

University of the Incarnate Word

## The Athenaeum

---

Theses & Dissertations

---

5-2021

### Effectiveness of Goal Setting on Teacher Evaluation

Julie White

University of the Incarnate Word, [jmwhite1@student.uiwtx.edu](mailto:jmwhite1@student.uiwtx.edu)

Follow this and additional works at: [https://athenaeum.uiw.edu/uiw\\_etds](https://athenaeum.uiw.edu/uiw_etds)



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Educational Leadership Commons](#)

---

#### Recommended Citation

White, Julie, "Effectiveness of Goal Setting on Teacher Evaluation" (2021). *Theses & Dissertations*. 387. [https://athenaeum.uiw.edu/uiw\\_etds/387](https://athenaeum.uiw.edu/uiw_etds/387)

This Dissertation is brought to you for free and open access by The Athenaeum. It has been accepted for inclusion in Theses & Dissertations by an authorized administrator of The Athenaeum. For more information, please contact [athenaeum@uiwtx.edu](mailto:athenaeum@uiwtx.edu).

EFFECTIVENESS OF GOAL SETTING ON TEACHER EVALUATION

by

Julie White

A DISSERTATION

Presented to the Faculty of the University of the Incarnate Word  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF THE INCARNATE WORD

May 2021

© Copyright by Julie White  
2020

## ACKNOWLEDGMENTS

First and foremost, I would like to give thanks to God for giving me the guidance and encouragement to complete this dissertation.

I would like to express my gratitude to my dissertation chair Dr. Stephanie Hartzell, for her expertise, guidance, and patience through this journey. I would also like to thank my dissertation committee Dr. Stephanie Grote-Garcia and Dr. Duard Mark Littleton, for your suggestions and encouragement.

I would like to extend my gratitude to the faculty of the Dreeben School of Education. You have each made my time at University of the Incarnate Word enjoyable. To all my colleagues that I have met on my journey, thank you for all the support and laughs.

I am extremely grateful for my sons, Kade and Miller, for their love, prayers and sacrifices. I would also like to give great gratitude to my husband, Cody, for his love, and patience. I am grateful to my parents for their love and support through my life and especially during this process. I would like to extend a thank you to my sister, brother in law, my brother and sister in law for the words of encouragement and prayers. Last but not least I would like to thank all my friends for their continued prayers and support.

Julie White

## EFFECTIVENESS OF GOAL SETTING ON TEACHER EVALUATION

Julie White

University of the Incarnate Word, 2020

The purpose of this study was to determine effectiveness of goal setting protocol procedures implemented by three school districts on teacher evaluation scores using the Texas Teacher Effectiveness Evaluation System (TESS). It determined which school's teacher evaluation ratings increased and the effect goal setting had on the increased goals. The evaluation system uses formative and summative goals to support teachers. T-TESS added the goal setting professional development process for all teachers to reflect on their practice and identify needs of improvement. With these goals, teachers should become more effective teachers. The teacher and the administrator develop a plan to reach goals and personal progress toward goals. Then they reflect on these goals throughout the year to track personal growth and the effects on student growth (TEA, 2014). Since the goal setting and professional development process is a new aspect to the appraisal system and their little available data of its effectiveness, research on this topic is warranted.

## TABLE OF CONTENTS

LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
CHAPTER 1: INTRODUCTION .....	1
Context of Topic .....	1
History.....	1
Current Problem.....	4
Purpose of Study .....	5
Research Questions .....	5
Theoretical Framework.....	5
Significance of the Study .....	5
Personal Background .....	6
CHAPTER 2: LITERATURE REVIEW .....	8
Value-Added System .....	9
Advantages of the Value-Added Model .....	10
Implementation of the Value-Added Model.....	11
Disadvantages of the Value-Added Model.....	11
Teacher Manipulation. ....	13
Pay for Performance .....	14
Implementation of Pay for Performance. ....	15
Measurement of Effective Teachers Project.....	18

## Table of Contents—Continued

## CHAPTER 2: LITERATURE REVIEW

Danielson Framework.....	21
Student Placement.....	22
Danielson’s Four Domains .....	23
Goal Setting .....	24
SMART Goals.....	26
Motivation.....	27
Extrinsic Motivation.....	28
Intrinsic Motivation.....	29
Teacher Perception.....	30
Professional Development Appraisal System.....	32
Texas Teacher Appraisal System.....	33
Texas Teacher Evaluation Support System .....	35
CHAPTER 3: METHODOLOGY .....	38
Research Questions.....	38
Population and Sample .....	38
Instrument.....	40
Data Collection .....	40
Data Analysis .....	41
Role of the Researcher .....	41
Reliability.....	42

## Table of Contents—Continued

CHAPTER 4: FINDINGS .....	43
Research Question 1: Which Goal Setting Protocol is More Effective in Increasing Teacher Evaluation Scores Measured Over a 4-Year Time Period? .....	43
Research Question 2: What is the Relationship Between Goal Setting Protocol Satisfaction and Overall Teacher Performance?.....	47
Factor Analysis. ....	47
Correlations.....	50
Research Question 3: What is the Teacher’s Perception of Professional Development Focused on Teacher Goals? .....	51
Summary.....	54
CHAPTER 5: DISCUSSION.....	55
Research Question 1: Which Goal Setting Protocol is More Effective in Increasing Teacher Evaluation Scores Measured Over a 4-Year Time Period? .....	55
Research Question 2: What is the Relationship Between Goal Setting Protocol Satisfaction and Overall Teacher Performance?.....	56
Research Question 3: What is the Teacher’s Perception of Professional Development Focused on Teacher Goals?.....	58
Theoretical Framework.....	60
Professional Recommendations .....	61
Limitations.....	61
Delimitations.....	61
Future Research .....	62
REFERENCES .....	63
APPENDICES .....	68
Appendix A: Invitation to Participate .....	69



Table of Contents—Continued

APPENDICES

Appendix B: Survey.....	71
Appendix C: IRB .....	76

## LIST OF TABLES

Table	Page
Table 1. Descriptive Statistics.....	44
Table 2. Levene’s Test of Equality of Error Variances .....	45
Table 3. Test of Homogeneity of Variances .....	46
Table 4. Teacher Evaluation Scores by Year.....	46
Table 5. Structure Matrix.....	49
Table 6. Descriptive Statistics for Satisfaction of Goal Setting Protocol Items .....	50
Table 7. Relationship Between Goal Setting Protocol Satisfaction and Overall Teacher Performance (2017-2018). .....	51

LIST OF FIGURES

Figure	Page
Figure 1. Evaluation Scores Over 4-Year Time Period.....	47
Figure 2. Cattell's Scree Test.....	48

## **Chapter 1: Introduction**

### **Context of Topic**

Teacher evaluation systems have been under construction for several years. Improved student outcomes have become key elements in these systems (Cohen & Goldhaber, 2016). Many schools have begun investigating teacher effectiveness using numerous measures for improved student achievement (Martínez, Schweig, & Goldschmidt, 2016). Though there has been a consensus that teacher evaluation systems need to reform, a debate on the structure of the evaluations continues (Cohen & Goldhaber, 2016). The Charlotte Danielson Framework has been widely used as a model to create teacher evaluation systems. Using theoretical research, this framework used teacher responsibilities to improve student outcomes (Danielson, 2013). The four domains are planning and preparation, environment, instruction, and professional responsibilities. Assessment in these domains create systems of research-based evaluation methods used throughout the nation.

Texas was under pressure to create a new evaluation system using student growth as one of the evaluation measures (Association of Texas Professional Educators, 2014). Using the Danielson Framework, Texas created the Texas Teacher Evaluation and Support System (T-TESS) which included the following three components: goal setting, the evaluation cycle, and student growth measure (Texas Education Agency, 2016). Creators of evaluation systems are also moving toward using formative and summative goals to support struggling teachers, offering incentives for high performance, providing instructional models, and implementing professional development policies and practices (Sledge & Pazey, 2013).

### **History**

In the 1700s, education was not thought to be a profession, so clergy were the ones to

hire and train teachers because of their extensive studies (Marzano, Frontier, & Livingston, 2011). By the 1800s education gained popularity and clergy soon realized they were not the right fit to supervise. Teachers with the most extensive background in education became the leader or principal of the school (Marzano et al., 2011). Blumberg and Flaherty (1985), stated pedagogical skills were an important component of effective teaching, so instruction needed to be the key focus in order to increase student learning. In the 20th century Dewey and Taylor (as cited in Marzano et al., 2011) brought two different viewpoints to education. With the help of the theorists, administrators began to look at various measures through a scientific approach of schooling (Marzano et al., 2011). Administrators used measures to analyze data to show optimal growth of the student and using administrative feedback to guide teachers' instructional strategies. After World War II there was a shift from the scientific model of teaching to the teacher as an individual and looked at all aspects of the teacher in the role of educator and a person of the community (Marzano et al., 2011). Matthew Whitehead (1952) identified six areas of supervision and noted that observational practices needed much improvement for optimal effective teaching. Whitehead summarized his position by explaining that "administrators should pay more attention to the chief aim of education—effective teaching" (p. 106). It was the recognition of the importance of classroom observation that laid the foundations for one of the most influential movements in supervision.

In 1969, Goldhammar (as cited in Marzano et al., 2011) created five phases of clinical supervision to enhance feedback between the supervisor and the teacher. Much like the model that is currently used today, the five phases consisted of pre-observation conference, classroom observation, analysis, supervision conference, analysis of the analysis. The dialogue that was encouraged in Goldhammar's model ultimately became the demise of the five phases. It became

more of a checklist than an open conversation between the supervisor and teacher. This movement lead into Madeline Hunter's lesson plan cycle. Using her model, supervisors would determine the effectiveness of the teacher through preconference, observation, and post conference using the lesson plan cycle as a guide.

The mid-eighties theorist William Glathorn (as cited in Marzano et al., 2011) used a supervisory model that highlighted career goals of teachers. He felt teachers' input was an important aspect of their development. Professional development was provided for teachers based on their individual needs. Thomas McGreal (as cited in Marzano et al., 2011) used this same idea to place teachers either in intense training if they showed significant instructional deficiencies or moved teachers to tenure. Glickman, Gordon, and Ross-Gordon (2001) stated that teachers' optimal growth occurs when the goals of the school and the needs of the teacher are targeted through a systemic approach to the supervisory process. The RAND (Marzano et al., 2011) group studied the supervisory practices in varies school in America. The study found the evaluation process to be dedicative and left little room for the teachers input to enhance the pedagogical development (Marzano et al., 2011).

This brings us to the Danielson model which, unlike the other models, sought to capture the full capacity of teaching. The framework's purpose was to distinguish a universal language for professional conversations using self-assessment and reflection to enhance professional practices. From planning to reporting achievement, it was said to encompass all the components of teaching yet is flexible enough to be used with any content or grade level. Using this model there is a clear shift from supervision to evaluation (Danielson, 2013).

There has also been a shift from focusing on teacher behaviors to student achievement. Studies showed there was a clear link between teacher effectiveness and students gains, therefor

students' performance was essential to teacher overall achievement. Though another report by Toch and Rothman (2008) determined new evaluation systems do not address quality of instruction nor do they measure student learning. The No Child Left Behind requirements only focused on teaching credentials and not teacher quality. In the Widget effect study (Weisberg, Sexton, Mulhern, & Keeling, 2009) it is assumed all teachers are the same. It further stated it does not look at individual needs of the teacher and using their strengths and weaknesses. Without looking a differentiation in the needs of teachers, the evaluation systems fail to identify specific developmental needs of teachers.

### **Current Problem**

Many of the current models include a value-added aspect, where teachers are evaluated on the students' performance. While teacher effectiveness does seem to impact students' scores, there is little evidence of how new models support teacher growth (Derrington, 2016). By providing teacher feedback after observations and identifying needs of growth, teachers are given the tools to become effective teachers (Derrington, 2016). It is the responsibility of both principals and teachers to collaborate to define the strategies that make these new evaluation systems work (Derrington, 2016).

T-TESS has added the goal setting professional development process for all teachers to reflect on their practice and identify needs of improvement. The teacher and the administer develop a plan to reach goals and personal progress toward goals. The teacher and administrator reflect on these goals throughout the year to track personal growth and the effects on student growth (TEA, 2016). Goal setting and the professional development process is a new aspect to the appraisal system and there was little available data of its effectiveness.

## **Purpose of Study**

The purpose of this study was to determine effectiveness of goal setting protocol procedures on teacher evaluation scores implemented by three school districts using the T-TESS.

## **Research Questions**

RQ1: Which goal setting protocol was more effective in increasing teacher evaluation scores measures over a 4-year time period?

RQ2: Which of the three school districts' goal setting protocols was most effective?

RQ3: What was the relationship between goal setting protocol satisfaction and overall teacher performance?

## **Theoretical Framework**

The purpose of this study was to understand how goal setting effects teacher evaluations using the T-TESS. The goal setting theory uses feedback targeted to specific goals which will increase task performance (Locke & Latham, 2006). The essential features of goal setting theory are working towards a goal is essential to the job at hand, using clear and focused goals are greater motivators and will achieve better performance, goals are realistic and measurable, and using appropriate feedback will increase overall performance and job satisfaction.

## **Significance of the Study**

The findings in this study showed the most effective goal setting protocol and how goal setting impacted teacher effectiveness. Individuals' self-efficacy is important to how a person completes a task. The greater the confidence, the higher the effort put forth to complete a challenging task.

This study will identify the importance of self-efficacy on obtaining goals. It will also show the growth of teachers over time, using the goal setting protocol and the process teachers



and administrators use to modify and adjust goals throughout the process to achieve higher teacher evaluation scores. The study will give the administration a deeper understanding of the most effective way to set and track teacher goal setting and professional development to increase teacher quality.

### **Personal Background**

As a teacher and administrator, I had the unique opportunity to work with several different appraisal systems. While most of Texas was using the Professional Development Appraisal System, my school took part in the TAP System for the Teacher Student Advancement program put in motion by the Milken family in 1999. The purpose was to improve teacher quality through four domains: Teaching Skills, Knowledge, and Responsibilities Performance Standards Overview, instruction, design and planning, the learning environment. We were compensated for student performance using the value-added model. This was a little tricky for teachers who did not teach grade levels with end of year testing. They were compensated a set value depending on the whole school's overall scores. Though the extra money at Christmas was a good bonus, overall it did not affect my teaching. The students were the ultimate reason for wanting to increase scores.

TAP also required grade level teachers to meet weekly to discuss best practices through professional development. The items that were discussed were upcoming classroom lessons and strategies to implement those topics. Most of the time the meetings were relevant to classroom needs because we discussed data and upcoming lessons weekly.

The next school district I joined used the old PDAS system. What's PDAS? Teachers were evaluated once or twice a year and received some feedback through an email. If there was a concern, the principal would conference with the teacher. Teachers were provided a stipend for

student achievement however did not set yearly goals. Professional development was not targeted on teacher's identified needs, but on what the district felt was needed. Staff meetings were boring and ambiguous most of the time. Teachers stated their time was not valued and there was no true learning involved.

When I became an assistant principal, my campus piloted the T-TESS. The principal and I divided the teachers into two groups. Each of us met individually with the teachers in our group and created individualized professional development plans based on their professional goals which aligned with the appraisal rubric. After setting the goals, it was our intention to meet with the teachers' midyear to adjust their goals. Though we did have biweekly individualized meeting with each teacher, we neglected to review their goals and check for progress. Goal setting seemed like a good idea but with all the observations, evaluations, meetings, and budgeting, it was nearly impossible to follow up and follow through with everyone's goals.

## Chapter 2: Literature Review

Teacher evaluation systems have been under review in the last few decades. The previous evaluation systems relied heavily on a checklist system and with little emphasis on teacher growth or retention (Steinberg & Kraft, 2017). In 2001, the No Child Left Behind Act asked school districts to create a system for ensuring teacher preparation and accountability systems were in place for high academic learning. Policyholders created various evaluation systems to support the most effective learning environment for all students (Steinberg & Kraft, 2017). The new evaluation systems focused on using a multiple measure performance rating, professional support, and incentive pay structures.

Melinda and Bill Gates started the Measurement of Effective Teachers project that used the value-added model to show teacher effectiveness using student assessment scores (Kane, McCaffery, Miller, & Staiger, 2012). Debate continues over the value-added system effectiveness. Policy makers have also used the pay for performance model using incentives to motivate teachers through monetary incentives (Mintrop, Ordenes, Coghlan, Pryor, & Madero, 2017). There is little evidence to support their continued use in the school system (Mintrop et al., 2017).

The framework that has shown to be the most effective is the Framework for Teaching created by Charlotte Danielson (2010). This framework uses multiple domains to observe teacher effectiveness. In 2013, Texas adopted an evaluation system based on the Danielson model called the T-TESS. An added element to the new evaluation system is goal setting. Goal setting was created to become an open conversation between teachers and administration. Zimmerman (2006) stated that goal setting reinforces motivation and student growth. Although goal setting has shown to be effective in school systems, the T-TESS does not give guidelines for

administrators to follow in order for the system to be the most effective for teacher and student growth. Decisions are left to the districts to determine how to set goals with teachers to obtain the most student academic growth. Without clear guidelines, there is not an effective way to show which goal setting protocol is the most effective. A goal setting protocol would allow stakeholders to monitor the effectiveness of the T-TESS. The background of the teacher evaluation systems needs to be laid out in order to understand goal setting in the new evaluation system.

This literature review will give an overview of the creation of the T-TESS and why other systems have helped guide the creation of the program. It will detail the programs that have had a positive and negative impact on the evaluation system. The literature review will also present the goal setting category and its implementation in the new system. A progression of systems is presented to show the overall development of the evaluation system and how each system is used to evaluate teachers' effectiveness.

### **Value-Added System**

The popularity of a value-added system to determine teacher effectiveness began in 1996, primarily due to the research completed by Sanders and Rivers (1996). The value-added system was created to use student test scores to measure teacher effectiveness (Haertel, 2012). Students are tested at the end of the year and then averaged together as a class. The teachers are then compared to each other based on the classroom averages (Haertel, 2012). Haertel (2012) indicated that the system involved more complex formulas with assumptions and violations that make it difficult to actually determine an accurate value-added score. For instance, analyzing the Tennessee Value-Added Assessment System and the relationship between teacher effectiveness and student growth, Sanders and Rivers (1996) collected data from mathematics classes in grades

three to five in two metropolitan areas. The study consisted of four phases. In the first phase of the study, over a 3-year period student scores were tracked, and teacher effectiveness was determined using a longitudinal analysis. Using the students' scores and the teacher effectiveness ratings, the researchers were able to track student progress from year to year and determine the effects of the previous teacher on student test scores and growth. In the second phase, the three grade levels split into quintiles with a possibility of 125 combinations. The results showed that students receiving instruction from an ineffective teacher one year show significant gains if they have an effective teacher the following year. However, there are also lingering effects of the ineffective teachers in subsequent years. In phase three the evaluation of the quintiles noted that overall, the low-achieving students made more gains than the high-achieving students. In phase four, the focus was on minority groups. When there were complaints about ineffective teachers, the administration tended to move these teachers to schools within the school district with high minority rates. The schools with a higher minority population demonstrated a higher rate of discipline referrals and lower funding for classroom materials. Essentially, the ineffective teachers were moved to poorer school districts. The results showed that even though this was the case, the scores of the ethnic groups were comparable to each other. Overall, the study showed that the one factor that had the most impact on student growth as teacher effectiveness.

### **Advantages of the Value-Added Model**

Administration can identify teachers' strengths and weaknesses and provide appropriate professional development to increase student growth using the value-added model. In 2001, the value-added model became a common practice with the introduction of the No Child Left Behind Act, a policy to close achievement gaps. Using the value-added model would help schools to use student scores and teacher evaluations over an extended period to determine teacher

effectiveness and close student achievement gaps (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004). Closing the achievement gap is the ultimate goal of all stakeholders, however using a one-time test score may not show the teacher's total effectiveness.

### **Implementation of the Value-Added Model**

Though the value-added system does not use a single snapshot date for student achievement, the it was relatively confusing to implement. A value-added model was to replace the snapshot, which does not take into consideration mobility and socioeconomic status of student test scores to determine the effectiveness of a teacher (Chambers & Tate, 2013). Timmermans, Snijders, and Bosker (2012) stated that the value-added model looks at the range of dates students were enrolled in a school and compared them to other students in the same timeframe at other schools. Geiger (2018) stated that measured growth can be summed up to the teacher's effectiveness on students' growth and achievement over time. He further stated that value-added has shown most accurate when using 3-years of data; where many schools are only using one year to terminate teachers. Koedel and Betts (2009) added that using several years' data showed a more accurate growth of teachers, but many school districts' policies only address a single year of data. Furthermore, Geiger (2016) stated that the value- added model would be more effective if attached to other measures. Teachers can use portfolios to track student performance and teacher learning goals to improve overall teacher performance in the value-added model (Haertel, 2012). Using this data would be all-encompassing of the student diversity in each classroom from one year to the next instead of using one snapshot date.

### **Disadvantages of the Value-Added Model**

As of 2015, 30 states used a value-added practice in their teacher evaluation process (Darling-Hammond, 2015). Several researchers have indicated a variety of ways the value-added

model was proven unreliable (Amrein-Beardsley & Holloway, 2017; Ballou & Springer, 2015; Darling-Hammond, 2015). Using a normal distribution for all teachers does not show if one teacher is effective or not, but instead if they are better or worse compared to other teachers.

Value-added models come with systematic errors leaving fluctuating variables such as student placement, socioeconomic, and race which can affect the testing results. Within the value-added model, students' scores are compared to previous years' test takers, and teachers are evaluated using the scores of the student which falls into a particular distribution. It is not taken into consideration the number of students in the class, the teacher's actual effectiveness through progress monitoring students and teacher evaluation scores, or yearly academic student performance. A teacher with 100 students is evaluated the same as a teacher with 26 students. The value-added model also assumes that learning is linear, and all students learn at the same rate, and because of this, expected student growth cannot be measured appropriately (Darling-Hammond, 2015).

Some experts are skeptical of the value-added system being useful as an indicator of teacher effectiveness (Darling-Hammond, 2015; Goldhaber, 2015; Hill, 2009). Hill (2009) indicated that observational research shows there is little evidence linking the value-added model with teacher quality. Hill (2009) argued student test scores could be because of former teachers and bring an error in measurement because of the one-time student assessment, and students' scores have little variance. Hill also indicated in a survey using low and high-stakes testing, the value-added model showed small amounts of evidence to determine teacher effectiveness. Darling-Hammond (2015) stated because of the No Child Left Behind Act, assessments are not at their level, but only on grade-level standards, students above and below grade level cannot show growth through their assessments. These tests do not take into consideration other variables

such as English language learners and special population students. Not only does the student population need to be taken into consideration, but so does student attendance, health, culture, student gains, and losses. Using these variables, Darling-Hammond (2015) argued that teachers only hold about 10% of variation on student achievement. Goldhaber (2015) stated outcomes from documented studies identified the unpredictability of students in each class year to year and from test to test. Some tests showed students' scores were in the top 20% one year and fluctuated to the bottom 15% the following year. These statistics show there is variation in student scores and teaching ability.

**Teacher manipulation.** Teachers have consciously and unconsciously manipulated the system to their advantage (Haertel, 2012). Haertel stated because teacher effectiveness is based on student achievement using the value-added model, teachers have been known to use the system to their advantage. The value-added model was used to spread ratings of teachers evenly and to dismiss ineffective teachers. Because the standardized tests show what the student knows on grade level and does not monitor what they have learned on the individual student's level within the year regarding depth and complexity, teachers with struggling students do not have accurate data to show gains of these students (Amrein-Beardsley & Holloway, 2017). In one study it was noted that teachers were manipulating the test rosters leaving out low-performing students (Amrein-Beardsley & Holloway, 2017). Research showed that students test better with their classroom teacher (Ballou & Springer, 2015). The value-added model does not take into consideration the concept which teachers test their student and are subject to cheat in many ways



such as: changing students' answers, teaching students codes, or unconsciously teaching strategies for the day of the test (Ballou & Springer, 2015). The students may adversely affect other students in their classroom academically and socially, which could increase or decrease performance (Amrein-Beardsley & Holloway, 2017). All of these distractors make it challenging to use student test scores to determine teacher effectiveness or determine teacher growth accurately.

Using the current value-added model, teachers also work against each other in order to have a higher value-added score. Instead of sharing ideas with other teachers, Haertel (2012) stated teachers were keeping lesson plans and materials for themselves. Within the system, teachers are given labels of effective or ineffective and ranked against each other. Since there is an equal distribution of teachers, there have to be ineffective teachers among all the teachers.

This could essentially lead some effective teachers to be labeled as ineffective and therefore create animosity against other teachers. On the other hand, schools with only ineffective teachers must label some of them effective due to the distribution of ranking (Haertel, 2012). Student growth does not progress when the teacher and administration work against each other.

### **Pay for Performance**

As with the debate on the value-added model, there is also a debate over incentive pay or pay for performance. In 2008 the federal government created the Teacher Incentive Funds (TIF) initiative, to use the value-added model to compensate teachers for their overall classroom evaluation and student test scores (Mintrop et al., 2017). Mintrop et al. (2017) conducted a study from 2011-2014 focusing on three secondary schools using the TIF initiative. The study interviewed 52 teachers and 15 administrators over the three-year time period. School districts

used several indicators to classify teachers into different teaching levels and rewarded those teachers with incentives according to overall evaluation, student test scores, and participation in school functions. The idea behind the TIF initiative and other teacher compensation models, was to use extrinsic motivation to promote teacher instruction such as, motivating teachers to take on leadership roles, bring in stronger teachers, dismiss ineffective teachers, and increase student performance (Rice, Malen, Jackson, & Hoyer, 2015).

Using the expectancy theory infused with the goal-setting theory, Rice et al. (2015) studied the teachers' view of the payout distribution. The expectancy theory uses valiance, expectancy, and instrumentality as factors that lead to motivation. Using this theory and the goal setting theory, which establishes reasonable and attainable goals to increase motivation, Rice et al. (2015) stated that teachers were motivated intrinsically by the amount of the merit, and must be attainable within a reasonable timeline using fair and just measures.

**Implementation of pay for performance.** Pay for performance has been implemented in both educational and non-educational settings. For educational settings, pay for performance depends on teachers' buy-in which is an important factor. When introducing differentiated salaries, teacher buy-in is lower because they are uneasy about unfair compensation (Rice et al., 2015). A study in Texas found that teachers who received little or no money in addition to their base salary were more likely to leave the profession or move jobs over teachers who were well compensated at least 4% of their base salary (Rice et al., 2015). This comes with some caution as large compensations could lead to undesirable behaviors such as cheating (Rice et al., 2015). Rice et al. claimed the largest payout came to teachers who took hard-to-staff positions and however based on student achievement was the smallest payout. Though monetary incentives influence where teachers work or if they remain in the profession, studies show that teachers are

more motivated by increasing students' performance.

Public sector pays for performance. Pay for performance is not only used in school systems but gained popularity in the public sector as well. Glassman, Glassman, Champagne, and Zugelder (2010), stated experts were at odds about the effectiveness of money for motivation. In their findings, there was a 44% increase in production when a glass company switched from a salary-based pay to a pay for performance pay system. The study showed this was due to inadequate employees leaving because they were unable to meet the goals required to receive payment for their performance and also there was an influx of productivity due to the adequate workers speeding up production to receive an increase in pay. Process theorists believe pay can be an external motivation as long as employees feel payment is fair and within a timely manner after the goal is achieved (Glassman et al., 2010). Mintrop et al. (2017) stated the sales industry uses clear job expectation to motivate employees to achieve a bonus payout externally.

Pay for performance systems are also known to fail because of poorly managed systems (Glassman et al., 2010). Performance can increase for monetary reasons only in simple task situations, but for complex tasks, the effect of these programs is the adverse effect (Kappan, 2010). However, Glassman et al. (2010) noted pay for performance can be effective if the merit is an increase of 5% in their salary and be lucrative to the employee (Glassman et al., 2010).

With these compensation increases, employees should see the merit pay funds within a reasonable timeframe, if not productivity decreases. Glassman et al. (2010) give the example of a nurse receiving incentive pay to attend to a patient's needs within a specific timeframe. Meeting the time, it did not matter to the patient, which makes the merit pay not equal to the outcome.

Public services have used pay for performance to emphasize extrinsic motivation and value of career and status, but there is a disconnect due to Herzberg's theory of motivation

(Mintrop et al., 2017). In Herzberg's theory for motivation (Robbins & Judge, 2015) money does not meet a basic need, and therefore cannot be used as a motivation for an individual unless the basic needs are not with the individual's salary. A relationship can be established between pay and motivation if there is a need that must be met (Glassman et al., 2010). Hertzberg further suggested that money will only relieve dissatisfaction, and ultimately, it is an intrinsic motivation that motivates an individual's desire to work and fill a need (Robbins & Judge, 2015).

***Advantages of pay for performance.*** When the pay for performance plan is inclusive, there is support for the program by all stakeholders. Teachers feel that monetary awards are more attainable when the standards are laid out in a comprehensible manner (Rice et al., 2012). Rice et al. (2012) also stated teachers also felt that tying professional development to the incentive program gave them an overall positive correlation to the pay for performance program. For a successful pay for performance program to work there needs to be significant funds at the district level. With the supportive staff to complete the observations and the financial backing, teachers feel the pay for performance programs promote teacher growth (Rice et al., 2102).

***Disadvantages of pay for performance.*** Recent studies have highlighted why pay for performance is not successful within the educational field. Mintrop et al. (2017) stated that intrinsic motivation and overtly complex tasks make it challenging to create a well-defined merit system in the education profession. With the implementation of incentive programs, increasing extrinsic motivation can be detrimental to long-term intrinsic motivation (Hulleman & Barron, 2010). Glassman et al., (2010) also stated that pay for performance could decrease the motivation to do what is good for the company and one's overall intrinsic motivation and focus more on completing the task "no matter what" for the money. Another reason this system may fail is due to the nature or task of the job (Glassman et al., 2010).

Understanding how to use valid data to connect to pay and the reliability of the process is a complicated process (Mintrop et al., 2017). According to Hulleman and Barron (2010), former President Obama created the Race to the Top American Recovery and Reinvestment Act in 2009. The purpose of the program was to pay highly effective teachers and principals with a \$4.35 million grant. This followed the footsteps of the Teacher Incentive Funds, which paid the teacher for performance in high needs school (Liang & Akiba, 2015).

Despite the rising interest in these programs, literature contradicts the success of such programs (Liang & Akiba, 2015). Furthermore, in some instances, it led teachers to use the system to their advantage instead of teaching the necessary skills to increase student comprehension (Liang & Akiba, 2015). Some studies have identified success in the pay for performance on student achievement, but there is an uncertainty if there is actual student achievement gain or are teachers teaching to the test (Liang & Akiba, 2015).

Yuan et al. (2013) stated in order to motivate students to increase scores on standardized tests, teachers must be motivated to make changes. The study looked at three different incentive pay programs. The first one was to change the way the teacher taught, the second was to build teams, and the last program was to attract highly qualified teachers through incentive pay. As a result, each of these programs increased student performance (Yuan et al., 2013). In two studies conducted in Kenya and India, teachers' practice had no significant difference. In the United States, the results were mixed, but the teachers were also participating in assessment-based reforms which could add to the skewed results (Yuan et al., 2013).

### **Measurement of Effective Teachers Project**

Both the value added and the pay for performance models have indicated that student growth and teacher effectiveness can be achieved using a multiple-measure teacher evaluation

model. Bill and Melinda Gates created the Measure of Effective Teaching (MET) project in order to look at the components of teacher effectiveness (Kane et al., 2012). The MET project measured effectiveness under three other premises; should student achievement be linked to teacher evaluation, should classroom observations and student feedback be related to student gains, and feedback should support teacher growth and development (Kane et al., 2012). The first idea looks at if the value-added model should be used to determine student achievement. Researchers suggest many factors can lead to student achievement such as: teacher inspiration, parenting, and prior learning (Jilliam, Tocci, Holtzman, & Williams, 2013). Therefore, to only look at the current grade level achievement is not necessarily a fair assessment of the students' growth.

The MET project collected data from grades 4 through 8 in English and Mathematics and the high school end of course students taking algebra, biology, and English. The project used five measures to determine the effectiveness of teachers (Kane et al., 2012). The first measure included multiple tests given to the students and reviewed for growth including the state recommended test. In addition to showing achievement gains and losses through testing, the second measure focused on classroom observation and teacher reflections. Teacher raters were trained through a training program provided by the Educational Testing Services (Kane et al., 2012). The raters were trained to score lessons using the Classroom Assessment Scoring System, Framework for Teaching, Mathematical Quality of Instruction, Protocol for Language Arts Teaching Observations, and Quality of Science Teaching. The third measure evaluated teacher knowledge of content and understanding of student development. In the fourth measurement, students gave feedback on their overall classroom experience. The final measure completed a survey about the teachers' overall experience including student environment, instructional

support, and working conditions (Kane et al., 2012).

The MET study also looked at various components to add in order to measure teacher effectiveness. The project used student perception, classroom observations, assessments of teacher knowledge using the value-added model (Kane et al., 2012). One way to minimize error while measuring teacher effectiveness was to place students in classrooms randomly, but the researcher could not determine if there was bias in the value-added scores. The second idea was to add domains to the teacher evaluation and not a simple checklist for the principals to mark through (Kane et al., 2012). The last component of the MET project used feedback to support teacher growth. The high achieving teachers used the feedback to become teachers that are more effective, and the low performing teachers did not use the feedback to grow as a professional. Teachers needed to have attainable goals using feedback to create the most student growth capitalizing on the teacher's strengths and weaknesses (Kane et al., 2012).

The results of the MET project indicated student achievement had a direct correlation to teachers' previous year teaching using the value-added model. Although there is variance from class to class through the years, the value-added model still showed to be a predictor of future learning achievements (Mihaly, McCaffrey, Staiger, & Lockwood, 2013). Test scores indicated math scores were more influential than English language arts. Also, the value-added model promoted deeper conceptual understanding. The project also showed students know a teacher's strengths and weaknesses in different content areas (Mihaly et al., 2013). The overall conclusion of the project stated there are still many more bridges to cross, but the initial findings showed there was a direct correlation between students' achievement and students' feedback on teacher performance. According to Mihaly et al. (2013), evaluation systems need to use the teacher's strengths and weaknesses to create goals for student achievement. To create better evaluations

systems Jilliam et al. (2013) suggested training principals on how to accurately use the value-added model with the more in-depth understanding of each student and using data-driven instruction to enhance classroom performance.

### **Danielson Framework**

The most commonly adopted model for evaluating teacher effectiveness is the Charlotte Danielson Framework for Teaching (Nielsen, 2014). This framework has shown composite scoring is more reliable and able to identify specific strengths and weaknesses of individual teacher practice in order to increase teacher effectiveness (Kettler & Reddy, 2007). In 2009, The Gates Foundation decided to use the Danielson framework when looking for a measurement tool to observe teacher effectiveness. The purpose of Kettler and Reddy's (2107) study was to understand which elements of teacher practice was closely related to students' achievement. The framework included a much-needed rubric with academic language to evaluate each component. It was not to be used as a checklist like other evaluation systems in the past, but rather a guide for teacher practice (Danielson, 2010). By using these enhancements, administrators and teachers were able to determine a level of performance on an objective level and accurately rate each teacher (Danielson, 2010). The framework was used to enhance teacher effectiveness by equipping teachers with the correct tools to be successful. The Danielson framework created an overall composite score for a multidimensional evaluation system (Mihaly et al., 2013). Understanding how to use the composite score when evaluating teachers is a complicated process. There is not one formula or exact number that is universal for all composite scores. The composite score is used to determine teacher retention, compensation, and professional development goals (Mihaly et al., 2013). Each stakeholder has a different opinion on teacher quality, so it is essential to determine the exact components of an effective teacher. Kimball and



Milanowski (2009), disagreed on the validity of the composite score due to the fact teaching is not as simplistic and personal opinion of what effective teaching comes into play.

The reliability of the composite scores using the Danielson framework studied by Kettler and Reddy (2017). A convenience sample was taken from the United States Department of Education including 12 charter schools with 156 teachers. Within their study, the composite scores showed more teacher growth over time than traditional models. Kimball and Milanowski (2009) also did a study consisting of one school district with 88 schools and 3,300 teachers that used the Danielson framework for three years. The school district was unsatisfied with the standard evaluation system and wanted something with more components to guide teacher growth. Their findings were that there was variation between evaluators and a relationship between rating and student achievement. Teacher growth occurred when the teacher and principal established a relationship and set goals together. Overall, this study suggested that teachers' ratings did appear to differ depending on the evaluator. This could be due to how the evaluator uses motivation. Leniency used when an evaluator wants to retain certain employees or vice versa (Kimball & Milanowski, 2009). In another study by Borman and Kimball (2005) 400 teachers teaching grades 4-6 with about 7,000 students in the Washoe County investigated the distribution of students and achievement of teachers. The study found that teacher quality was not distributed evenly among classrooms. Lower-achieving teachers received the minority students. Higher evaluated teachers received the higher achieving students in their classrooms.

### **Student Placement**

Garrett & Steinberg (2015) mimicked the MET study in six school districts over a two-year period. The sample consisted of 834 teachers from grades 4-9 teaching math and English language arts. There was a correlation in composite scores and teacher growth using the

Danielson framework, but limited research of the placement of students and overall teacher performance. Teachers that were assigned the top students were more likely to receive a higher evaluation score than the teachers who had the low performing students. Garrett and Steinberg (2015) stated the framework was to be used to retain or dismiss teachers because there is a correlation in teacher performance and student achievement, however, this study recognized the Danielson framework does not take into consideration the non-randomization of student placement. Their study further discusses the other limitations of the framework such as the compositions of students in a classroom, behaviors, and students overall educational needs. They also noted that when standardized test scores or the value-added model are used to reflect on teacher effectiveness, it does not take into consideration the students assigned to the teacher and previous performance. Because of the many components that go into teacher effectiveness, there should be other measures to capture teacher performance over time instead of solely relying on one instrument.

The concept behind the Danielson framework was to implement an evaluation system across all grade levels and content areas. Danielson and McGreal (2000) discussed the evaluation system and its components. The Framework for Teaching system uses four domains which include rubrics with descriptions and possible teacher strategies to evaluate teacher effectiveness. The four domains consist of the following: a) planning and preparation, b) classroom environment, c) instruction and d) professional responsibilities totaling 66 elements.

### **Danielson's Four Domains**

In the planning and preparation domain, teachers are to create lessons using their expertise in their content area and to focus on instructional strategies based upon state standards (Danielson & McGreal, 2000). In the second domain, classroom environment, teachers are to

create an organized classroom environment to foster learning for all learners. In the instruction domain, teachers create lessons that engage all learners through learning strategies where students take the initiative for their growth. In the fourth domain, teachers are to maintain an ethical standard and take part in community involvement activities. Through these four domains, teachers can obtain one of four performance levels depending on classroom observations: distinguished, proficient, necessary, and unsatisfactory. To determine teachers' ranking within these performance levels a rubric based on their strengths and weaknesses within each domain. They are required to provide a portfolio of evidence, which can be in the form a binder or digital format, for their teaching in each domain and which strategies they used to become a highly effective teacher. In the portfolio, teachers will also document professional development and any mentor teacher observation completed within the year. Tenured teachers are required to submit a self-evaluation and evidence of professional development they completed throughout the year.

### **Goal Setting**

The state of Texas has included the goal setting measure to improve students' achievement and teacher effectiveness. The purpose of goal setting is for the teachers to reflect on their growth using the components of the evaluation system and review the goals with their administrators to create a professional development plan (Texas Education Agency, 2011). Locke and Latham (2006) stated that setting specific, attainable goals over vague goals creates higher work performance. They further state that employees are more satisfied in the workplace when charged with meeting challenging goals. Goals can motivate people to acquire new skills or recall existing abilities within one's self. According to Bandura's (1977) theory of stated that self- efficacy, self-satisfaction is achieved by setting and obtaining personal goals. Achievement goals give individuals ways to understand how they learn or process to gain knowledge and

master goals (Poortvliet & Darnon, 2010).

Studies of the goal setting theory evaluated in over 88 different tasks, with more than 40,000 participants in four continents in both laboratory and field settings (Locke & Latham, 2006). The studies have been tested in one minute to 25-year time spans using individuals and groups. They found goal setting was not effective when using monetary incentives, but when the group and individual goals were aligned, there was more personal growth (Locke & Latham, 2006). Along the same lines, when corporations and employees shared a vision, dysfunction decreased, and productivity increased. The study also discovered that setting goals using past experiences could create self-efficacy and setting higher specific goals can lead to higher achievement unless the goals are unattainable, and then work performance decreased. Bandura's social cognitive theory uses self-efficacy to motivate an individual to monitor and adjust their learning according to failures and gains in past experiences (Erikson, 2002). When an individual has a high self-efficacy their motivation to reach a goal is increased, and the individual becomes more challenged by the goal (Locke & Latham, 2006).

In the educational setting using challenging but attainable goals increase motivation and achievement (Locke & Latham, 2006). Morisano, Hirsh, Peterson, Pihl, and Shore (2010) investigated a goal-setting program for struggling college students. Students who wrote detailed strategies to obtain their goal over a four-month period had a higher academic achievement.

Academia also uses goal setting when planning curriculum. The backward by design model uses goal setting to create lessons and curriculum. Educators use the concepts of backward by design by focusing on the end of the course goals using standards the students have already learned and new standards and creating learning goals for the students to achieve. Teachers using the backward by design approach and creating learning goals had greater overall

student success. More so, students who set their own goals were more motivated to improve their learning (Zimmerman, 1990).

Goal setting has shown to be beneficial, but it requires diligent structure and guidance. Locke and Latham (2006) looked at the history of goal setting and discovered there were, in fact, several components which made goal setting successful. First goal complexity needed to be challenging enough to create the continued focus on the goal. Creating specific goals that addressed specific issues lead to more goal achievement. The goals need to have energy and a direction. When individuals can set their timeline to complete the task, they tend to work harder to complete it. As individuals work through these goals in order to be the most effective, there needs to be a time for feedback to refocus a goal and adjust accordingly. Because these goals are complex the knowledge to complete these tasks is not acquired yet, and with obtaining new information obtained there is a higher probability of meeting the goal.

Latham, Stajkovic, & Locke (2010) used the motivation hub to explain how individuals reach goals that included action, consistent personal goals, commitment, and self-efficacy. Individuals can obtain demanding goals by setting the timeline, continual feedback, and goal adjustment.

**SMART goals.** Educators can create a more effective classroom and show student growth by creating SMART goals: specific, measurable, attainable, real, timeline (Zimmerman, 2006). Zimmerman (2006) reviewed Bandura's (1986) social cognitive theory and Locke and Latham's (2006) goal setting theory that when teachers set goals with their students, the students showed more persistence and worked towards those goals with more attention. He found that by using the goal setting theory and the social cognitive theory, there are criteria that lead to obtaining goals. The six criteria included: specific goals, using a timeframe, hierarchically

organized, goals are consistent with self and others, feedback and reflection, and type of goal processor performance. Moeller, Theiler, and Wu (2011) showed evidence of goal setting success in the classroom through a study they performed using a purposeful sampling of teachers over several years from 2005-2009. There were 21 teachers and 1,273 students in the study. The evidence concluded that there was an increase over time of the mean goal, action plan, and reflection scores. The study also showed there was growth in goal setting and growth in proficiency. Using SMART goals to progress monitor showed evidence of student achievement. T-TESS is using this data to create a new criterion to help with teacher growth.

**Motivation.** The debate over teacher evaluation systems comes down to how to motivate teachers to be the most effective for student growth. The pay for performance and the value-added measure use extrinsic motivation to boost teacher effectiveness, while goal setting and professional development use intrinsic motivation. Teacher motivation is an essential component of teacher retention. Finnigan (2011) examined a low performing elementary school interviewing teachers between the years 1999 to 2001. In the study, Finnigan (2011) studied what factors motivated teachers. Cultivating a trusting relationship with the principal was one of the factors teachers associated with their motivation. When principals did not micromanage teachers they were able to build these trusting relationships. Principals who gave teachers the freedom to take a risk and use innovative ideas in the classroom fostered the trusting relationship that motivated teachers to grow as professionals.

Motivation is inspired by three factors: expectancy, value, and effective components according to Thoonen, Slegers, Oort, Peetsma, and Geijsel (2011). When teachers feel their beliefs are valued and can create self-efficacy, they tend to show higher levels of classroom performance and are open to new ways of thinking. Teachers who set goals and acquire new

knowledge to reach these goals are more committed to completing a difficult task. If teachers' personal goals are valued, teachers tend to be more supportive and motivated to accept the goals of the organization and adapt them as their personal goals. Creating and accepting personal and organizational goals directly affect teacher self-efficacy. Robbins and Judge (2015) stated that using the self-determination theory individuals feel control over their work and are more likely to finish the task. The cognitive evaluation theory states that using extrinsic rewards diminish intrinsic motivation (Robbins & Judge. 2015). According to this theory when individuals are paid to do a job, they do not feel autonomy and are less likely to want to complete the task. Using Maslow's hierarchy of needs, an individual need to move through five levels of needs to become motivated (Robbins & Judge. 2015 ). The lower needs such as physiological and safety can be extrinsically satisfied with money, and other rewards whereas the higher needs social, esteem, and self-actualization are satisfied within the person (Robbins & Judge. 2015).

Creating a safe place for sharing new ideas cultivates a nurturing professional environment. Thoonen et al. (2011) stated when teachers can create interactions with other professionals in a trusting environment, they were able to transform as teachers. These relationships gave the teachers emotional and psychological support needed to feel motivated in a stressful situation.

**Extrinsic motivation.** A variety of extrinsic motivations used in the educational setting. Firestone (2014) discussed several of these motivators which include: the principal-agent theory, career ladder, recruit and retain through pay, merit pay, bonus pay, and removing poor teachers. Using the principal-agent theory rewards are offered when goals met under specific conditions set forth by the authority. Schools have also used the career ladder approach to incentivize teachers to take on more prominent roles outside the classroom. Through the merit pay system,

teachers are rewarded either for student growth measures or teacher practice through evaluations. Springer (2009) stated the idea behind these incentives is to motivate teachers to put in the effort for the reward. Merit pay was also used to retain teachers in the profession because when teachers are not able to live off their monthly earnings, they are discouraged and move to other careers.

Although there are many positives to extrinsic motivators, there is also the downside to them. The system for creating a valid pay for performance is so critical and has yet to be designed without fundamental flaws (Baker, 2010). Furthermore, untested subjects are not easily able to allocate the proper value added. Firestone (2014) stated that students' scores are asked to measure items such as problem-solving skills that the standardized test does not measure, so there is a mismatch in the criteria in the assessment. A study conducted by the National Center on Performance Incentives involved a school in New York and another in Texas, demonstrated there was no significant evidence that the pay for performance system worked. Teachers felt the incentives did not motivate them to put in any more effort into their jobs (Marsh & Martin, 2011).

**Intrinsic motivation.** Intrinsically motivated individuals create challenges or goals to reward themselves. Firestone (2014) studied two motivation theories. He focused on an economist-based theory using extrinsic motivation and a psychology-based theory using intrinsic motivation. The intrinsic study showed that in order for an individual to take the initiative of their learning, they must have autonomy. Teachers have stated that even with the extrinsic motivations they feel more successful when they are guided to reach goals through administration feedback. The study showed teachers using the feedback could create intrinsic goals to guide their teaching in a clear direction.



The challenges of intrinsic motivation though seem to be less than extrinsic motivation they still do exist. Though teachers are receiving the adequate feedback needed to motivate them to continue teaching, the supplies need to teach is not there and teachers must use their own money to buy supplies (Firestone, 2014). When principals and teachers create effective goals together, it is hard to find appropriate professional development for teachers to attend. Firestone (2014) stated sometimes teachers' idea of intrinsic motivation might not fit the goals of the administration. Because intrinsic motivation takes autonomy, it is up to the individual to find the best way to motivate themselves within a school system.

**Teacher perception.** Though there are many changes in the evaluation system, it is essential to understand how teachers perceive the evaluation system. In a study, Hopkins (2016) used a stratified random sample of 5,000 teachers which focused on four categories: utility, feasibility, accuracy, and propriety standards. One hundred sixty-six teachers were selected out of the 5,000 in the study and showed a national distribution of teachers. The survey focused on the standardized testing impact on teachers, students, and value if professional development was included. In the findings, teachers noted when they were familiar with the evaluation system, students' performance went up versus teachers who were not familiar with the evaluation system. Teachers stated that student data helped teachers grow as professional educators eliminating evaluator bias. On the other hand, teachers who had not experienced the evaluation system felt the student data might be the main focus of their evaluation. Taken as a whole, teacher buy-in is a significant contributor to teacher perception of evaluation systems.

Jiang, Spote, and Luppescu (2015) studied the My Voice My School survey given by the University of Chicago Consortium on Chicago School Research to teachers and principals between the years 1991 to 2009. The purpose of the study was to evaluate the new evaluation

system implemented by schools in Chicago. The new system was based on the Framework for Teaching model and called Recognizing Educators Advancing Chicago Students. The survey showed teachers were concerned about using student growth, but overall felt positive about the new system. The teachers have a higher positive perception of the new system after the first year of implementation. Background knowledge plays a crucial role in teacher buy-in into any program. Principals and teachers felt the feedback sessions were useful and helped them grow as educators. When a specific growth plan was in place, the teachers felt more confident and were able to improve their practice.

The teachers' main concern was using student growth in their evaluation. Teachers were unsure of the fairness when using student growth. They were concerned with the burden of testing on the student and teacher. They were also concerned with the test not showing accurate student growth, and the students' home lives not taken into consideration. On the other hand, Hopkins' (2016) study found that teachers felt that using student performance for their evaluation enhanced their overall evaluation. They also felt the feedback received after the evaluation was beneficial to their professional development (Hopkins, 2016). Teachers also felt observations were not a detraction to the learning environment and understood the overall value of the program when student performance was part of the process (Hopkins, 2016). The role of the administration is also another vital part of understanding and embracing any evaluation system (Jiang, Spote, & Luppescu, 2015). Principals that participated and created a social construction of the policy were more likely to get teachers to buy into the new evaluation system (Jiang et al., 2015). It is vital for principals to embrace the shift from the old to new evaluation systems in order for teachers to adapt to the new changes.

When it comes to pay for performance, teachers' perceptions were not favorable. Farrell

and Morris (2004) included 330 teachers from the United Kingdom in their study and found teachers felt pay for performance was a way for the government to control their teaching. The teachers also expressed the extra pay only pulled good teachers into leadership roles instead of growing them as better teachers. Teachers also felt pay for performance put extra unneeded stress on the students.

When the focus is on collaboration and team building teacher, perspective is more favorable. In a study of 86 graduate students in an educational leadership program, all participants were employees in local school districts teachers who were surveyed using the Professional Appraisal System Survey (Zimmerman & Deckert-Pelton, 2003). According to the study, collaboration is a critical element of evaluation systems. Teachers felt better equipped when they received the constructive feedback for classroom refinements. Teachers use principal guidance to focus their goals on growing student performance. When teachers saw value in their evaluation, they felt their pedagogical background grew.

Each school district was able to use any goal setting protocol the principal seemed fit for each school. Because of this factor, each school collected data differently and set up professional development in various ways. With such varying factors, the data will show which goal setting protocol is most effective.

Each of these models was a foundation to the T-TAS. The models were vetted, and the elements identifying teacher effectiveness were used to create the Texas model.

### **Professional Development Appraisal System**

The Professional Development Appraisal System (PDAS) was adopted by Texas as the annual appraisal system for teachers. This system allowed school districts to evaluate teachers less frequently if they scored proficient or higher on the previous years' appraisal (Region 13

ESC, 2011). Teachers who were not satisfied with their appraisal scores could ask for a rebuttal within 10 days of the appraisal. The guidelines of the appraisal system had to be followed within the timeline or all appraisals would be void (Region 13 ESC, 2011).

The overview of the appraisal system consisted of 45-minute observation, data collected from all domains over the school year, walkthroughs, and teacher reflection report (Region 13 ESC, 2011). The four domains categories included: active, successful student participation in the learning process, learner centered instructions, evaluations and feedback on student progress, management of student discipline, instructional strategies, time and materials. Teachers received a rating of exceeds expectation, proficient, below expectation, or unsatisfactory. Proficient was the teacher's targeted score.

In order for a teacher to have received a proficient rating a teacher had to meet expectations 80% of the time. Though the appraisal system was to be objective because of the standards laid out by PDAS, it led appraisers to subjective professional judgments when administrators would evaluate teachers.

### **Texas Teacher Appraisal System**

The Texas Teacher Appraisal System (T-TAS) was a tool used to standardize the evaluation system in order to implement a career ladder and to improve teacher instruction in the classroom. The overall intent of the appraisal system was to help with teacher retention or dismissal. It was used as a guide for appraisers to improve teacher craft no matter the skill level of the teacher (Ovando & McCleary, 1991). Ovando and McCleary (1991) evaluated the T-TAS to determine its effectiveness. In the study, 455 graduate students were surveyed on teacher attitudes. The students consisted of early childhood education, middle school, high school teachers, and school administrators. T-TAS was a checklist of 71 items a principal used four

times a year to rate the teachers' effectiveness (Ovando & McCleary, 1991). The results determined teachers felt the system interfered with their teaching. It also implied the system had no criteria or rubric and the scores were not reliable and left the teachers questioning their skills with little feedback to monitor growth. Career teachers also felt they should be paid for advanced degrees and number of years teaching using the career ladder instead of appraisal scores. The ineffectiveness gave many stakeholders reasons to revamp the evaluation system.

The Texas Education Agency (1989) conducted a study of T-TAS to determine the effectiveness and to help create a better-quality system. The study showed there was a conflict of opinion between the teachers and the administration about different components of the appraisal system. The teachers wanted their points awarded for each level of criteria, but the principals did not want to provide the extra data needed to award the points for the teachers' evaluation score. For any score below satisfactory, documentation had to be included in the evaluation. It was clear the points would show the separation of effective and ineffective teachers, but because of the documentation, principals were reluctant to spend more time on appraisals and gave satisfactory scores to all teachers. Although there was evidence that showed conferences between teacher and appraiser were effective, they were few and far between because the only time conferences occurred was when a teacher received an unsatisfactory evaluation. One of the attributes of the T-TAS was to move up the career ladder. If teachers received an inaccurate score, then compensation was not received. The failed system's goal was to increase teacher effectiveness and student growth, which did not occur.

A review of the evaluation systems determined that T-TAS was an effective way to assess teacher effectiveness but followed with fidelity is required (Texas Education Agency, 1984). The career ladder component showed an adverse effect and should be eliminated from the

evaluation system. In order to create a more efficient system, recommendations were laid out for a more uniform system. Individual teacher assignment and student outcomes should be reviewed to determine student growth. It suggests more criteria should be added to the system to evaluate critical thinking and assign points to teachers for these criteria. The system was to take away the career ladder and not appraise teachers until their third year of teaching. It was also suggested not using the system for contract decisions but create a monetary program to pay for performance (Texas Education Agency, 1984).

When teachers and administrators use the system to create trusting relationships and resources are provided for teacher success then teacher performance was achieved (Ovando & McCleary, 1991). Maintaining a professional, trusting environment for all teachers created the most effective environment for teacher and student growth.

### **Texas Teacher Evaluation Support System**

Teacher evaluation has changed once again in Texas with the adoption of the T-TESS in 2014. The Texas Education Agency (2016) examined the evaluation system which was created by not only policymakers, but also educators, and administration. After reviewing feedback from 57 school districts which piloted the system in 2014, the evaluation system was cultivated to create the most effective evaluation for teacher refinement and student growth.

These standards give professionals a guide to follow for understanding by all stakeholders (Texas Education Agency, 2016). The evaluation system has core components that support teacher growth, including the revised Texas teacher standards of: 1) instructional planning and delivery, b) knowledge of students and student learning, c) content knowledge and expertise, d) learning environment, e) data-driven practices, f) professional practices and responsibility.

The purpose of these elements is to promote teacher growth using specific feedback. Unlike the previous evaluation system, PDAS, T-TESS is not a checklist but uses well-developed rubrics to create professional development plans specific to the teacher. The final observation uses a matrix that weights teacher evaluation for 80% and student growth for the remaining 20% of the evaluation.

The Texas Education Agency (2016) created the rubric that provides detailed descriptors for each performance level used to rate teacher effectiveness. Teachers and administrators participate in ongoing training and exposure to the rubric to better understand implementation. When schools provide adequate training and establish a routine to promote implementation with safeguards for learning to occur, the evaluation process should improve instruction and establish a support system. The implementation of instructional strategies identified as best practices, create one way to most significantly impact student growth (Texas Education Agency, 2016). Instructional strategies implemented by the teacher should fall within the guidelines provided by the rubric to create teacher and student growth.

To give the school districts ownership of T-TESS, they are given the flexibility to create a timeline that is most effective for them. The Texas Education Agency (2016) stated within these timelines, districts determine the number and length of observations, the requirements of each administrator, professional development guidance and reflection requirements, data collection, reinforcement and refinement processes, and the process for end of the year scoring. Giving this flexibility to the districts helps all stakeholders understand the system with more clarity. It also forces district stakeholders to collaborate and to be immersed in the growth process.

T-TESS roles do not stop with teacher and administration but involves all departments

including human resources, curriculum and instruction, technology, communication, and finance. Texas Education Agency (2016) also stated each department has individualized roles, that lead to the overall growth of the district. The teachers' role is to create learning goals and a professional development plan that focuses on reinforcement and refinements created during observations or goal setting conferences. Teachers are to use metacognition about their pedagogy and teaching skills. The cycle of reflection involves: teacher self-assessment, goal setting and establishing a professional development plan, professional development plan, formative reviews, prepare for the end of the year conference, end of year conference, primary goal setting and planning for the following year.

During the goal-setting process, teachers create professional development plans to understand their own strengths and weaknesses and create professional goals. Attending professional development opportunities aligned to these goals in order to increase student academic and social-emotional needs is encouraged. Through collaboration with administrators, teachers receive meaningful feedback in regard to their goals. To accomplish this growth schools must use the T-TESS rubric to fidelity and align professional development around the teacher's goal setting (Texas Education Agency, 2016). According to the T-TESS rubric teachers are responsible for creating and tracking growth through strategic goal setting. After teachers identify their target goals, they meet with their appraiser to create a support system through a goal-setting conference (Texas Education Agency, 2016). This process ensures the appraiser and teacher understand the goal and how to obtain the goal through subsequent actions. Throughout the year, the teacher and appraiser review the goals and actions to reach set goals (Texas Education Agency, 2016). In order for the evaluation system to be successful, the learning community continuously improves individuals' school-wide.



### **Chapter 3: Methodology**

Teacher evaluation systems have been studied over the years, but little research has focused on the goal setting portion of the T-TESS. The purpose of this study was to understand the effectiveness of goal setting protocols on teacher evaluation using the T-TESS. A concurrent nested approach was used. Within a current nested design both quantitative and qualitative data were collected at the same with more emphasis on one more than the other. The quantitative data was emphasized more in this study with the qualitative data serving to provide descriptive information. A quantitative study was chosen to determine the effectiveness of goal setting on teacher evaluation scores using the given data from the teacher's evaluation scores and teachers rating of goal setting protocol. Data was collected from four south Texas school districts that have participated in the T-TESS over the last four years.

The ANOVA compared goal setting methods used at each of the three school districts and within the individual school district. T-TESS does not give a specified protocol for school districts to follow. Therefore, it was determined by the school principals the protocol they use for their campus. The purpose was to see if there was a difference in the protocols and what affect it has on teacher overall evaluation score.

#### **Research Questions**

RQ1. Which goal setting protocol was more effective in increasing teacher evaluation score measures over a 4-year time period?

RQ2. Which of the three school districts' goal setting protocol and most effective?

RQ3. What was the relationship between goal setting protocol satisfaction and overall teacher performance?

#### **Population and Sample**

This study took place in three small school districts in south Texas. Each school district is made of up four schools: a high school (grades 9-12), middle school (grades 6-8), elementary

school (grades 3-5), and primary school (grades k-2).

The demographics in District A included 2,044 students with a 63.5% economic disadvantage with 193 teachers and 12 principals including assistant principals. The teachers in District A comprised of four beginning teacher, 37 teachers with 1-5 years of experience, 30 teachers with 6-10 years of experience, 52 teachers with 11-20-year of experience, 31 teachers with over 20 years of experience, with an average class size of 13 students. Within the state of Texas accountability system this school district met the standards in 2017. In 2018, the state of Texas changed the accountability system to a letter grade. This school district received a C rating with elementary and primary schools reporting improvement required.

District B had 1,085 students with a 77.6% economic disadvantage with 79 teachers and five principals. The teachers in District B included eight beginning teachers, 29 teachers with 1-5 years of experience, 18 teachers with 6-10 years of experience, 18 teachers with 11-20-year experience, eight teachers with over 20 years of experience, with an average class size of 14 students. Their accountability rating for 2017 indicated improvement required for the high school campus and met the standard for all other campuses. In 2018 with the new accountability system they received a F rating with the elementary and primary campuses improvement required.

District C had 1,979 students with a 57.5% economic disadvantage with 180 teachers and nine principals. The teachers in District C comprised of four beginning teacher, 31 teachers with 1-5 years of experience, 26 teachers with 6-10 years of experience, 57 teachers with 11-20-year experience, 23 teachers with over 20 years of experience, with an average class size of 12 students. In 2017 all campuses in the district met the standard. In 2018 with the new accountability system, the district received a B rating and all campuses met the standard.

District D enrolls 2,246 students with a 53.7% economic disadvantage with 143 teachers

and eight principals. The teachers in District D comprised of seven beginning teacher, 20 teachers with 1-5 years of experience, 30 teachers with 6-10 years of experience, 54 teachers with 11-20-year experience, 34 teachers with over 20 years of experience, with an average class size of 14 students. In this school district in 2017 all campuses met standard. In 2018 with the new accountability rating system this district received a B with all campuses meeting the standard. All the teachers were surveyed. Each school district has participated in the T-TESS over the past four years.

### **Instrument**

The researcher created a goal setting survey. The survey uses a Likert-type scale (0 – 5) and include three open-ended questions. It consists of basic demographic information and 28 research items. The survey took no more than 15 minutes to complete. It measured the effectiveness of the goal setting protocol of each school and school district and teacher growth after implementing goal setting. The topics of the survey included goal setting protocol, goal feedback, evaluation ratings, and goal setting attributes. Teachers answered questions using the Likert scale 0 being disagree and 5 being highly agree. Three open ended question detailed the goal setting process, and teachers' perception about the effectiveness of the process.

### **Data Collection**

To obtain access to the participants, a face-to-face meeting was arranged with the superintendent. Then permission from the Institutional Review Board was acquired. The survey was sent to the principals through the superintendent's desk using a survey monkey link to ensure the highest return rate at the beginning of December. The principals sent the link to the teachers the same week. Teachers received notification that all data collected was anonymous and for research purpose only through a cover letter with the survey link. Teachers had two

weeks to return the survey. A reminder email was sent a week before the survey was due, and another email was sent the day before it was due for any participants that have not returned the survey. A final email was sent a week after the survey was due for any surveys that may have not been turned in by the due date.

### **Data Analysis**

Data were collected using survey monkey. Data analysis was run directly after all data had been gathered. The first research question asked which goal setting protocol was more effective in increasing teacher evaluation scores in three school districts over a four-year period of time, was analyzed using a mixed between within between ANOVA. The independent between subjects would consist of the three different school districts, one categorical independent within subject variables of the four-time periods and one continuous dependent variable teacher evaluation scores for each year. The second question asked if there a relationship between goal setting protocol satisfaction and overall teacher evaluation score used a Pearson product-moment correlation. To understand the teachers' perceptions of how goal setting impacted professional development, a survey question was added to the quantitative survey. Then answers were coded by hand into themes. Once themes were identified the researcher was also able to determine the overall consensus of the teachers by district.

### **Role of the Researcher**

The researcher has no affiliation with the schools and obtained permission from the superintendent to conduct the study. The researcher emailed documents and had no influence on the participants. The researcher analyzed and interpreted the data. A report was given to each school district by the researcher.

**Reliability**

To check for reliability the researcher ensured the number of cases was correct in the case processing summary and the number of items was correct in the reliability statistics table. The inter-item correlation matrix was conducted to identify negative values. Then a Cronbach alpha was run. Values above 0.7 were acceptable. Scores lower than 0.7 were checked for incorrectly scored items and then removed. The researcher also conducted a factor analysis as a data reduction technique.

## Chapter 4: Findings

In this chapter a current nested design used both quantitative and qualitative data collected at the same with more emphasis on one more than the other. The quantitative data was emphasized more in this study with the qualitative data serving to provide descriptive information. A quantitative study was chosen to determine the effectiveness of goal setting on teacher evaluation scores using the given data from the teacher's evaluation scores and teachers rating of goal setting protocol.

Data were collected from three south Texas school districts that have participated in the T-TESS over the last four years. A factorial ANOVA analysis compared goal setting methods used at each of the three school districts and a correlational analysis was used to determine the relationship between teacher satisfaction of goal setting protocols and teacher evaluation scores. The T-TESS does not give a specified protocol for school districts to follow, therefore the protocols used were determined by the school principals.

### **Research Question 1: Which Goal Setting Protocol is More Effective in Increasing Teacher Evaluation Scores Measured Over a 4-Year Time Period?**

In order to respond to the first research question, a mixed between-within subjects' ANOVA was conducted to assess the overall evaluation scores from the T-TESS over a four-year period of time between three school districts. The file was divided by school districts.

Descriptive statistics for each school district including means and standard deviations of self-reported teacher evaluation scores are summarized by year in Table 1. After a cursory observation of District B, there seems to be a visually observable decrease in evaluation rating for the 2017-2018 school year. At District A, there does not seem to be much difference in the

mean scores across the four years. District C seems to have a relatively stable set of evaluation scores.

Table 1

*Descriptive Statistics*

Year		Mean	Std. Deviation	N
Evaluation score for the 2014-2015 school year	District B	3.3	0.7	14.0
	District A	3.3	0.6	61.0
	District C	3.2	0.8	21.0
	Total	3.3	0.6	96.0
Evaluation rating for the 2015-2016 school year	District B	3.2	0.7	14.0
	District A	3.2	0.6	61.0
	District C	3.2	0.7	21.0
	Total	3.2	0.6	96.0
Evaluation rating for the 2016-2017 school year	District B	3.2	0.4	14.0
	District A	3.2	0.4	61.0
	District C	3.3	0.5	21.0
	Total	3.2	0.4	96.0
Evaluation rating for the 2017-2018 school year	District B	3.0	0.6	14.0
	District A	3.3	0.5	61.0
	District C	3.2	0.4	21.0
	Total	3.2	0.5	96.0

The Levene's Test of Equality of Error Variances was checked to see if the assumptions of homogeneity of variance were violated. The significance value was greater than  $p = .05$ , which indicated no significant violations (Table 2). The homogeneity of variance significant values across school are: 2014-15 ( $p = .704$ ), 2015-2016 ( $p = .923$ ), 2016-2017 ( $p = .165$ ) and 2017-2018 ( $p = .321$ ) The Box's Test of Equality of Covariance Matrix was then checked, and it indicated the assumption was not violated ( $p = .009$ ), so a multivariate test was completed.

Table 2

*Levene's Test of Equality of Error Variances*

Year of Evaluation	Levene Statistic	df1	df2	Sig.
2014-2015	0.35	2.00	93.00	0.70
2015-2016	0.08	2.00	93.00	0.92
2016-2017	1.84	2.00	93.00	0.16
2017-2018	1.15	2.00	93.00	0.32

The Multivariate Test was evaluated to assess the interaction between time and school districts. The results indicate that there was a significant interaction between the three school districts and the four time periods (Wilks' Lambda = 0.87,  $F(3,182) = 2.25$ ,  $p = .040$ , partial eta squared = .069). Because the interaction demonstrated significance the simple main effects were calculated.

First, simple main effects of each year across districts were researched by using a one-way analysis of variance (ANOVA) for each time period. After first verifying that there were no violations of homogeneity of variance by using the Levene's Test (Table 3), the differences between districts for each of the time periods was measured. There were no significant differences between the district teacher evaluation scores for any of the four years (Table 4). This indicates that the teachers from each of the districts had similar evaluation scores for each year.



Table 3

*Test of Homogeneity of Variances*

	Levene Statistic	df1	df2	Sig.
2017-2018	1.315	2	97	.273
2016-2017	1.422	2	97	.246
2015-2016	.104	2	96	.901
2014-2015	.352	2	93	.704

Table 4.

*Teacher Evaluation Scores by Year*

Year		Df	Mean Square	F		Sig.
2017-2018	Between Groups	2	.406	1.617		.204
	Within Groups	97	.251			
	Total	99				
2016-2017	Between Groups	2	.140	.720		.490
	Within Groups	97	.195			
	Total	99				
2015-2016	Between Groups	2	.013	.037		.964
	Within Groups	96	.360			
	Total	98				
2014-2015	Between Groups	2	.026	.064		.938
	Within Groups	93	.402			
	Total	95				

Then the simple main effects for each district across years were analyzed by using a one-way repeated measures ANOVA for each district (Figure 1). In District A there was not a significant difference between the four time periods Wilks' Lambda = .877,  $F(3,58) = 2.70$ ,  $p = .054$ , partial eta squared = .123. In District B there was not a significant difference between the time periods (Wilks' Lambda = .772,  $F(3,11) = 1.08$ ,  $p = .396$ , partial eta squared = .22). In

District C there was no difference over time (Wilks' Lambda= .875,  $F(3,18) = 1.08$ ,  $p = .483$ , partial eta squared = .125). This indicates that teachers' evaluation scores remained consistent across the four-year time period.

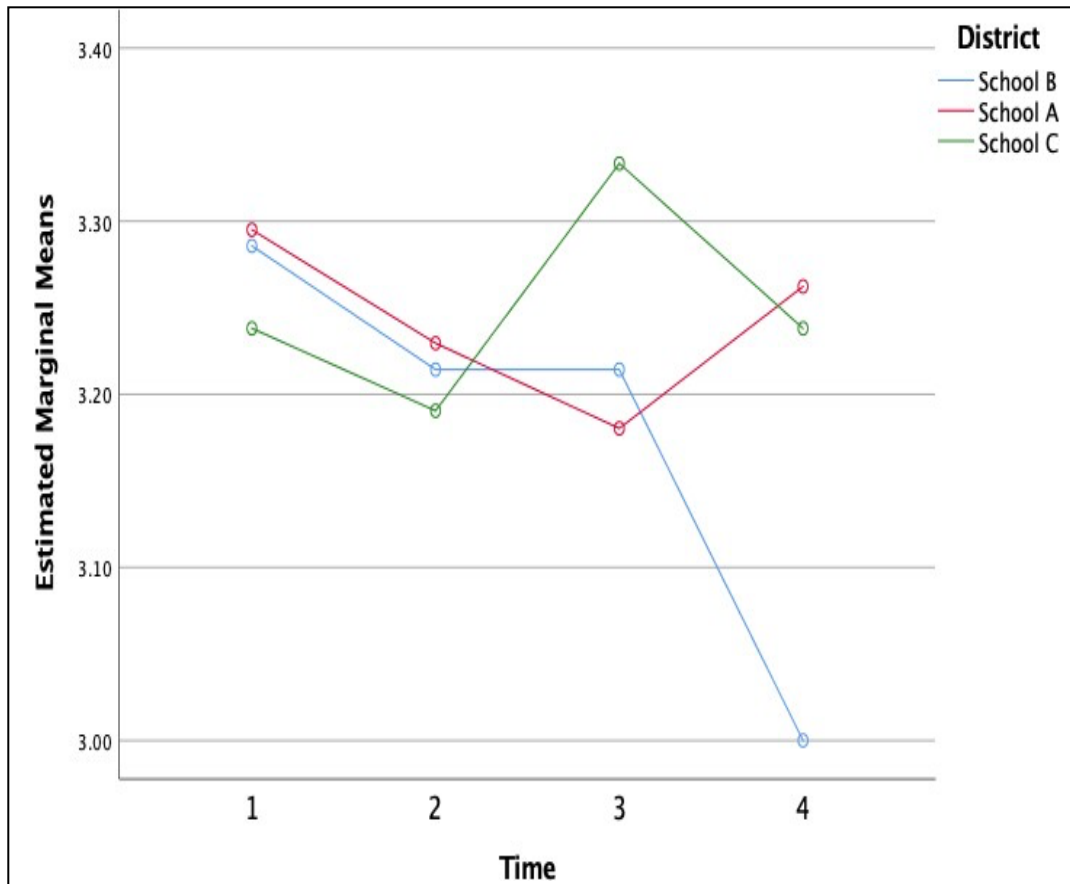


Figure 1. Evaluation scores over 4-year time period.

### Research Question 2. What is the relationship between goal setting protocol satisfaction and overall teacher performance?

In order to respond to the second research question, the teachers' goal setting protocol satisfaction was measured. Before using the survey responses to goal setting protocol satisfaction, an exploratory factor analysis was done to determine the reliability of the survey factors, and then correlations between the factors were run to determine relationships.

**Factor analysis.** In order to determine the validity of the goal setting protocol

perspectives measurement, an exploratory factor analysis was done to determine the viability of the survey factors. Inspection of the correlation matrix revealed the presence of many coefficients of 0.3 and above, indicating an appropriate level of item level relationships. The Kaiser-Meyer-Olkin value was 0.85, exceeding the recommended value of 0.6, and the Bartlett's Test of Sphericity did reach statically significance ( $p < .001$ ) which indicated sampling sufficiency. Principal components analysis revealed the presence of four components with eigenvalues exceeding 1, explaining 47.45%, 9.22%, 7.97%, and 5.17% of the variance respectively. An inspection of the scree plot (Figure 2), revealed a break after the second component. Using the Cattell's scree test it was decided to retain the two factors for further investigation.

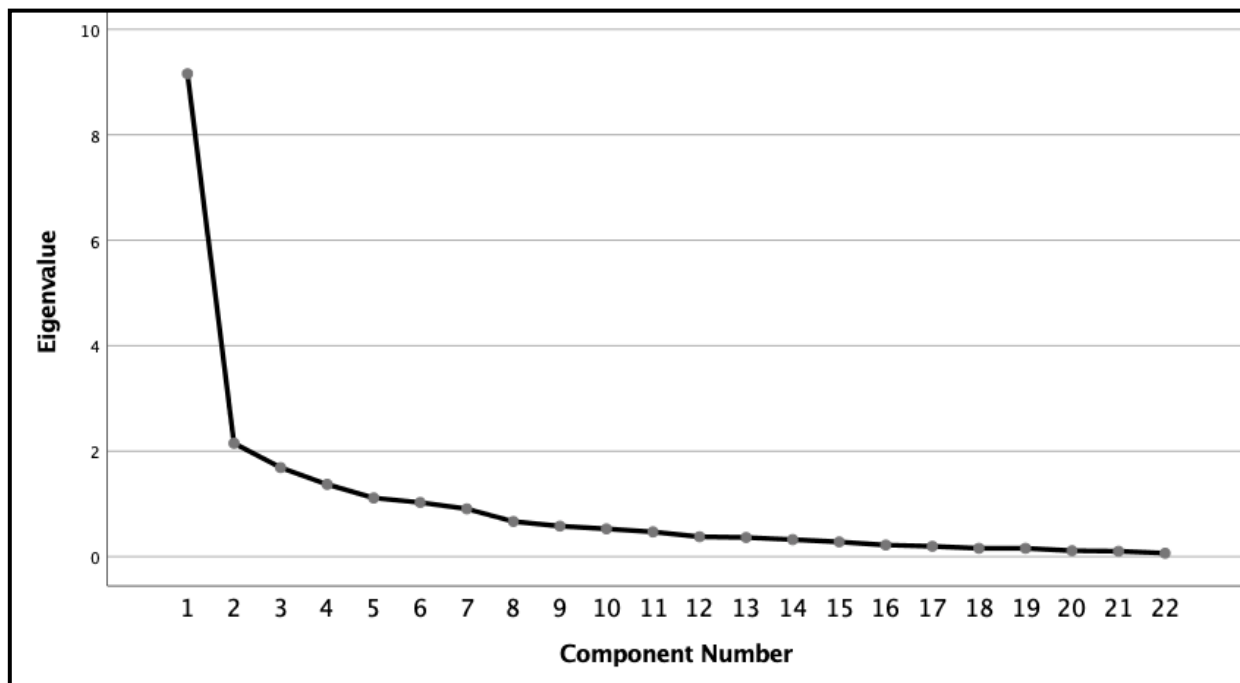


Figure 2. Cattell's scree test.

The two-component solution explained a total of 56.67%, with Component 1 contributing 47.45% and Component 2 contributing 9.22%. To aid in the interpretation of these two

components, an oblimin rotation was performed (Table 5). Goal setting items loaded strongly on Component 1 and evaluation items loading on Component 2. There was a weak negative correlation between the two factors ( $r = -.05$ ). The results of this analysis support the use of the goal setting items and the evaluation items as separate scales. Four items were not used in the analysis, as they cross loaded on the factors.

Table 5

*Structure Matrix*

	Component	
	1	2
Informal goal related feedback was frequent.	0.85	
Goal setting	0.83	
Formal goal related feedback was frequent.	0.82	
Goal setting policies were clear.	0.80	
Pre-conference	0.79	
Artifacts (lesson plans, materials, home/school communications)	0.79	
Self-evaluation	0.77	
Professional development activities	0.76	
Time was allotted for professional development.	0.75	
Feedback was given in a timely manner	0.73	
Training related to my goals was provided	0.70	
Ample time was allotted for goal setting.	0.67	
Observation of your classroom performance	0.64	0.33
My Professional Development plan was created using my goals.	0.64	
Student performance	0.58	
My goals were clear to me.	0.56	0.37
I think that my goals were created using the T-TESS standards.	0.50	0.46
I think that goal setting conferences were the same for all teachers.		
The overall impact of your evaluation on your professional development.		0.75
The overall quality of your evaluation.		0.63
I think that goals were the same for all teachers.		-0.47
Peer evaluations	0.40	-0.45

**Correlations.** After calculating the descriptive statistics for the satisfaction of goal setting protocol items (Table 6), the relationship between goal setting protocol satisfaction and overall teacher performance was investigated using the Pearson product-moment correlation coefficient (Table 7).

Table 6

*Descriptive Statistics for Satisfaction of Goal Setting Protocol Items*

District		N	Min.	Max.	Mean	Std. Dev
District B	Evaluation score for 2014-2015	14	2.00	5.00	3.29	.73
	Evaluation rating for 2015-2016	14	2.00	5.00	3.21	.70
	Evaluation rating for 2016-2017	14	3.00	4.00	3.21	.43
	Evaluation rating for 2017-2018	14	2.00	4.00	3.00	.55
District A	Evaluation score for 2014-2015	61	2.00	5.00	3.29	.56
	Evaluation rating for 2015-2016	63	2.00	5.00	3.22	.55
	Evaluation rating for 2016-2017	64	2.00	4.00	3.19	.43
	Evaluation rating for 2017-2018	64	2.00	5.00	3.27	.51
	Valid N (listwise)	61				
District C	Evaluation score for 2014-2015	21	1.00	5.00	3.24	.77
	Evaluation rating for 2015-2016	22	1.00	4.00	3.18	.66
	Evaluation rating for 2016-2017	22	3.00	4.00	3.39	.48
	Evaluation rating for 2017-2018	22	3.00	4.00	3.23	.43
	Valid N (listwise)	21				

Preliminary analyses were performed to ensure no violation of assumptions of normality, linearity, and homoscedasticity. District A demonstrated a small, non-significant, negative correlation ( $r = -.18$ ,  $n = 38$ ,  $p = .17$ ) between the goal setting variable and the teacher performance evaluation score. District B showed a non-significant negative relationship between the variables ( $r = -.51$ ,  $n = 13$ ,  $p = .77$ ). District C demonstrated a medium, yet non-significant, relationship between the variables ( $r = -.30$ ,  $n = 17$ ,  $p = .23$ ). This analysis indicates that none of

the districts demonstrated a significant relationship between goal setting protocol perceptions and teacher evaluation scores.

Table 7

*Relationship Between Goal Setting Protocol Satisfaction and Overall Teacher Performance (2017-2018).*

District			Goal setting	Eval rating
District B	Goal setting	Pearson Correlation	1	-.507
		Sig. (2-tailed)		.077
		N	13	13
	Evaluation rating	Pearson Correlation	-.507	1
		Sig. (2-tailed)	.077	
		N	13	14
District A	Goal setting	Pearson Correlation	1	-.183
		Sig. (2-tailed)		.170
		N	58	58
	Evaluation rating	Pearson Correlation	-.183	1
		Sig. (2-tailed)	.170	
		N	58	64
District C	Goal setting	Pearson Correlation	1	.308
		Sig. (2-tailed)		.230
		N	17	17
	Evaluation rating	Pearson Correlation	.308	1
		Sig. (2-tailed)	.230	
		N	17	22

### **Research Question 3. What is the teacher's perception of professional development focused on teacher goals?**

The qualitative data were gathered from the open-ended question on the survey which

gathered from the open-ended question on the survey which inquired about the teachers' campus goal setting process for the 2017-2018 school year. It was coded using three general codes: principal focus, teacher collaboration, and staff development meetings. The three codes were then developed into themes: principals and teachers create clear focused goals together, teachers create broad goals together, and PLC meetings focus on overall data to create campus goals. Although the general coding was completed on all of the responses, the themes were developed by district, as the responses were generally grouped according to the district processes. The following descriptions describe the findings that were common across each district.

In District A, 52 of the 73 teachers who responded to the survey, responded to the open-ended item. These teachers stated that goal setting was done in a faculty meeting at the beginning of the year or on an individual basis. One teacher stated, "our goal setting process was discussed at the beginning of the year during meetings to just put them in, but not a discussion as to how they should be produced. Those goals were not discussed at all during annual evaluation." The teachers felt the goal setting was not effective because although they were asked to set them individually, or as a staff, but they were never discussed again throughout the year. The teachers indicated that they used data from the previous years and created goals for the campus. One teacher stated, "it was based on my personal previous experiences and knowing what areas I needed to improve in." They were asked to set their personal goals on their own time using the previous year's data. Out of the 52 teachers that replied, 15 teachers stated they had to complete the goals on their own with no administration assistance. Seven teachers stated they created goals in a staff meeting and six teachers stated they created goals with other teachers. No teachers stated they created goals with a principal or other member of the school administration.

At District B, 13 out 18 teachers responded to the open-ended item. They felt that goal

setting was different this year than in previous years as it was more focused and consistent. One teacher stated “They were more focused”. Eleven of the teachers felt the goal setting was different. There were three different types of responses from teachers in this district. First, that the principals helped them create focused goals, such as a teacher who said the “Principal and teacher reviewed observations and determined goals. Goals were clear and time professional development was provided in order to help me accomplish the goals.”

Also, some teachers stated that they created goals in staff development meetings, as stated by a teacher who wrote, “I did feel that it was beneficial to goal set as an entire campus, but I also feel that maybe it should have been talked about with each grade level during PLC.” Describing the difference noticed in the referenced year, one teacher stated that “goals were clear and this time professional development was provided in order to help me accomplish the goals. The principals and leadership teams worked with the teachers to create goals for the school and the teachers.” Out of the 13 teachers that responded four teachers stated their goals were created with a principal or other administrator, one teacher stated they created goals with other teachers and four teachers stated they set goals during a staff meeting.

At District C, 17 out of 31 teachers responded to the open-ended item. All teachers stated that goals were the same from the previous years. Goals were either created during a PLC or the principal sent an email stating a deadline, or they were to be completed by the individual teachers. Examples of each of these perspectives can be observed in the quotes: “We were walked through the T-Tess Goal requirements and we worked during PLC on setting our personal goals,” and “We got an email with the deadline of when our goal needed to be set. We did receive a little guidance as to what our goals should be.” Eight teachers agreed that the goal setting process was effective, and 10 teachers felt it was not effective. They felt the process had



not changed over the 4-year time period. One teacher stated, “yes, the school administration made sure our goals were always in our minds and re-enforced by them. This made them effective.” While another teacher stated, “this goal was only slightly effective. It wasn’t clear from the start of the year and kept changing throughout the year. Professional development was not well-provided for young teachers like me.” This indicated a potential difference between the schools within the district in terms of goal effectiveness.

### **Summary**

This chapter has demonstrated the findings associated with each of the research questions. Results from the between-within factorial ANOVA, the factor analysis, the correlational analysis, and the qualitative coding have been presented. In the next chapter, there will be a discussion of these findings and they will be placed within a context of the current literature.

## Chapter 5: Discussion

The purpose of this study was to determine if goal setting and professional development procedures implemented by school districts affected teacher evaluations and to determine, within the school districts, how goal setting procedures vary between school district.

### **Research Question 1. Which goal setting protocol was more effective in increasing teacher evaluation scores measured over a 4-year time period?**

The results did not show any variations within the school districts from one year to the next except for a visible, yet non-statistically significant, change in the scores for District B in 2017-2018 school year from the 2016-2017 school year. After reviewing the teachers' surveys there did not seem to be a reason for the change. A few teachers did indicate there was a change in the protocol, but none of the reasons were in a consistent pattern. Teachers stated various reasons such as "Goals were more detailed and frequent", "They were more frequent" and "I worked with the principal to decide goals". Teachers from other districts shared "They were not allowed to build on previous goals." The teachers from each district stated the goals were set the same way each year and there was no follow through from the administration. One teacher stated "We were just told to enter the goals, they were not discussed with again at any time." Creating specific goals that addressed specific issues lead to more goal achievement (Latham et al., 2010). The goals need to have energy and a direction. When individuals can set their timeline to complete the task, they tend to work harder to complete it (Latham et al., 2010) As individuals work through these goals in order to be the most effective, there needs to be a time for feedback to refocus a goal and adjust accordingly. Because these goals are complex the knowledge to complete these tasks is not acquired yet, and with obtaining new information obtained there is a higher probability of meeting the goal (Poortvliet & Darnon, 2010).

Teachers would have likely been more successful at meeting goals if there was a set goal setting protocol with follow through. Latham et al. (2010) used the motivation hub to explain how individuals reach goals that included action, consistent personal goals, commitment, and self-efficacy. When asked about their goal setting protocol teachers did feel that when they worked with the principal, they were more successful. “Teacher are to look at their weakness, or ant aspect they want to improve in conjunctions with district wide initiative.” Individuals can obtain demanding goals by setting the timeline, continual feedback, and goal adjustment. Some teachers felt there was no follow through or guidance from the administration “I do not feel it is effective, I feel like all we do is paperwork and no one really looks at it unless there is an angry parent.”

Other teachers indicated they met with administration either through professional learning communities or during faculty meetings to set goals. But these goals are not on an individual bases and are more focused on a district or campus level. The T-TESS goal setting protocol was set up for individual teacher goal attainment. According to the T-TESS rubric teachers are responsible for creating and tracking growth through strategic goal setting. After teachers identify their target goals, they meet with their appraiser to create a support system through a goal setting conference (TEA, 2016). The results indicated that this process was not followed through by administration. For example, one teacher commented “There was more explanation the first year but most of it did not make sense.”

## **Research Question 2. What is the relationship between goal setting protocol satisfaction and overall teacher performance?**

There were also no significant correlations between teacher satisfaction with goal setting protocols and evaluation scores. This further substantiated the lack of evidence linking goal

setting with teacher evaluations. According to the teacher responses, goal setting was often left up to the teachers to do, and it was only sporadically monitored by administrators. “I felt there wasn’t feedback given to me in a timely manner to work on skill that needed improvement.” This lack of focus on goal attainment could explain the lack of relationships demonstrated in the quantitative data.

Within all three districts there was no established goal setting protocol. Teachers indicated that they had to create their goals on their own or set them with other teachers or as a district. “There was no communication, whether the goals were stated correctly or not and were discussed during the evaluation.” Locke and Latham (2006) stated that setting specific, attainable goals over vague goals creates higher work performance. They further state that employees are more satisfied in the workplace when charged with meeting challenging goals. Goals can motivate people to acquire new skills or recall existing abilities within one’s self. Because there was no set protocol this can be ambiguous and leave the teachers with uncertainty or they could feel that it is just a check mark on their paperwork pile on the corner of their desk.

Teachers indicated because there was no set protocol, they did not feel the goal setting was effective. One teacher stated “I am still unclear on how the evaluation system works.” To create effective goals, there needs to be something to be obtained at the end of meeting that goal and progress monitoring should be completed along the way. According to Bandura’s (1978) theory self-efficacy, self-satisfaction is achieved by setting and obtaining personal goals (Rachman & Bandura, 1978). Another teacher stated, “very little was explained about what was expected.” Achievement goals give individuals ways to understand how they learn or process to gain knowledge and master goals (Poortvliet & Darnon, 2010).

Across the three schools there was no change in the evaluation scores over the 4-year

time period. Teachers from two school districts indicated that there was no change in goal setting protocol over the 4-year time period. District B showed a visible decrease in evaluation scores in the 2017-2018 school year; however, it was not a statistically significant difference. Eight out of 13 teachers from District B indicated goal setting protocol changed in this time period for various reasons. Only one teacher indicated they worked with their principal to set goals. Goal setting has shown to be beneficial but requires diligent structure and guidance (Locke & Latham, 2006).

### **Research Question 3. What is the teacher's perception of professional development focused on teacher goals?**

The qualitative data was gathered from the open-ended question on the survey which inquired about the teachers' campus goal setting process for the 2017-2018 school year. It was coded using three general codes: principal focus, teacher collaboration, and staff development meetings. The three codes were then developed into themes: principals and teachers create clear focused goals together, teachers create broad goals together, and PLC meetings focus on overall data to create campus goals. Although the general coding was completed on all of the responses, the themes were developed by district, as the responses were generally grouped according to the district processes. The following descriptions describe the findings that were common across each district.

In District A, 52 of the 73 teachers who responded to the survey, responded to the open-ended item. These teachers stated that goal setting was done in a faculty meeting at the beginning of the year or on an individual basis. One teacher stated:

Our goal setting process was discussed at the beginning of the year during meetings, to just put them in, but not a discussion as to how they should be produced. Those goals were not discussed AT ALL during annual evaluation.

The teachers felt the goal setting was not effective because although they were asked to set them individually, or as a staff, but they were never discussed again throughout the year. The teachers indicated that they used data from the previous years and created goals for the campus. One teacher stated, “it was based on my personal previous experiences and knowing what areas I needed to improve in.” They were asked to set their personal goals on their own time using the previous year’s data. Out of the 52 teachers that replied, 15 teachers stated they had to complete the goals on their own with no administration assistance. Seven teachers stated they created goals in a staff meeting and six teachers stated they created goals with other teachers. No teachers stated they created goals with a principal or other member of the school administration.

At District B, 13 out of 18 teachers responded to the open-ended item. They felt that goal setting was different this year than in previous years as it was more focused and consistent. One teacher stated, “they were more focused.” Eleven of the teachers felt the goal setting was different. There were three different types of responses from teachers in this district. First, that the principals helped them create focused goals, such as a teacher who said, “principal and teacher reviewed observations and determined goals. Goals were clear and time professional development was provided in order to help me accomplish the goals.”

Also, some teachers stated that they created goals in staff development meetings, as stated by a teacher who wrote, “I did feel that it was beneficial to goal set as an entire campus, but I also feel that maybe it should have been talked about with each grade level during PLC.” Describing the difference noticed in the referenced year, one teacher stated that “goals were clear, and this time professional development was provided in order to help me accomplish the goals. The principals and leadership teams worked with the teachers to create goals for the school and the teachers.” Out of the 13 teachers that responded four teachers stated their goals

were created with a principal or other administration, one teacher stated they created goals with other teachers and four teachers stated they set goals during a staff meeting.

At District C, 17 out of 31 teachers responded to the open-ended item. All of the teachers stated that goals were the same from the previous years. Goals were either created during a PLC or the principal sent an email stating a deadline, or they were to be completed by the individual teachers. Examples of each of these perspectives can be observed in these quotes: “We were walked through the T-Tess Goal requirements and we worked during PLC on setting our personal goals,” and “We got an email with the deadline of when our goal needed to be set. We did receive a little guidance as to what our goals should be.” The teachers from this district had mixed feelings on the effectiveness of the goal setting process. Eight teachers agreed that the goal setting process was effective, and 10 teachers felt it was not effective. They felt the process had not changed over the 4-year time period. One teacher stated, “yes, the school administration made sure our goals were always in our minds and re-enforced by them. This made them effective,” while another teacher stated, “this goal was only slightly effective. It wasn’t clear from the start of the year and kept changing throughout the year. Professional development was not well-provided for young teachers like me.” This indicates a potential difference between the schools within the district in terms of goal effectiveness.

### **Theoretical Framework**

For the purpose of this study the Locke’s goal setting theory of motivation was used (Locke & Latham, 2006). This theory is based on task performance. Using feedback targeted to specific goals will increase task performance (Locke & Latham, 2006) The essential features of goal setting are working towards a goal is the main reason of the job, using clear and focused goals are greater motivators and will achieve better performance, goals are realistic and

measurable, and using appropriate feedback will increase overall performance and job satisfaction.

### **Professional Recommendations**

School districts seemed to be unclear on how to set obtainable goals for teachers on an individual level. There needs to be a set protocol established by the Texas Education Agency in order for all districts to have clear guidelines to set individuals goals using the SMART goal guidelines. The training needs to be valid for all participants. Districts need to ensure campus administrators are trained on goal setting protocols and are following through with fidelity. Teachers need to create personal goals and create a plan with administrators to become more successful teachers. As teachers meet goals, or fall behind on goals, administrators and the teacher could come up with a plan of action to move forward or to remedy any failures in the goal setting process. This plan may include individualized professional development.

### **Limitations**

There were several limitations encountered during the research process. A major issue was the lack of participants. Although it was initially conceptualized to include four school districts and individual campuses within the districts, participant constraints forced the number of districts to be reduced to three, and the individual campuses did not garner enough participants to warrant that level of study. Another issue was found in regard to the survey. After starting to analyze the qualitative data, the researcher would have appreciated if there had been a specific place for the teachers to write the actual protocol.

### **Delimitations**

The information in many of the qualitative responses was limited. The small sample size may have been a contributing factor to the lack of statistical signification on multiple tests.



Another delimitation may have been the location of all the schools. They were located in the same area and composed of the same population.

### **Future Research**

In the future, more research should be done on individual campuses and their goal protocols. There should also be research on effective protocols created at other school districts and how they rolled them out to their campuses. Researchers need to work closely with the Texas Education Agency to create a well-rounded goal setting protocol and then do a follow up 4-year study on how the goal setting protocol affected the retention of teachers. This topic deserves a much larger scale investigation.

## References

- Amrein-Beardsley, A., & Holloway, J. (2017). Value-added models for teacher evaluation and accountability: Commonsense assumptions. *Educational Policy*, 33(3), 516–542. doi:10.1177/0895904817719519
- Association of Texas Professional Educators (2014). Retrieved from <https://www.atpe.org/en/Home>
- Baker, M. (2010). Career confidence and gendered expectations of academic promotion. *Journal of Sociology*, 46(3), 317–334. doi:10.1177/1440783310371402
- Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance. *Educational Researcher*, 44(2), 77–86. doi:10.3102/0013189x15574904
- Blumberg, P., & Flaherty, J. A. (1985). The influence of noncognitive variables in student performance. *Academic Medicine*, 60(9), 721–3. doi:10.1097/00001888-198509000-00009
- Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal*, 106(1), 3–20. doi:10.1086/496904
- Chambers, S., & Tate, K. (2013). Value-added measures in restructuring urban schools. *Education and Urban Society*, 47(6), 723–738. doi:10.1177/0013124513508582
- Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, 45(6), 378–387. doi:10.3102/0013189x16659442
- Danielson, C. (2010). Framework for teaching. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. doi:10.4135/9781506326139.n275
- Danielson, C. (2013). Danielson’s framework for teaching for classroom observations. Retrieved from [https://k12education.gatesfoundation.org/download/?Num=2792&filename=Danielson-FFT\\_10\\_29\\_101.pdf](https://k12education.gatesfoundation.org/download/?Num=2792&filename=Danielson-FFT_10_29_101.pdf)
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation: To enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher*, 44(2), 132–137. doi:10.3102/0013189x15575346
- Derrington, M. L. (2016). Implementing teacher evaluation. *Journal of Research on Leadership Education*, 11(2), 181–199. doi:10.1177/1942775116658689
- Erikson, T. (2002). Goal-setting and entrepreneurial self-efficacy. *The International Journal of Entrepreneurship and Innovation*, 3(3), 183–189. doi:10.5367/000000002101299187

- Farrell, C., & Morris, J. (2004). Resigned compliance. *Educational Management Administration & Leadership*, 32(1), 81–104. doi:10.1177/1741143204039301
- Finnigan, K. S. (2011). Principal leadership in low-performing schools. *Education and Urban Society*, 44(2), 183–202. doi:10.1177/0013124511431570
- Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100–107. doi:10.3102/0013189x14521864
- Garrett, R., & Steinberg, M. P. (2015). Examining teacher effectiveness using classroom observation scores. *Educational Evaluation and Policy Analysis*, 37(2), 224–242. doi:10.3102/0162373714537551
- Geiger, R. L. (2018). Postwar American higher education: Documenting recent history. *Curriculum, Accreditation, and Coming of Age of Higher Education*, 139–150. doi:10.4324/9780203793947-6
- Glassman, M., Glassman, A., Champagne, P. J., & Zugelder, M. T. (2010). Evaluating pay-for-performance systems: Critical issues for implementation. *Compensation & Benefits Review*, 42(4), 231–238. doi:10.1177/0886368710374391
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2001). *Supervision and instructional leadership* (5th ed.). Boston, MA: Allyn and Bacon.
- Haertel, H. E. (2012). *Reliability and validity of inferences about teachers based on student scores*. Retrieved from <https://eric.ed.gov/?id=ED560957>
- Hill, H. C. (2009). Evaluating value-added models: A validity argument approach. *Journal of Policy Analysis and Management*, 28(4), 700–709. doi:10.1002/pam.20463
- Hopkins, P. (2016). Teacher voice: How teachers perceive evaluations and how leaders can use this knowledge to help teachers grow professionally. *NASSP Bulletin*, 100(1), 5–25. <https://doi.org/10.1177/0192636516670771>
- Hulleman, C. S., & Barron, K. E. (2010). Performance pay and teacher motivation: Separating myth from reality. *Phi Delta Kappan*, 91(8), 27–31. doi:10.1177/003172171009100806
- Hunter, M. (1991). Testing and teaching. *Educational Considerations*, 18(2), 29–30. doi:10.4148/0146-9282.1548
- Jiang, J. Y., Spote, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform. *Educational Researcher*, 44(2), 105–116. doi:10.3102/0013189x15575517
- Jilliam, J., Tocci, C., Holtzman, S., & Williams, J. (2013). Foundations of observation: Considerations for developing a classroom observation system that helps districts achieve consistent and accurate scores. *MET Project, Policy and Practice Brief*. Seattle, WA.

- Kane, T. J., McCaffrey, D. F., Miller, T. M., & Staiger, D. O. (2012.). *Have we identified effective teachers? Validating measures of effective teaching using random assignment - K-12 education*. Retrieved from <https://k12education.gatesfoundation.org/resource/have-we-identified-effective-teachers-validating-measures-of-effective-teaching-using-random-assignment/>
- Kettler, R. J., & Reddy, L. A. (2017). Using observational assessment to inform professional development decisions: Alternative scoring for the Danielson framework for teaching. *Assessment for Effective Intervention*, 44(2), 69–80. <https://doi.org/10.1177/1534508417745628>
- Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34–70. doi:10.1177/0013161x08327549
- Koedel, C., & Betts, J. (2009). *Value-added to what? How a ceiling in the testing instrument influences value-added estimation*. doi:10.3386/w14778
- Latham, G. P., Stajkovic, A. D., & Locke, E. A. (2010). The relevance and viability of subconscious goals in the workplace. *Journal of Management*, 36(1), 234–255. <https://doi.org/10.1177/0149206309350777>
- Liang, G., & Akiba, M. (2015). Characteristics of teacher incentive pay programs: A statewide district survey. *Journal of Educational Administration*, 53(6), 702–717. doi:10.1108/jea-09-2013-0106
- Locke, E. A., & Latham, G. P. (2006). *New developments in goal setting and task performance*. London, United Kingdom: Routledge.
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81(1), 59–77. doi:10.1348/000709910x503501
- Martínez, J. F., Schweig, J., & Goldschmidt, P. (2016). Approaches for combining multiple measures of teacher performance. *Educational Evaluation and Policy Analysis*, 38(4), 738–756. doi:10.3102/0162373716666166
- Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: supporting the art and science of teaching*. Alexandria, VA: ASCD.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29(1), 67–101. doi:10.3102/10769986029001067
- Mihaly, K., McCaffrey, D. F., Staiger, D. O., & Lockwood, J. R., (2013). A composite estimator of effective teaching - K-12 education. <https://k12education.gatesfoundation.org/resource/a-composite-estimator-of-effective-teaching/>

- Mintrop, R., Ordenes, M., Coghlan, E., Pryor, L., & Madero, C. (2017). Teacher evaluation, pay for performance, and learning around instruction: Between dissonant incentives and resonant procedures. *Educational Administration Quarterly*, *54*(1), 3–46. doi:10.1177/0013161x17696558
- Moeller, A. J., Theiler, J. M., & Wu, C. (2011). Goal setting and student achievement: A longitudinal study. *The Modern Language Journal*, *96*(2), 153–169. doi:10.1111/j.1540-4781.2011.01231.x
- Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, *95*(2), 255–264. doi:10.1037/a0018478
- Nielsen, L. D. (2014). Teacher evaluation: Archiving teaching effectiveness. <https://journals.sagepub.com/doi/abs/10.1177/0027432114536736>
- Ovando, M. N. & McCleary, L. (1991). The supervision of teaching within a formal teacher appraisal system. *School Organization*, *11*(2), 145-151. <https://doi.org/10.1080/13632434.1991.10384446>
- Poortvliet, P. M., & Darnon, C. (2010). Toward a more social understanding of achievement goals. *Current Directions in Psychological Science*, *19*(5), 324–328. doi:10.1177/0963721410383246
- Rachman, S., & Bandura, A. (1978). *Perceived self-efficacy: Analyses of Bandura's theory of behavioral change*. Oxford, England: Pergamon Press.
- Region 13 ESC: Positively affecting the future (2011). Retrieved from <http://www4.esc13.net/>
- Rice, J. K., Malen, B., Jackson, C., & Hoyer, K. M. (2015). Time to pay up. *Educational Evaluation and Policy Analysis*, *37*(1), 29–49. doi:10.3102/0162373714524622
- Robbins, S. P., & Judge, T. A. (2015). *Organizational behavior*. Boston, MA: Pearson.
- Sanders, W., & Rivers, J. (1996) Cumulative and residual effects of teachers on future. Retrieved from [https://www.beteronderwijsnederland.nl/files/cumulative and residual effects of teachers.pdf](https://www.beteronderwijsnederland.nl/files/cumulative%20and%20residual%20effects%20of%20teachers.pdf)
- Sledge, A., & Pazey, B. L. (2013). Measuring teacher effectiveness through meaningful evaluation. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *36*(3), 231–246. doi:10.1177/0888406413489839
- Springer, M. (2009). Rethinking teacher compensation policies: Why now, why again? In M. G. Springer (Ed.), *Performance incentives: Their growing impact on American K-12 education*. Washington, DC: Brookings Institution Press.

- Steinberg, M. P., & Kraft, M. A. (2017). The sensitivity of teacher performance ratings to the design of teacher evaluation systems. *Educational Researcher*, 46(7), 378–396. doi:10.3102/0013189x17726752
- Texas Education Agency. (2016). Welcome to Texas Education Agency: Texas Education Agency. Retrieved from <https://tea.texas.gov/>
- Thoonen, E. E. J., Slegers, P. J. C., Oort, F. J., Peetsma, T. T. D., & Geijsel, F. P. (2011). How to improve teaching practices. *Educational Administration Quarterly*, 47(3), 496–536. doi:10.1177/0013161x11400185
- Timmermans, A. C., Snijders, T. A. B., & Bosker, R. J. (2012). In search of value added in the case of complex school effects. *Educational and Psychological Measurement*, 73(2), 210–228. doi:10.1177/0013164412460392
- Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Washington, DC: Education Sector.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project.
- Whitehead, M. (1952). *Teachers look at supervision - ASCD*. Retrieved from [http://www.ascd.org/ASCD/pdf/journals/ed\\_lead/el\\_195211\\_whitehead.pdf](http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_195211_whitehead.pdf)
- Yuan, K., Le, V.-N., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2013). Incentive pay programs do not affect teacher motivation or reported practices. *Educational Evaluation and Policy Analysis*, 35(1), 3–22. doi:10.3102/0162373712462625
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3-17. Doi: 10.1207/s15326985ep2501\_2
- Zimmerman, B. J. (2006). Development and adaption of expertise: The role of self-regulatory processes and believes. In K. A. Ericsson, N. Charness, P. H. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 705-722). New York, NY: Cambridge University Press.
- Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin*, 87(636), 28–37. doi:10.1177/019263650308763604

## Appendices

## Appendix A: Invitation to Participate

### **Dear Sir or Madam,**

You are invited to participate in a research study about your awareness of experiences with evaluations, goal setting, and professional development. The information obtained from this survey will be used to create some “best practices” in the fields of teacher goal setting and professional development in order to help teachers to progress to their greatest potentials. Filling out this relatively short survey will take about 20 minutes. Your participation is completely voluntary and you may decline to take this survey if you choose. Please note there is no direct benefit that will accrue to you from taking this survey; however, your participation will contribute greatly to the body of knowledge about this topic and hopefully help teachers and districts in the future.

### **Things you should know**

Your responses to this survey will be anonymous and the research findings from the data collected will be reported in aggregate form. Since we are not collecting any personally identifying information from you, your responses will not be linked back to you.

### **Taking the survey**

Completing and submitting this survey represents informed consent to participate in the research study. You may choose to opt out of the study at any time and simply close the survey. To take the survey, please click on the link below and follow the directions. This survey will be available for your response until December 31, 2018. <https://www.surveymonkey.com/r/BQSNTZS>

If you have questions at any time about the study or survey, you may contact Dr. Stephanie Hartzell at [hartzell@uiwtx.edu](mailto:hartzell@uiwtx.edu).

For questions about your rights as a research participant or to discuss problems, complaints or



concerns about a research study, or to obtain information or offer input, contact the UIW Institutional Review Board (IRB) at (210) 805-3036. This research and survey tool has been approved by the UIW IRB (IRB # 1382 - 3 1382)

Thank you in advance for your time.

Sincerely,

Stephanie Hartzell, PhD

## Appendix B: Survey

## Consent

\* 1. We would like to learn about your experiences with evaluations, goal setting, and professional development. As you are a teacher in your school district you are qualified to participate. If you decide to participate in this study, you will be asked to complete a survey which should take no longer than 15 minutes. Your responses will be kept anonymous and only aggregated and non-personally identifiable information will be reported back to your district office. We are hoping that the results of this study will help to create some "best practices" in the fields of teacher goal setting and professional development in order to help teachers to progress to their greatest potentials. Your participation will help both your district and other districts within the state.

Participation in this study is completely voluntary and you can stop at any time. If you have any questions about the study, please contact me at [hartzell@uiwtx.edu](mailto:hartzell@uiwtx.edu). For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information or offer input, contact the UIW Institutional Review Board (IRB) at (210) 805-3036. This research and survey tool has been approved by the UIW IRB (IRB #XX-XX-XXX).

Do you agree to participate in this research study?

Yes  No

## Please describe your Teaching Experience

\* 2. Including the current year, how many years have you taught in your current district?

\* 3. Including the current year, how many years have you taught in any district?

\* 4. Your current teaching assignment grade level

- pk-2
- 3-5
- 6-8
- 9-12

\* 5. Your gender

- Male
- Female

## Evaluation Information

\* 6. What was your overall evaluation rating for the 2017-18 school year?

\* 7. In 2017-18, how satisfied were you with:

	Very Satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
The overall quality of your evaluation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The overall impact of your evaluation on your professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 8. How many classroom observations did you have in 2017-18?

\* 9. What was your overall evaluation rating for the 2016-17 school year?

\* 10. In 2016-17, how satisfied were you with:

	Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very Dissatisfied
The overall quality of your evaluation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The overall impact of your evaluation on your professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 11. What was your overall evaluation rating for the 2015-16 school year?

\* 12. In 2015-2016, how satisfied were you with:

	Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
The overall quality of your evaluation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The overall impact of your evaluation on your professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 13. What was your evaluation score for the 2014-15 school year?

\* 14. In 2014-15, how satisfied were you with:

	Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
The overall quality of your evaluation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The overall impact of your evaluation on your professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### Rating Attributes of Goal Setting for 2017-2018

\* 15. Describe the attributes of the procedures used during your most recent goal-setting conference.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I think that my goals were created using the T-TESS standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My goals were clear to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that goals were the same for all teachers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that goal setting conferences were the same for all teachers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My Professional Development plan was created using my goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 16. The following sources were used for your most recent EVALUATION RATING:

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Observation of your classroom performance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pre-conference	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Artifacts (lesson plans, materials, home/school communications)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Student performance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peer evaluations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Self-evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Goal setting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Professional development activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 17. Describe the attributes of the feedback you received during your most recent EVALUATION EXPERIENCE:

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Formal goal related feedback was frequent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informal goal related feedback was frequent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Feedback was given in a timely manner	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 18. Describe these attributes of the evaluation context:

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Ample time was allotted for goal setting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Time was allotted for professional development.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Training related to my goals was provided	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Goal setting policies were clear.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 19. Describe your campus goal setting process for the 2017-18 school year?

\* 20. Was the goal setting process different than previous years? If so, explain?

\* 21. Did you feel the goal-setting process was effective? Why or Why Not?

## Appendix C: IRB

**Principal Investigator Signature**

I certify that the information above is accurate and complete. I will request prior IRB approval for any changes to the approved protocol and/or informed consent forms, and will not implement those changes until I receive IRB approval. I will report any adverse effects to the IRB immediately. I agree to comply fully with the ethical principles and regulations regarding the protection of human subjects in research.

Principal Investigator

**Signed:** This form was signed by Mrs Julie White (jmwhite1@student.uiovtx.edu) on 12/15/2018 2:45 AM

I certify that the student named above is knowledgeable of the regulations and policies governing research with human subjects and has sufficient training and experience to conduct this study as described in the proposed protocol.

I furthermore certify the following:

- I have reviewed this application;
- I will maintain knowledge of the direction and completion of the project;
- I will assure the student investigator remains in compliance with UIW and federal human subjects protection policies;
- I assure the student investigator will promptly file for revision, amendment, annual continuing review, or completion of the supervised protocol and will provide assistance to them as needed;
- I assure both the student investigator and I will promptly report any significant or untoward adverse effects to the UIW IRB;
- If this protocol is to be conducted as part of a course, I will ensure the student investigator is informed of the requirement to file appropriate documents at the end of the course; and
- If at any time I am unable to proceed as Faculty Supervisor (e.g., end of the course during which research was planned, sabbatical leave, or exit from the University), I will assist the student in designating an alternate Faculty Supervisor for the remainder of the study.

Faculty Supervisor

**Signed:** This form was signed by Dr. Stephanie Hartzell (hartzell@uiovtx.edu) on 12/15/2018 2:52 AM