into instructional lessons as an explicit learning objective; rather, literacy is presented as a single, stand-alone skill.	Literacy, the practice of reading, writing, spelling, listening and speaking, is embedded in ALL content as an explicit learning objective.	Literacy, the practice of reading, writing, spelling, listening and speaking, is embedded in ALL content as an explicit learning objective and its definition is expanded to include visual representations, expressions of ideas, making decisions and solving problems.	Includes the narrative descriptions in performance category 4, plus the additional definitional components of literacy to include: innovative use of multimedia, computer, information analysis and technology.
Does not appear to value or recognize that literacy is the "bonding agent" for all learning.	Demonstrates weak recognition of the importance of literacy as the "bonding agent" for all learning.	Displays basic recognition of the importance of literacy as the "bonding agent" for all learning.	Leverages literacy as the "bonding agent" for all learning	

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

8	Domain: **Instructional Effectiveness**		Dimension: **Common Core Standards**	
Teacher understands and optimizes the delivery focus of Common Core State Standards and the expectations derived from same on student learning and achievement.				
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Neither understands nor participates (at even the "conversation / awareness" level) in the multi-year conversion process from PASS to CCSS.	Neither understands nor participates (at even a minimal implementation level) in the multi-year conversion process from PASS to CCSS.	Understands and participates in the multi-year conversion process from an emphasis on PASS to an emphasis on CCSS as evidenced by use of alternate instructional strategies and modified content focus aligned with CCSS.	Has participated in available learning opportunities to assure a strong foundation of understanding the conversion process from PASS to CCSS and regularly and routinely uses alternate instructional strategies and modified content focus aligned with CCSS.	Includes the narrative descriptions in performance category 4, plus serves as a "change agent" and/or grade level, curricular area, building-wide, or departmental presenter / facilitator for the implementation of the conversion from PASS to CCSS. This participation level could be initiated via volunteering or being asked.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

9	Domain: **Instructional Effectiveness**		Dimension: **Involves All Learners**	
Teacher uses active learning, questioning techniques and/or guided practices to involve all students.				
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Students are not mentally engaged in active learning experiences during any significant portion of the class.	A few students dominate the lesson, and only a few students are minimally engaged in active learning experiences 50 percent of the class time.	Engages most students in active learning experiences 80 percent of the class time.	An overwhelming majority of students are cognitively engaged and exploring content in active learning experiences 80 percent of the class time.	All students are cognitively engaged and exploring 80 percent of the class time. Students initiate or develop their own activities to enhance their learning.
Does not ask any type of questions or use questioning techniques during the lesson to involve all learners.	All or most questions used are recall questions.	Uses questioning techniques throughout the lesson, scaffolding to at least the mid-level of Bloom's taxonomy. Provides wait time for some student response and does random checking to ensure the involvement of all learners.	Uses consistently high-quality and varied questioning techniques, scaffolding to the higher levels of Bloom's taxonomy and providing adequate wait time for most students to respond.	Uses consistently high-quality and varied questioning techniques, scaffolding to the higher levels of Bloom's taxonomy and leading students to formulate many of their own questions. Provides adequate wait time for most students to respond.
Student participation is not monitored or the teacher response is inconsistent, overly repressive or does not respect the student's dignity.	Typically calls on students who raise their hands first and responds to students who blurt out answers.			
Displays no knowledge of students' interests and skills.	Displays little knowledge of students' interests and skills and rarely uses them as a strategy to engage them.	Engages students by incorporating their general skills and interests into the lesson.	Engages students by incorporating their individual skills and interests into the lesson.	Engages students by incorporating and expanding their individual skills and interests.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

10	Domain: **Instructional Effectiveness**		Dimension: **Explains Content**	
	Teacher teaches the objectives through a variety of methods.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Does not use cooperative learning activities, advance organizers, or other teaching strategies that foster student participation and an understanding of the objectives.	Uses limited cooperative learning activities, advance organizers, or other teaching strategies that foster participation and an understanding of the objectives.	Uses cooperative learning activities, advance organizers, or other teaching strategies that foster participation and an understanding of the objectives.	Uses knowledge of student skills and interests when selecting and using cooperative learning activities, advance organizers, and other teaching strategies that foster participation and an understanding of the objectives.	Uses all of the characteristics of Level 4. In addition, continually seeks out new strategies to support instructional outcomes and cognitively challenge diverse learners. Willingly shares discoveries and successes with colleagues. Students are included in planning for methods of instructional delivery.
Students are provided with activities from the textbook, specific to the content, but there is no attempt to use a variety of activities to support instructional outcomes and no attempt to differentiate tasks to address a variety of student needs/learning styles / multiple intelligences.	Attempts, but does not successfully use a variety of activities (e.g. modeling, visuals, hands-on activities, demonstrations, gestures, body language and thematic instruction) to support instructional outcomes and meet varied student needs/ learning styles / multiple intelligences.	Uses a variety of activities (e.g. modeling, visuals, hands-on activities, demonstrations, gestures, body language and thematic instruction) to support the instructional outcomes and meet varied student needs/ learning styles / multiple intelligences.	Successfully uses a variety of activities (e.g. modeling, visuals, hands-on activities, demonstrations, gestures, body language and thematic instruction) to support the instructional outcomes and meet varied student needs/ learning styles / multiple intelligences. The activities maximize student potential and most require significant cognitive challenge.	
Technology is not used as designed and not used as an instructional tool.	Technology is rarely included in the planning process to support instruction, and technology is not used on a regular basis as an instructional tool.	Technology is included in the planning process to support instruction, and technology is used on a regular basis as an instructional tool.	Technology is woven into / serves as a foundational base in the planning process to support instruction, and technology is used on a common-place basis as an instructional tool.	

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

11	Domain: **Instructional Effectiveness**		Dimension: **Explains Directions**	
	Teacher gives directions that are clearly stated and relate to the learning objectives.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Directions and procedures are confusing to students.	Directions and procedures are initially confusing to students and are not clarified.	Provides directions and procedures, in a variety of delivery modes, e.g., verbal, modeling, visual, demonstration, etc., that are clearly stated / presented and relate to the learning objectives.	Directions and procedures, in a variety of delivery modes, are clear to students. Anticipation of possible student misunderstanding and/or confusion is incorporated in the initial direction and clarified.	Uses all of the characteristics of Levels 3 and 4. Facilitates students in constructing their own understanding of how the directions relate to the learning objectives.
Does not offer alternative, clarifying directions. Does not give students directions for transitions and does not plan for transitions.	Attempts to give students directions for transitions but does not plan for transitions.	Gives students directions for transitions and includes transitioning in the planning process to optimize academic learning time.	Gives clear directions for transitions between lessons and between instructional activities while optimizing academic learning time.	Plans for smooth, structured transitions between lessons and instructional activities and gives clear, concise directions to accomplish same while optimizing academic learning time.
Spoken language is inaudible or written language is illegible. Spoken or written language contains errors of grammar or syntax. Vocabulary may be inappropriate, vague, or used incorrectly causing students to be confused.	Spoken language is audible and written language is legible. Usage of both demonstrates many basic errors (mispronunciation, misspelled words, etc.). Vocabulary is correct, but limited, or is not appropriate to the students' ages or backgrounds.	Uses spoken and written language that is clear and correct, conforms to standard English, vocabulary, and is appropriate to students' ages and interests.	Spoken and written language is clear and correct and conforms to standard English. Vocabulary is appropriate to the students' ages and interests. Teacher finds opportunities to extend students' vocabularies.	Spoken and written language is correct and conforms to standard English. It is also expressive with well-chosen vocabulary that enriches the lesson and extends students' vocabularies. Teacher seizes opportunities to enhance learning by building vocabulary skills and experiences based on student interests or a spontaneous event.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

12	Domain: **Instructional Effectiveness**			Dimension: **Models**
	Teacher demonstrates / models the desired skill or process.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Does not demonstrate or model the desired skill or process.	Demonstration or modeling of the desired skill or process is infrequent and unclear to students.	Provides demonstrations and modeling of the desired skill or process that are clear and precise to students.	Demonstrations are clear and precise to students with anticipation and preemptive action to avoid possible students' misunderstanding.	Demonstrations will match all characteristics of Level 4. Additionally, teacher's modeling will assist students in achieving the lesson's stated objective. Students will demonstrate the skill or process.

13	Domain: **Instructional Effectiveness**			Dimension: **Monitors**
	Teacher checks to determine if students are progressing toward stated objectives.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Never moves around the room while students are working on guided practice.	Seldom moves around the room while students are working on guided practice to promote and reinforce positive student behaviors. When movement happens it is to the same area of classroom.	When appropriate, moves to all areas of the room while students are working on guided practice to promote and reinforce positive student behaviors.	Moves to all areas of the room with efficiency and effectiveness while students are working on guided practice to promote and reinforce positive student behaviors. Makes eye contact with all students often.	Moves throughout the room to assure optimal instructional impact while students are working on guided practice to promote and reinforce positive student behaviors. When a problem is observed reviews / re-teaches it to the whole class.
Never uses student response techniques to increase active engagement.	Seldom uses student response techniques to increase active engagement.	Uses different types of student response techniques, both individual / group. Uses student response techniques to increase active engagement.	Routinely uses developmentally appropriate student response techniques to increase active engagement by the students.	Delivers upon all of performance category 4 and varied response techniques are used to provide immediate feedback to re-teach / review the concept(s) misinterpreted or not learned, while actively engaging all students.
Never uses feedback concerning student's understanding.	Seldom uses feedback concerning student's understanding.	Student's understanding is evaluated by feedback.	Uses immediate feedback concerning student's understanding.	Delivers upon all of performance category 4 and is able to assess when question / wait time is no longer effective and employs a different strategy / technique.
Never uses wait time after voicing a question to the students.	Seldom uses wait time after voicing a question to the students.	Uses wait time of 3-5 seconds (more for more complex questions) after voicing the question. Provides opportunity for students to formulate more thoughtful responses and allows time for the student to consider supporting evidence.	Routinely uses wait time of 3-5 seconds (additional time for more complex questions) after voicing the question. Provides opportunity for students to formulate more thoughtful responses and allows time for the student to consider supporting evidence. Re-phrases the question after hearing student response to probe for deeper understanding of concept utilizing appropriate wait time.	

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

14	Domain: **Instructional Effectiveness**		Dimension: **Adjusts Based Upon Monitoring**	
	Teacher changes instruction based on the results of monitoring.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Does not adjust instructional plan to meet the needs of students. Lesson pace is too fast or slow to accommodate for students' questions or interest.	Inconsistently monitors student involvement and makes some effort to adjust instructional plans to engage more students.	Consistently monitors student involvement and makes efforts to adjust instructional plans to engage more students.	Is aware of student participation and smoothly makes appropriate adjustments to the lesson successfully accommodating student questions or interests.	Is always aware of student participation and successfully engages all students in the lesson. Is able to successfully make adjustments to the lesson to accommodate student questions or interests.
Does not assess mastery of the new learning to determine if independent practice or re-teaching is appropriate.	Inconsistently assesses mastery of the new learning to determine if independent practice or re-teaching is appropriate without making adjustments as necessary.	Assesses mastery of the new learning to determine if independent practice or re-teaching is appropriate and makes adjustments to lessons.	Assesses mastery of the new learning using a variety of methods to determine if independent practice or re-teaching is appropriate and restructures lessons to address various learning needs.	Assesses mastery of the new learning using a variety of methods to determine if independent practice or re-teaching is appropriate. Works with individual students or small groups to reteach. Uses peer tutoring to facilitate mastery of skills.
There is no evidence that the teacher uses data from various assessments to modify instruction and guide intervention strategies.	There is little evidence that data is used from various assessments to modify instruction and guide intervention strategies.	Reviews data from assessments to modify instruction and guide intervention strategies.	Uses data from various assessments to modify instruction and to determine what additional interventions can be implemented to assist students.	Multiple classroom evaluations, assessments and formal State assessments provide ample and varied opportunity for all students to demonstrate their knowledge and skill set levels. Ongoing assessment is systematically used to modify instruction and guide intervention strategies.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

15	Domain: **Instructional Effectiveness**		Dimension: **Establishes Closure**	
	Teacher summarizes and fits into context what has been taught.			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
There is no ending to the lesson. Students disengage at the end of the class with no teacher direction.	The teacher ends the lesson without a summary of the main points of the segment of instruction or day's learning/activity.	Ends the day's learning / activity by summarizing the lesson or asking students to summarize the lesson.	Ends the day's learning / activity by summarizing the lesson in a variety of ways. Students are able to summarize in a variety of ways and reflect on their own learning.	Ends the day's learning / activity by facilitating students in summarizing and discussing main ideas.
Does not connect what is learned to prior learning and does not relate how the learning will be needed in the future.	Does not connect what is learned to prior learning and does not relate how the learning will be needed in the future.	Connects what is learned to prior learning.	Relates instruction to prior and future learning.	Students are able to connect the lesson to prior learning and articulate how learned skills can be used in the future. Linkages with real world situations are woven into the lessons.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

16	Domain: **Instructional Effectiveness**		Dimension: **Student Achievement**	
Effective development and use of modified assessments and curriculum for special education students and other students experiencing difficulties in learning.				
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
There is no evidence that the teacher is knowledgeable of the IEP or that the teacher modifies instruction for all students on an IEP regardless of student's learning goals.	There is some evidence that the teacher is aware of the IEP; however, the IEP is not being used to guide instruction for the student.	Modifies assessments for special education student populations in alignment with the IEP.	Modifies assessments for special education student populations as indicated in the IEP and as needed, working with individual students to develop a mutually acceptable plan for "success."	Modifies assessments and curriculum for special student populations as indicated in the IEP and as needed, working with individual students to develop a mutually acceptable plan for "success."
		Provides required feedback to student, roster teacher and/or parent.	Provides frequent / timely feedback to student, teacher or parent.	Provides frequent/timely feedback to student, roster teacher and parent of the results of modifications on student progress and participates as a team member in recommending needed changes in modifications.
		Assures that all students have access to standard / common core / district curriculum.	Assures that all students have access and modifications to standard /common core /district curriculum.	The teacher consistently advocates for all special needs students to have direct access to standard /common core /district curriculum.
Gives up, blames the student, or blames the student's home environment if the student has difficulty learning.	When a student has difficulty learning, the teacher makes an ineffectual effort and quickly gives up or blames the student or the student's home environment.	Accepts responsibility for the success of all students.	When a student has difficulty learning, the teacher perseveres to identify effective approaches to reach the student, drawing on a broad repertoire of strategies.	Perseveres in seeking effective approaches for students who need help using an extensive repertoire of strategies and soliciting additional resources from the school and community. Maintains contact with the student to monitor and support the student's success even after the student has moved on to another class.

17	Domain: **Professional Growth and Continuous Improvement** **Uses Professional Growth as a Continuous Improvement Strategy**			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Does not participate in professional development that updates their content knowledge and professional practices.	Participates in a portion of the required minimum hours of professional development. The professional development does not update their content knowledge and current professional practices.	Participates in the required minimum hours of professional development updating their content knowledge and current professional practices.	Participates in the required hours of professional development and seeks additional training to update their content knowledge and professional practices beyond what is required.	In addition to participating in the required hours of prof. development and add'l training, the teacher makes a substantial contribution to the profession through activities such as, coaching and mentoring new teachers, training teachers in professional practices, making presentations, conducting action research, working towards Master Teacher Certification and/or writing articles for grade level, department level, internal / school-wide and/or external publication. Writings that could be used as "models" may include classroom newsletters, parent / community communications, etc.

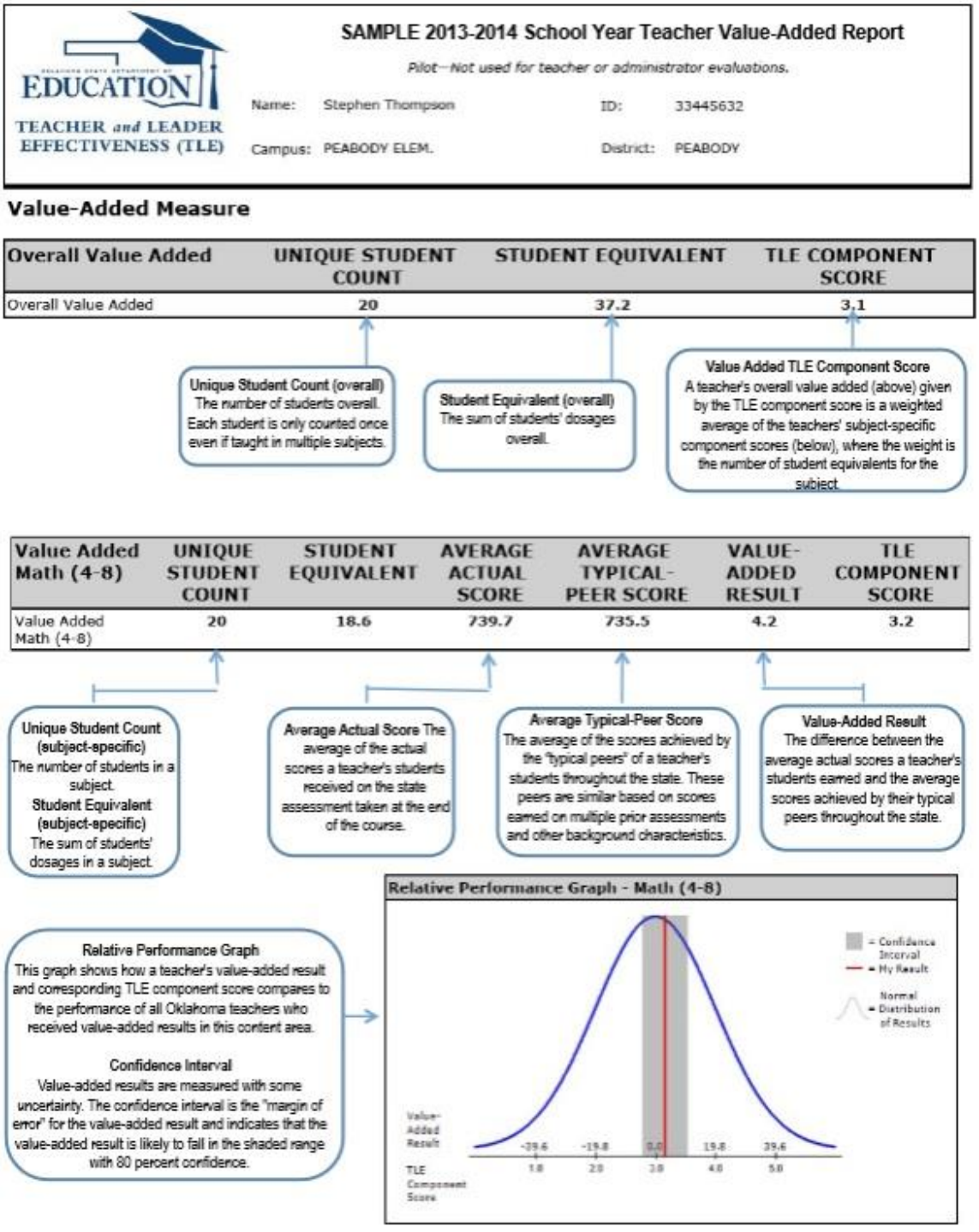
18	Domain: **Professional Growth and Continuous Improvement** **Exhibits behaviors and efficiencies associated with professionalism.**			
1 Ineffective	**2 Needs Improvement**	**3 Effective**	**4 Highly Effective**	**5 Superior**
Exhibits documentable patterns of repeated inconsistent reliability-based behavior patterns as delineated in performance category 3 – Effective.	Exhibits inconsistent reliability-based behavior patterns as evidenced by flawed punctuality and dependability; not adhering to prescribed arrival and departure times; not following notification and reporting procedures for absences; not complying with reporting timelines and other time sensitive info./compliance requests.	Exhibits consistent reliability-based behavior patterns as evidenced by punctuality and dependability; adhering to prescribed arrival and departure times; following notification and reporting procedures for absences; complying with reporting timelines and other time sensitive info./compliance requests.	Exhibits highly consistent reliability-based behavior patterns as evidenced by punctuality and dependability; adhering to prescribed arrival and departure times; following notification and reporting procedures for absences; complying with reporting timelines and other time sensitive info./compliance requests.	Serves as a model and mentor exhibiting consistent reliability-based behavior patterns as evidenced by punctuality and dependability; adhering to prescribed arrival and departure times; following notification and reporting procedures for absences; complying with reporting timelines and other time sensitive info./compliance requests.

Appendix A – Tulsa Model Observation and Evaluation Rubric for Teachers cont.

19	Domain: **Interpersonal Skills** **Effective Interactions and Collaboration with Stakeholders.**				
1 **Ineffective**	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Provides minimal or no information to families and colleagues and makes no attempt to engage them in the educational program.	Appears to be inconsistent and inaccurate in providing information to families and colleagues and engaging them in the educational program.	Interacts with families and colleagues in a timely, consistent, positive and professional manner. Complies with school procedures for communicating with families and colleagues and makes an effort to engage them in the educational program.	Communicates frequently and sensitively with families and colleagues and engages them in the educational program.	Communicates consistently and sensitively with families and colleagues and uses diverse methods to engage them in the educational program and supports their participation. Communication is clearly understood by diverse stakeholders.
Does not consult or collaborate with other staff members.	Plans and makes decisions assuming the result will be positive for everyone. Consults infrequently with other staff members.	Collaborates appropriately and makes decisions that reflect genuine professional consideration.	Maintains an open mind and participates in collaborative planning, reflection and decision making, respecting and considering the thoughts of colleagues.	Takes a leadership role in ensuring that all collaborative decisions, planning and reflection activities with colleagues are based on the highest professional standards. Seeks out the expertise and opinion of other professionals before considering collaborative decisions.

20	Domain: **Leadership** **Exhibits Positive Leadership through Varied Involvements.**			
Ineffective	**2** **Needs Improvement**	**3** **Effective**	**4** **Highly Effective**	**5** **Superior**
Consistently declines becoming involved in school or district events when asked.	Avoids becoming involved in school or district events.	Agrees to participate in school or district events when asked.	Volunteers or eagerly accepts an invitation to substantially contribute to a school or district event.	Develops or leads important school or district events.
Impedes colleagues' efforts to share their knowledge or assume professional responsibility.	Makes no effort to assume professional responsibilities or share professional knowledge with colleagues in the school or district.	Finds ways to contribute to the profession and follows through.	Actively participates in assisting other educators in their growth as professionals.	Initiates important activities contributing to the profession, such as mentoring new teachers, writing articles for publication or making presentations.
Perpetuates biased, negative or disrespectful attitudes or practices in the school that impede the school's ability to serve all students.	Rarely contributes to the modification of school practices that would result in students being better served by the school.	Assumes a proactive role in addressing student needs.	Works within a team of colleagues to ensure that all students have a fair and equal opportunity to learn and succeed in school.	Leads others to challenge and reject biased, negative or disrespectful attitudes or practices in the school that impede the school's ability to serve all students.

Appendix B – Sample Value-Added Report



Accuracy of student roster subject assignments may depend on whether you and/or your district chose to participate in roster verification in Spring 2014. The value added results in this report are based on the roster and assignment as listed.

Appendix B – Sample Value-Added Report cont.

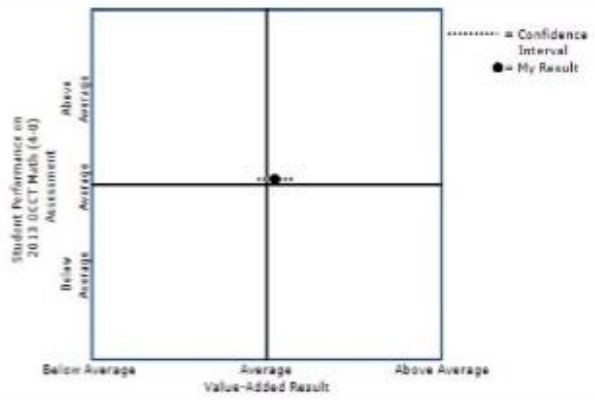
Student Performance vs. Teacher Value-Added Result

This graph shows a teacher's value-added result and the performance of that teacher's students relative to the statewide average on the state assessment. With value-added measures, teachers can be identified as high performers regardless of their students' achievement levels.

Confidence Interval

The confidence interval is the "margin of error" for a value-added result and indicates that the value-added result is likely to fall in the shaded range with 80 percent confidence. A smaller confidence interval means the value-added result is a more precise measure of teacher effectiveness. In general, a result based on a larger number of students will have a smaller confidence interval.

SAMPLE Student Performance vs. Teacher Value-Added Result Comparison Graph - Math (4-8)



% of Year

The portion of the year the student was enrolled in the course a teacher taught in this subject from September through April.

% of Instruction

The percentage of responsibility for instruction a teacher was assigned for the student during the period he or she was enrolled in the course from September through April.

Dosage

The amount of instructional time a teacher spent with a student. This is equal to the percent of instructional time multiplied by the percent of the year with a teacher. Dosage is used to weight students in the value-added calculation.

Student Roster

The roster includes students who contribute to a teacher's value-added result for the subject. Students who are not eligible to be included in the value-added model are excluded from this list. For example, students must have valid pre-test and post-test scores to be included. The roster is based on data from roster verification, when available.

PARTIAL Student Roster - Math (4-8)

STUDENT (Please note: Student roster has been truncated to save space on sample report only)	% of Year	x	% of Instruction	=	Dosage
JOE BARKLEY	100%		100%		100%
SUSIE SMITH	50%		100%		50%
JIMMY JONES	100%		100%		100%

Note: All percentages have been rounded to the nearest ten percentage points for display on this report. For example, an instruc=onal percentage of 75 percent was rounded to 80 percent. For accuracy, all calcula=ons related to value-added results were done using unrounded numbers.

Prior Achievement

Based on student test scores from the prior grade and year. For grades 4 through 8 math, Algebra I, Geometry, or Algebra II, this is another mathematics assessment. For grades 4 through 8 reading, or English III, this is another reading/ELA assessment.

Value Added By Sub-Groups - Math (4-8)

Prior Achievement	Unique Student Count	Student Equivalent	Value-Added Result
Advanced	5	***	***
Proficient	10	9.2	Average
Limited Knowledge	3	***	***
Unsatisfactory	2	***	***
Additional Groups	**Unique Student Count**	**Student Equivalent**	**Value-Added Result**
ELL	4	***	***
Not ELL	16	14.4	Average
IEP	4	***	***
Not IEP	16	14.4	Average

***

Indicates that there is insufficient data to yield a result for this metric.

Accuracy of student roster subject assignments may depend on whether you and/or your district chose to participate in roster verification in Spring 2014. The value added results in this report are based on the roster and assignment as listed.

Appendix B – Sample Value-Added Report cont.

SAMPLE Value Added Reading(4-8)	UNIQUE STUDENT COUNT	STUDENT EQUIVALENT	AVERAGE ACTUAL SCORE	AVERAGE TYPICAL-PEER SCORE	VALUE-ADDED RESULT	TLE COMPONENT SCORE
Value Added Reading (4-8)	20	18.6	720.5	721.8	-1.3	2.9

Unique Student Count (subject-specific)
The number of students in a subject.
Student Equivalent (subject-specific)
The sum of students' dosages in a subject.

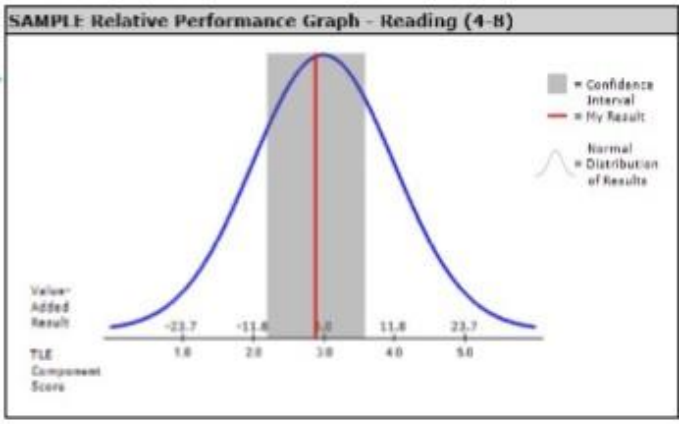
Average Actual Score The average of the actual scores a teacher's students received on the state assessment taken at the end of the course.

Average Typical-Peer Score
The average of the scores achieved by the "typical peers" of a teacher's students throughout the state. These peers are similar based on scores earned on multiple prior assessments and other background characteristics.

Value-Added Result
The difference between the average actual scores a teacher's students earned and the average scores achieved by their typical peers throughout the state.

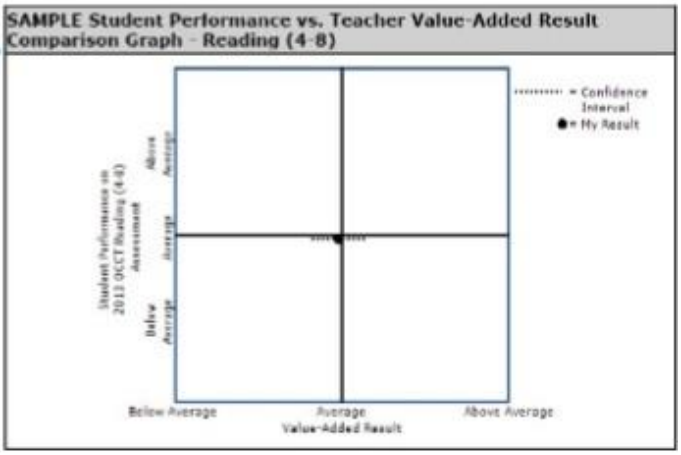
Relative Performance Graph
This graph shows how a teacher's value-added result and corresponding TLE component score compares to the performance of all Oklahoma teachers who received value-added results in this content area.

Confidence Interval
Value-added results are measured with some uncertainty. The confidence interval is the "margin of error" for the value-added result and indicates that the value-added result is likely to fall in the shaded range with 80 percent confidence.

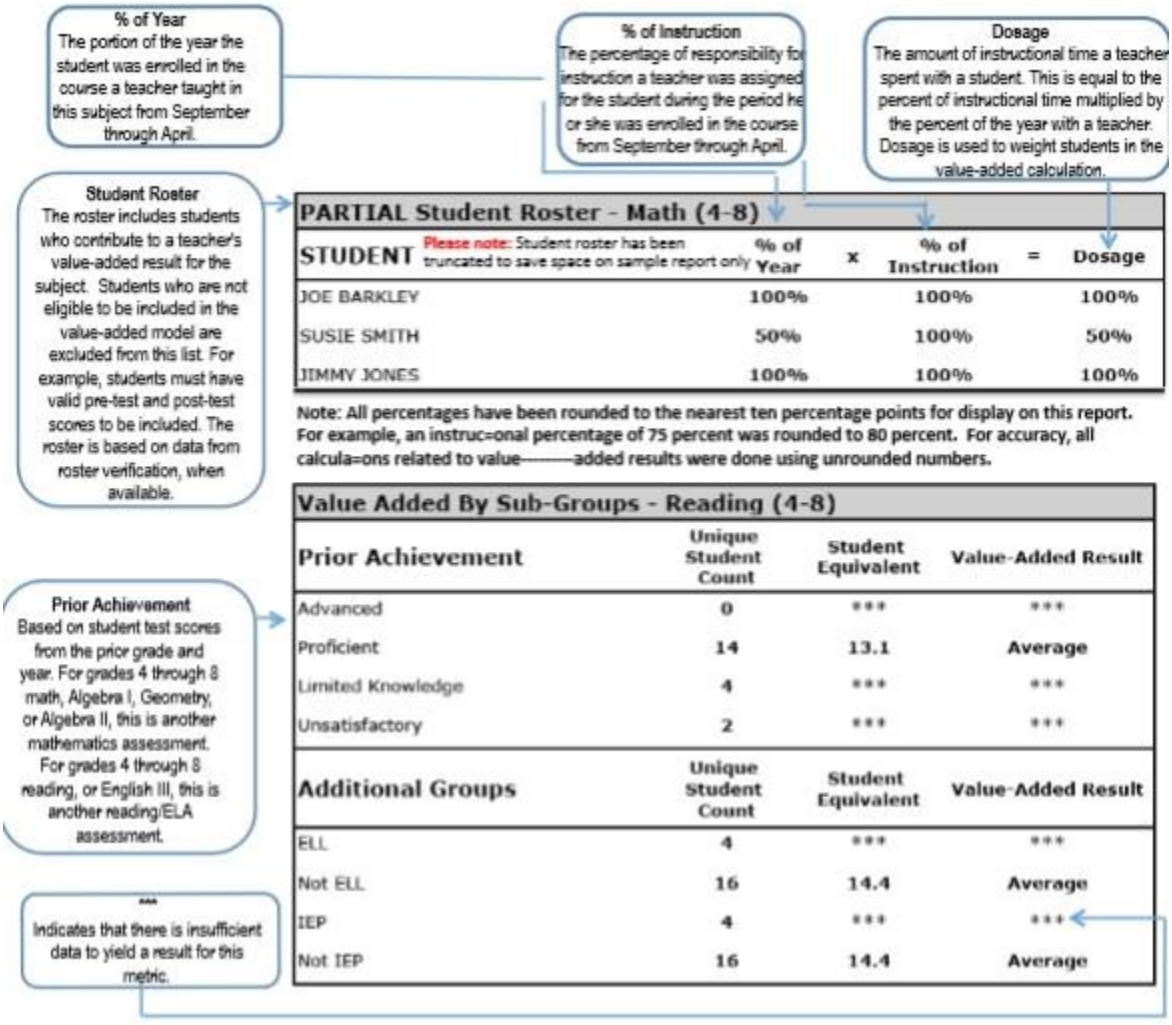


Student Performance vs. Teacher Value-Added Result
This graph shows a teacher's value-added result and the performance of that teacher's students relative to the statewide average on the state assessment. With value-added measures, teachers can be identified as high performers regardless of their students' achievement levels.

Confidence Interval
The confidence interval is the "margin of error" for a value-added result and indicates that the value-added result is likely to fall in the shaded range with 80 percent confidence. A smaller confidence interval means the value-added result is a more precise measure of teacher effectiveness. In general, a result based on a larger number of students will have a smaller confidence interval.



Appendix B – Sample Value-Added Report cont.



VITA

Amy Castro Braun

Candidate for the Degree of

Doctor of Education

Dissertation: EXAMINING TEACHER EFFECTIVENESS THROUGH VALUE-ADDED SCORES AND OBSERVED TEACHING PRACTICES

Major Field: School Administration

Biographical:

Education:

Completed the requirements for the Doctor of Education in School Administration at Oklahoma State University, Stillwater, Oklahoma in May, 2022.

Completed the requirements for the Master of Science in Educational Leadership at University of Central Oklahoma, Edmond, OK/USA in 2013.

Completed the requirements for the Master of Science in Instructional Leadership and Academic Curriculum at the University of Oklahoma, Norman, OK/USA in 2004.

Completed the requirements for the Bachelor of Science in Education at the University of Oklahoma, Norman, OK/USA in 2002.

Experience:

University of Central Oklahoma - Adjunct Professor, Summers 2019 and 2020
Moore Public Schools - Principal, 2016 - present; Assistant Principal, 2014 - 2016; District Facilitator of Teacher/Leader Effectiveness Evaluation System, 2014 – 2016; 1st Grade Teacher, 2005 – 2013

Professional Memberships:

Oklahoma Education Association, Cooperative Council for Oklahoma School Admin.