PREDICTING EDUCATOR PERCEPTIONS OF TENNESSEE'S TEACHER EVALUATION SYSTEM: THE ROLE OF TEACHER GENDER, YEARS OF EXPERIENCE, AND SCHOOL LEVEL TAUGHT

by

Leonida Bell

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

2024

PREDICTING EDUCATOR PERCEPTIONS OF TENNESSEE'S TEACHER EVALUATION SYSTEM: THE ROLE OF TEACHER GENDER, YEARS OF EXPERIENCE, AND SCHOOL LEVEL TAUGHT

by Leonida Bell

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

2024

APPROVED BY:

Jeffrey S. Savage, Ed.D., Committee Chair

Melinda B. Carver, Ph.D., Committee Member

ABSTRACT

This quantitative, correlational study aims to determine if a predictive relationship exists between teacher perceptions of Tennessee's teacher evaluation and the linear combination of teacher gender, years of experience, and school-level taught (elementary, middle, high). Although teachers have recognized the current teacher evaluation system, many educators expressed concerns about it. This study is vital in determining teachers' perceptions of the evaluation system. Participants of this study included a sample of 108 certified elementary, middle, and high school teachers from one of the public school districts in southwestern Tennessee. Data were collected via an anonymous, self-paced online survey utilizing a modified Teacher Evaluation Profile (TEP) questionnaire. Multiple regression was used to measure the correlation between the predictor variables and the criterion variable. The findings revealed that no significant relationship existed between the predictor variables, taken together as a model, and the criterion variables ($F(3,104) = 1.86$, $p = .14$, $R^2 = .05$). This study findings emphasize the critical need to incorporate teachers' perspectives in developing a fair teacher evaluation system, rather than relying on a one-size-fits-all approach. Recommendations for future research include conducting a qualitative research design incorporating teacher interviews to uncover teachers' experiences and classroom decisions regarding the Tennessee teacher evaluation system.

Keywords: student achievement, self-determination theory, Tennessee Education Acceleration Model, teacher evaluation, teacher perceptions, Tennessee value-added assessment, teacher quality

Copyright Page

Dedication

I want to dedicate my work to my entire family in the United States and the Philippines, especially to my dad, Pedro Tubil, who fills me with pride and joy with his unwavering support. I am also deeply grateful to my husband, Curtis, and son, Curtis Jr., for their constant love and understanding that have been my rock throughout my doctorate journey.

I also dedicate this work in memory of my mom, Antonia Tubil, and sister, Joy, who, in life, instilled in me the values of hard work, determination, and dedication. These values have been the guiding light in my academic journey, and I am forever grateful for their influence. Mom and Sis, I know you are my angels from the heavens, and you are greatly missed!

Acknowledgments

I would like to express my deepest gratitude to Dr. Jeffrey Savage, my dissertation chair, for his unwavering guidance, profound understanding, and constant encouragement throughout the dissertation stage. His positive energy was a beacon of light that guided me through the challenges of completing the dissertation. I am also indebted to Dr. Melinda Carver for her insightful feedback, which was instrumental in honing my writing skills. Additionally, I am truly appreciative of Dr. Laura Mansfield for her meticulous effort and unwavering dedication in reviewing and approving my research proposal.

Further, I sincerely acknowledge the school district personnel, principals, and teachers who contributed to the success of my study. Lastly, I am beyond grateful to my best friend, whose support and empowerment have contributed to my success, and to my brothers and sisters in the Philippines, who constantly prayed for me and believing my capacity to fulfill my dreams.

Above all, I thank God for giving me wisdom, good health, and guidance throughout my doctorate journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Basic Psychological Need Theory (BPNT)

Elementary and Secondary Education Act (ESSEA)

End-of-Course (EOC)

Every Student Succeeds Act (ESSA)

First to the Top (FTTT)

Instructional Growth for Effectiveness and Results (TIGER)

Level of Effectiveness (LOE)

No Child Left Behind (NCLB)

Professional Development (PD)

Self-Determination Theory (SDT)

Teacher Effectiveness Measures (TEM)

The Tennessee Department of Education (TDOE)

The Tennessee Education Research Alliance (TERA)

Tennessee Comprehensive Assessment Program (TCAP)

Tennessee Evaluation Acceleration Model (TEAM)

Tennessee Value-Added Assessment System (TVAAS)

Value-Added Models (VAMs)

CHAPTER ONE: INTRODUCTION

Overview

This quantitative, predictive correlational study aims to determine if a predictive relationship exists between teacher perceptions of Tennessee's teacher evaluation and the linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high). Chapter One provides a background for the teacher evaluation system in a public school setting. The background includes an overview of the historical and social context and the theoretical framework for this study. The problem statement examines the scope of the current literature on this topic. The significance of the present study follows its statement. Finally, the research question is introduced, and definitions pertinent to this study are provided.

Background

Teacher quality delivers excellence in teaching, which produces student academic achievement (Kirschner & Hendrick, 2020. If research is detailed on anything, it is this one point: effective teaching matters in student academic achievement (Hattie, 2009, 2023; Kirschner & Hendrick, 2020). Warner (2016) stated that an excellent teacher inspires lifelong learning for the students and embodies organizational skills, proper application of instructional pedagogies, strong classroom management, high expectations, and moral and ethical actions and decision-making. Faced with public demands for assessing teacher quality, policymakers in the United States strived to enhance the quality of teaching in schools, which led to the passing of the No Child Left Behind (NCLB) legislation in 2001 (Sartain & Steinberg, 2016). The goal of the NCLB proposition was to shrink or eliminate the achievement gap among various groups of students (Heise, 2017), which required states to reach a series of student-achievement benchmarks through a yearly standardized assessment. As part of education's ongoing reform

efforts, the Every Student Succeeds Act (ESSA) replaced NCLB in 2015, superseding the nation's K-12 education law that nonetheless still focused on accountability through the lens of teacher quality and student achievement (Ravitch, 2016). To meet the federal expectations in ensuring public school accountability for teacher quality and student learning, policymakers across the nation adopted standards-based observation-driven evaluation systems (Williams & Hebert, 2020) . Tennessee pivoted to use the Teacher Evaluation Accelerator Model (TEAM) to measure teachers' effectiveness without carefully considering educators' perceptions before implementing the evaluation system (Glatthorn et al., 2019).

Historical Overview

American teacher evaluation can be traced to the colonial period when groups of citizens toured the schools to observe how teachers managed the classroom (Stronge & Tucker, 2013). In the 1800s, administrative positions became more prevalent, and master teachers and full-time school administrators assumed responsibilities for evaluation; however, evaluators typically performed evaluations informally and with no written procedures (Stronge & Tucker, 2013). During the 1960s, President Lyndon Johnson passed the Great Society legislation of the Elementary and Secondary Education Act (ESEA) of 1965 (Alford, 1965), which ratified educational programs to encourage underprivileged children to succeed in school (Mertens, 2001). Before the 1970s, teacher evaluation's focus was primarily summative, with little or no feedback on teacher practice (Stronge & Tucker, 2013). Several decades have elapsed using traditional teacher evaluations, and many changes have transpired, in which teacher evaluations were tools to assess teacher competence and were expected to foster educators' professional development and growth (Goldstein, 2014). Since there was a rapid increase in constituents' dissatisfaction with measuring teachers' ability to educate their children, the No Child Left

Behind Act of 2001 (U.S. Department of Education, 2002) required states to implement plans to ensure highly qualified teachers teach in core subjects (Mills, 2008). The NCLB mirrored the ESEA of 1965, which aimed to deliver quality education for all students, and in doing so, the NCLB required states to fill every public-school classroom with highly qualified teachers. Under the NCLB Act, highly qualified teachers were required to meet state qualifications and licensing criteria for the grade levels and subject areas taught (Goldstein, 2014).

Despite NCLB's robust intentions to determine highly qualified teachers, there were discourses on appropriate models to utilize and implement its requirements in the K-12 school setting. Teacher evaluations became the focal policy target across the states (Hazi & Rucinski, 2009). Former U. S. President Obama passed and signed the Every Student Succeeds Act (ESSA) in 2015 to permit states to decide how administrators evaluated teachers (Amrein-Beardsley, 2020). Such authorization prompted state policymakers to adopt new and rigorous teacher evaluation systems. Under ESSA, teacher evaluations were no longer correlated with student proficiency and standardized test scores as previously mandated by NCLB. However, ESSA suggested performing annual assessments to check student performance. Despite loosened ESSA restrictions, school district leaders in several states continuously utilized value-added models (VAMs) as part of teacher evaluations. Ballou and Springer (2015) mentioned how a current teacher evaluation system propelled by a competitive grant called Race to the Top (RttT), which took effect in 2011, incentivized states to rectify K-12 education encompassing low-performing schools and improved teacher effectiveness. Following the RttT, most states incorporated VAMs for teacher evaluations and high-stakes employment decisions, and several states and school districts were involved in court cases (Paige, 2020). In 2010, the Tennessee General Assembly passed the First to the Top (FTTT) Act (Koedel et al., 2017), requiring all

school leaders to evaluate school personnel annually. The Tennessee Department of Education (TDOE) approved the Tennessee Educator Acceleration Model (TEAM). It specified that teacher evaluations should be based on three factors: classroom observations comprising 50% of the score, students' standardized test scores comprising 35%, and 15% would be determined by student achievement as measured by a student's proficiency on the state-mandated assessment (Tennessee Department of Education, 2021).

Society-at-Large

Public school districts began piloting the new teacher evaluation systems at the behest of state lawmakers and aligning with the RttT, requiring administrators to evaluate all certified teachers annually based on instructional practices and student academic progress (Paufler & Sloat, 2020). Although school leaders can integrate evaluative measures, such as classroom observations and other professional artifacts, linking VAMs to the overall evaluation process broadens teacher effectiveness. Further, the new teacher evaluation system placed heavy emphasis and pressure on teacher accountability (Jewell, 2017) for student academic achievement, impacting educators' morale and wellbeing, resulting in low self-efficacy and decreased motivation. Jewell (2017) stated that low-performing teachers could face public scrutiny in educating the students in the community. Teachers' overall ratings could become public records that could impact future employment because the evaluation system was linked to teachers' promotion, retention, or dismissal. Derrington and Martinez (2019) stated that the teacher evaluation process harmed trusting relationships between teachers and administrators and prompted teachers to leave the profession, leading to an alarming shortage of teachers in the public school setting.

Amidst several negative connotations of the teacher evaluation system overall, many studies have shown a positive impact of teacher evaluations. Mireles-Rios et al. (2019) emphasized that teachers became classroom management experts through teacher evaluation and delivered engaging and challenging tasks for all students in the classroom. Further, Cairreiro (2020) noted that the evaluation outcomes allowed teachers to accept and meet high expectations and promoted reflective discourses for teachers' pedagogical improvement (Williams & Hebert, 2020).

Theoretical Background

The modern teacher evaluation system was linked theoretically with self-determination theory (Ryan & Deci, 2017). Self-determination theory supports the idea of professional learning openly within and across social communities (King & Paufler, 2020). When educators face a heavy workload, teachers' psychological and physical welfare is at risk, and teachers' wellbeing could impact students' academic learning. The self-determination theory framework affirmed the fundamental psychological needs in motivation, improvement, fitness, competence, and autonomy (Ryan & Deci, 2017). With the present teacher evaluation system, Anderson et al. (2019) noted that the process situates teachers into circumstances that negatively impact their well-being, and the evaluation system restrained teachers' autonomy from educating the students. Classroom educators experienced low levels of choice when significantly regulated by curriculum requirements, principal supervision, and administrative practices (Watt & Richardson, 2015). In theorizing the teacher evaluation system with self-determination theory, intrinsically motivated teachers enjoyed teaching and delivered exciting tasks, which resulted in students' academic achievement. From the self-determination perspective, classroom educators were motivated to provide quality education to students since teachers naturally become adept

and self-sufficient. Amidst ongoing criticisms and requests for teacher evaluation reform from the NCLB legislation to the RttT initiative, policymakers continuously strive to produce an objective evaluation system that measures teacher performance and helps teachers grow professionally (Hess & Noguera, 2021). Embracing high but complex standards to evaluate teachers' accountability, policymaking constituents might consider employing the evaluation process through self-determination theory to protect all public school teachers' physical and psychological needs (Wieczorek et al., 2022). Since teachers are explicitly affected by the evaluation system outcomes, educational advocates claim that teachers should participate in redesigning the teacher evaluation system. A lack of consideration in including teachers' voices in redesigning the evaluation system could lead to a never-ending debate about a suitable evaluation system to evaluate teacher quality (Hess & Noguera, 2021).

Problem Statement

Numerous studies in appraising teacher quality have documented teachers' nervousness and anxiety (de Lima & Silva, 2018; Gitomer, 2019; Woolfolk & Usher, 2022) when policies overlook teachers' professional values and competence. Many scholars have argued that teachers should be involved in the design and implementation of their evaluation system (Carreiro, 2020; Finster & Milanowski, 2018; Reddy et al., 2018; Warren & Ward, 2018), while others such as Kim and Youngs (2016) resisted such calls because teachers' perceptions about the teacher evaluation policies are sometimes in conflict with teachers' beliefs. To ensure that teachers grow professionally, they must perceive the multiple measures of accountability as fair, valid, and reliable (Finster & Milanowski, 2018; Reid, 2020), and they should be allowed to choose how they will be evaluated and to determine the areas of refinement for personal growth (Carreiro,

2020). Whether or not teacher involvement leads to more accurate evaluations or more positive reception of the evaluation by the teachers is yet to be seen.

Despite much empirical research about valid measures and various models used in the evaluation system (Amrein-Beardsley & Holloway, 2019; Close et al., 2020; Xu et al., 2016), few studies exist about teachers' perceptions of the evaluation context and the evaluation system's overall quality (Warren & Ward, 2018). Additionally, Skedsmo and Huber (2018) noted that the results of empirical research favored the evaluators and that direct inputs from teachers were limited—indicating a withdrawal of the teachers' voices from evaluation systems. According to Ford and Ware (2018), refocusing the design and implementation of teacher evaluation would begin restoring working conditions in schools that are more conducive to meeting the psychological needs of teachers. Considering teachers' voices, the evaluation system becomes credible for assessing teachers' effectiveness, and teacher evaluation matters because teachers matter (van der Steeg & Gerritsen, 2016; Xu et al., 2016). In Tennessee, a study's findings showed concern regarding value-added assessments (TVAAS) and the validity of providing student growth scores from the state achievement measures using a formula unknown to the public (Davis et al., 2016). While Derrington and Martinez (2019) briefly explored Tennessee's teacher evaluation perceptions, the survey responses and self-reported data were limited to the high school teachers' perceptions. Because the survey was administered at one point in time, it is outside of the scope of the research to determine if teacher perceptions have changed over time; thus, further study is vital to determine if there are differences in perceptions between elementary, middle, and high school teachers. The problem is that the literature has not fully addressed the need for teachers's perspectives on whether or not the current teacher

evaluation system will positively affect teachers' professional practices and whether or not sustaining the current evaluation system is practical.

Purpose Statement

This quantitative, predictive correlational study aims to determine if a predictive relationship exists between teacher perceptions of Tennessee's teacher evaluation and the linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high). The criterion variable of teacher perceptions of the Tennessee teacher evaluation refers to the teacher's individualized thoughts about the participants' credibility of analysis of the state's evaluation system (Derrington & Campbell, 2019). The three categorical predictor variables will include gender, years of experience, and school level taught by the respondents. The predictor variable gender refers to the possible response of female or male (Linqvist et al., 2020). The years of experience pertain to the number of years the participant has taught: 0 – 3 years (beginning), 4 – 5 years (transitioning), and more than 5 years (experienced) (Graham et al., 2020), while the school level refers to the respondent's instructional level, either in elementary, middle, or high (Dogan et al., 2021). The participants of this study will be randomly selected from elementary, middle, and high school teachers in one of the school districts in southwestern Tennessee, which adopted the current evaluation system in 2011 for measuring teacher performance. All participants are certified teachers who are subject to evaluation annually.

Significance of the Study

Across the United States, there is a shift to teacher accountability, which has led to the widespread adoption of formative and constructive teacher evaluation systems (Williams & Hebert, 2020). The state of Tennessee adopted a rigorous evaluation design linked to the First to the Top Act (FTTT) bill, requiring a thorough annual evaluation of all personnel (Close et al.,

2020; Derrington et al., 2021; Ford & Hewitt, 2020; Koedel et al., 2017) to determine highly-qualified teachers based on evaluation ratings. Under FTTT, the teacher ratings are associated with compensations, promotion, tenure, and certification decisions (Amrein-Beardlesy, 2020; Koedel et al., 2017), prompting teachers' concerns. This present study is an essential component of the theory of self-determination. Self-determination theory posits the fundamental psychological needs in motivation, development, competence, relatedness, and autonomy (Ford, 2019; Fradkin-Hayslip, 2021; Ryan & Deci, 2017). According to Ford et al. (2018), teachers are self-determined to improve their pedagogical skills by incorporating the feedback received after the evaluation.

Examining teachers' voices about the Tennessee teacher evaluation system could help school leaders understand why there is a high teacher turnover rate in the state and the impact of teacher evaluation on job satisfaction in terms of teacher gender, years of experience, and school-level taught, allowing policymakers to comprehend why teachers leave the profession. Smith and Holloway (2020) mentioned how holding teachers accountable for students' standardized test scores damages job satisfaction. Paufler and Sloat (2020) stated that the evaluation system is trustworthy when stakeholders respect educators as professional resources rather than unvalued subjects, so the process must understand teachers’ perspectives. This study will contribute to the limited existing research on teachers' perceptions of the current Tennessee evaluation system. Moreover, examining teacher perceptions could influence teachers' engagement in the revamping process, ensuring fair, accurate, reliable, and valid evaluation components and implementation, resulting in a teacher evaluation design that better enhances teacher professional practice and growth. Further, because of TEAM's components, teachers in Tennessee cannot be evaluated as a

whole. Some teachers do not teach tested subjects (e.g., music, art, and P.E.), and others teach at a school level where the value-added measures are not applicable (e.g., pre-K and K-3).

Research Question

The following research question will drive the study:

RQ: How accurately can teacher perception of Tennessee's teacher evaluation system be predicted from a linear combination of teacher gender, years of experience, and school-level taught (elementary, middle, high) for Tennessee educators?

Definitions

1. *Gender* – Gender refers to the characteristics of females and males (Lindqvist et al., 2020)
2. *Self-determination theory* – Self-determination theory presents motivation as a fundamental psychological need that occurs at the nexus of competence, relatedness, and autonomy (Ryan & Deci, 2017)
3. *School level* – School level refers to the teacher's instructional level, either in elementary, middle, or high school (Dogan et al., 2021)
4. *Student achievement* – Student achievement is a proficiency outcome of student learning and academic performance across tested subject areas such as English and Language Arts, mathematics, and science (Alexander et al., 2017)
5. *Teacher evaluation* – Teacher evaluation is the process of collecting data to measure teacher performance for the intentions of school decision-making related to teacher outcomes (Holloway, 2019)
6. *Teacher perceptions* – Teacher perceptions are teachers' individualized thoughts and credibility of analysis on the teacher evaluation system (Derrington & Campbell, 2019).

7. *Teacher quality* – Teacher quality is a teacher characteristic that influences student achievement (Kolman, 2017)
8. *Tennessee Educator Acceleration Model (TEAM)* – The Tennessee Educator Acceleration Model is one of the four tools used to measure teacher performance annually in the state of Tennessee (Tennessee Department of Education, 2021)
9. *Tennessee Value-Added Assessment (TVAAS)* – Tennessee Value-Added Assessment is an assessment used by the state of Tennessee to measure students' performance through standardized test scores (Tennessee Department of Education, 2021).
10. *Years of Experience* – Years of experience pertains to the number of years the teacher teaches (Campbell, 2020)

CHAPTER TWO: LITERATURE REVIEW

Overview

The purpose of this literature review is to present the vital elements of the teacher evaluation system in the United States and to review the teacher evaluation system implemented in the state of Tennessee. The chapter opens with the theoretical framework. This study is grounded first in Deci and Ryan's (1985) self-determination theory, which posits the fundamental aspects of intrinsic and extrinsic motivation. A thorough review of literature pertinent to the teacher evaluation system in public schools and the teacher evaluation system in Tennessee completes the chapter and ends with a summary.

Theoretical Framework: Self-Determination Theory

Although widespread research exists about teacher evaluation systems as a tool to support instructional practices, increasingly, the scholarship has pivoted toward a focus on accountability within teacher evaluations (Cairreiro, 2020; Mireles-Rios et al., 2019; Paufler & Sloat, 2020). The current study is based on a theory that educational stakeholders support teachers during the evaluation process regardless of teachers' gender, years of experience, and school-level taught. Teachers' need for support regarding time, autonomy, and professional development aligns with self-determination theory, which foregrounds the psychological needs of competence, autonomy, and relatedness (Ford, 2018).

The historical background and principles of self-determination theory inform essential functions in education, such as the teacher evaluation system. First studied in the 1970s, self-determination theory evolved from studies comparing intrinsic and extrinsic motives and was first introduced in 1985 by Edward Deci and Richard Ryan (Ryan & Deci, 2017). Self-determination theory asserts the fundamental psychological needs in motivation, development,

wellness, competence, relatedness, and autonomy (Ryan & Deci, 2017). Competence pertains to the person's urge to strengthen existing capabilities in hopes of future accomplishment (Ford, 2018). Relatedness refers to the need to be connected and cared for by colleagues and allocate a sense of acceptance to others in the community (Ford, 2018). Autonomy touches on actions for which impetus is generated not from the need to satisfy the external expectations but rather from self-advocate values and beliefs (Ford, 2018). The self-determination theory emphasizes the internationalization of social morals and that society contains various extrinsic rewards and controls that may not be suitable for the individual's quest for self-determination (Schunk, 2020). It circulates four mini-theories: cognitive evaluation, organismic integration, causality orientations, and basic needs that correlate with motivating and achieving health and wellbeing (Ryan & Deci, 2017). In applying this to the educational setting, self-determination theory supports teacher motivation and psychological health (Cuevas et al., 2018). It presumes that humans are innately curious, physically active, and intensely social beings (Ryan & Deci, 2017).

There are various reasons for teachers' motivation to perform well in their profession, such as to serve the students better and to provide quality education. Incorporating a framework of self-determination theory is crucial to ensure that the teacher evaluation system produces a balanced approach in terms of teachers' needs. Ford and Hewitt (2020) asserted that self-determination theory could help explain how evaluation feedback motivates teachers to improve or grow professionally because motivation drives teachers to become competent and autonomous. In a self-determination view, wellness is accompanied by a thriving characterized by vitality, apprehensions, access to, and the freedom to exercise one's capacities and proper self-regulation (Ryan & Deci, 2017). The teachers' self-determination to reach goals in refining educational pedagogies could lead to satisfaction and self-fulfillment. Satisfaction is assumed to

be a necessary condition for human flourishing, and frustration is injurious to well-being (Ryan & Deci, 2017).

Various studies have emphasized the usefulness of self-determination theory (Cuevas et al., 2018; Ford et al., 2018; Ford, 2019). In Ford et al.'s (2018) study, two of the mini-theories of self-determination theory, namely the basic psychological need theory (BPNT) and organismic integration theory (OIT), were linked to examine the influence of supportive evaluation practice on teacher job satisfaction. The psychological need theory predicts that intrinsic motivation will remain intact when psychological conditions of competence, autonomy, and relatedness are met (Ford et al., 2018). Current research supports that autonomy is connected to competence, relatedness, and job satisfaction, resulting in high teacher retention rates and student academic outcomes (Fradkin-Hayslip, 2021). The BPNT is relevant in understanding teacher motivation because of the uniqueness of the teaching profession, and creating conditions that reinforce and activate existing intrinsic motivation is one of the motivational strategies that policymakers should approach accountability (Ford et al., 2018).

However, Ford et al. (2018) asserted that the prevalence of extrinsic rewards and punishments tied to teacher performance creates a fundamental misalignment between the motivational context and the approach. This assertion was supported by Saeki et al.'s (2020) study of the influence of test-based accountability policies that the misalignment manifests in such ways as increased stress, decreased satisfaction, and increased teacher attrition. Further, the organismic integration theory helps explain how moving individuals from control to autonomy yielded substantially greater positive behavioral and psychological outcomes (Ryan & Deci, 2017). Ford et al.'s (2018) findings were rooted in SDT and provided evidence that there were

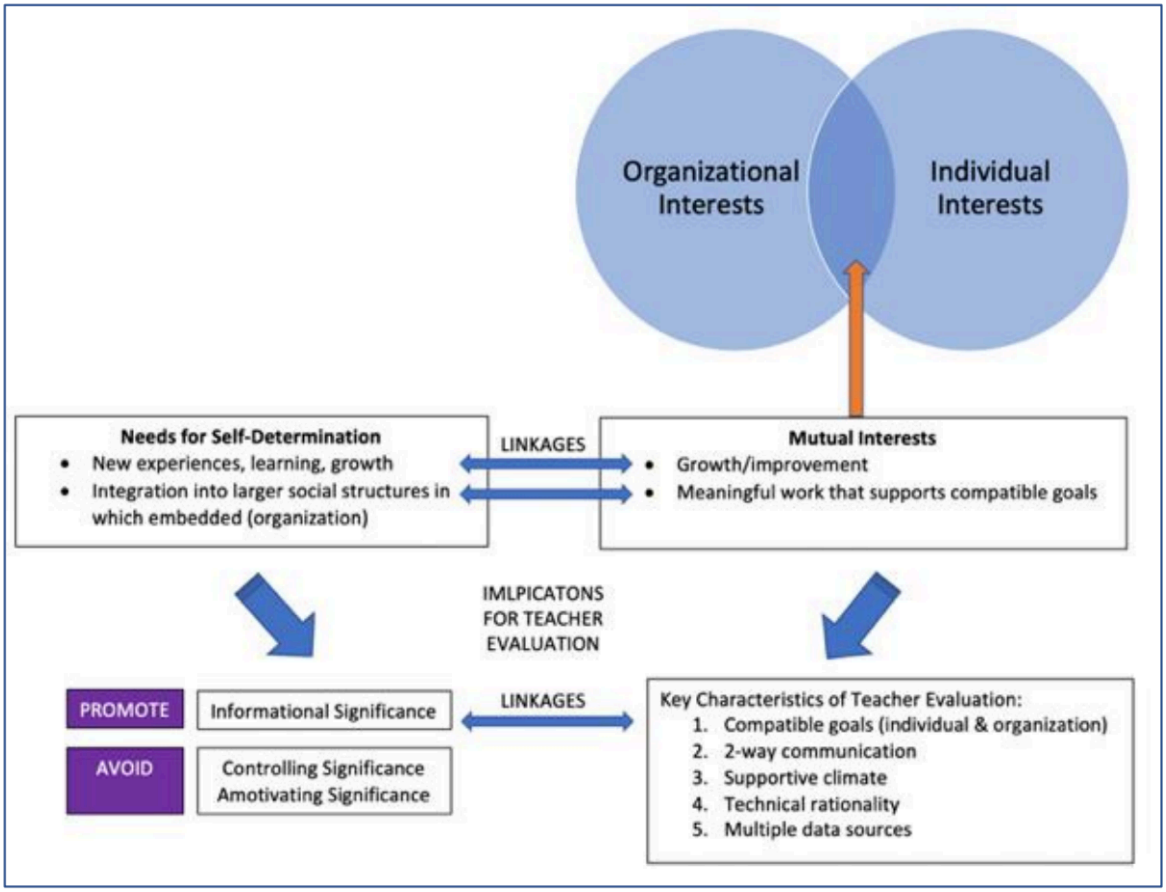
gains in job satisfaction among educators who experienced more supportive teacher evaluation processes.

Evidence suggests that teacher evaluation systems should be rooted in a thorough assessment of instructional practice and not simply based only on performance measures such as student test scores (Grissom & Youngs, 2016). Further, Stronge (1995) contended that the primary goals of evaluation—accountability and professional growth—are not only compatible, but connecting them to a unified teacher evaluation is desirable and that the teacher evaluation system should fulfill the needs of both the individuals and the organizations. From the lens of SDT, teachers could hone pedagogical skills through professional developments (PDs) to enhance classroom atmospheres, including all students' wellbeing and academic success. Ford (2019) strongly affirmed that all areas of SDT should be incorporated during professional development as teachers may need support in terms of competence and autonomy. In line with SDT, Cuevas et al. (2018) anticipated that autonomous motivation positively predicted vitality and negatively predicted exhaustion. The findings have shown that teachers who taught for pleasure or professional growth were likelier to report feeling full of life and less likely to feel exhausted (Cuevas et al., 2018). Conforming to SDT, Ford and Hewitt (2020) believed that using rewards and punishments to transform behavior was an effective way to motivate people. Additionally, integrating the theoretical framework of SDT and linking it with Stronge's improvement-oriented model could develop a significant lens through which to evaluate existing and emerging teacher evaluation systems (Ford & Hewitt, 2020). Figure 1 demonstrates that the focus of teacher evaluation should be the area of overlap between organizational and individual interests, which are growth or improvement, and works that support mutual goals (Ford & Hewitt, 2020). The choice of incorporating SDT in the study is to explain how teachers' views on

the teacher evaluation system could open pathways to self-motivation and autonomy in improving instructional capabilities.

Figure 1

Integrated Framework of Self-Determination Theory & Stronge's Improvement-Oriented Model



Note: The integrated framework: Intersection of SDT and Stronge's Improvement-Oriented Model. From "Better Integrating Summative and Formative Goals in the Design of Next Generation Teacher Evaluation Systems," by T.G Ford and K.K. Hewitt, 2020, *Policy Analysis Archives*, p. 7 (https://doi.org/10.14507/epaa.28.5024).

Related Literature

Over the past decades, each state has freely adopted its teacher evaluation system to measure teacher effectiveness (Kraft & Gilmour, 2017; Wieczorek et al., 2022). In Tennessee,

the current teacher evaluation system is driven by a competitive grant, which incentivizes states to reform K-12 education around low-performing schools and enhances teacher and principal effectiveness (Ballou & Springer, 2015; Reid, 2020; Sartain & Steinberg, 2016). One of the goals of the teacher evaluation system is to hold public accountability for both teachers and the schools. Under the new evaluation system, various measures are considered, including the use of mandated classroom rubrics during observations that entail instructional improvement (Ballou & Springer, 2015) and the incorporation of value-added measures (VAMs) to inform student growth and achievement (Sartain & Steinberg, 2016), teacher self-reports, student and parent-teacher surveys, and other classroom artifacts (Lavigne & Chamberlain, 2017).

The Teacher Evaluation

The practice of teacher evaluation has been pervasive in education since the late 2000s, with most states adopting new high-stakes teacher evaluation systems between 2011 and 2016 (Kraft et al., 2020). Holloway (2019) defined teacher evaluation as measuring teacher performance. Although Carreiro (2020) specified that teacher evaluation is an act of compliance that evaluators must accomplish according to predetermined timelines, it offers teachers endless opportunities for self-reflection and professional growth. The growing body of research categorized teacher evaluation as formative and summative (Ford & Hewitt, 2020). However, Mireles-Rios et al. (2019) argued that the recent teacher evaluation focused only on curriculum and grade-level summative standards assessments (Kraft & Gilmour, 2017). Regardless of the type of evaluation employed to assess teachers, the intent of the evaluation systems should remain intact (Mireles-Rios et al., 2019).

The existence of the teacher evaluation system is accompanied by targeted goals to aid teachers professionally and personally. With the passage of the Every Student Succeeds Act

(ESSA), states have been given significant authority to create an evaluation system to serve the community's needs better and adhere to state standards (Young et al., 2017). Initially, the teacher evaluation goals were to measure teacher competence and enhance teacher quality (Mireles-Rios et al., 2020). On the other hand, since there is a greater sense of teacher accountability, Williams and Hebert (2019) suggested that state lawmakers should reevaluate the purpose and components of the teacher evaluation to ensure clarity, validity, and reliability.

Elements of Teacher Evaluation Systems

The teacher evaluation system has been the center of much research, but additional studies are essential in examining teachers' perceptions of the evaluation system. Empirical research showed that teacher evaluation systems incorporating structured classroom observations and detailed feedback on instructional practice could improve student achievement (de Lima & Silva, 2018; Sartain & Steinberg, 2016). Measuring teachers' attitudes and beliefs about their evaluation is critical for assessing teacher evaluation's cogency and improving their implementation (Reddy et al., 2018). Teachers' perspectives and concerns regarding the current evaluation systems must be understood and considered if policymakers can seize the opportunity to redesign the systems (Paufler & Sloat, 2020) to have tremendous potential for teacher quality and positive student learning outcomes. Several factors could potentially shape how educational stakeholders redesign the teacher evaluation system (Reid, 2020), such as the characteristics of a teacher, the evaluator, the evaluation process, the feedback, and the evaluation context.

The Characteristics of a Teacher

Teachers were facilitators of learning, providing students with a quality education and opportunities to improve academically. Teachers' attributes contribute to the academic success of students. Teachers' strengths of professional practice, the ability to take risks, willingness to

experiment, openness to criticisms and change, and knowledge of curriculum content and instructional strategies could be valuable assets for effective teachers. However, Schuls and Flores (2020) noted that teaching is one of the most stressful professions, leading to difficulty in retaining teachers. Even retaining novice teachers became a crisis within the public schools in the country (Warsame & Valles, 2018). However, district leaders could distinguish between effective and ineffective teachers through a teacher evaluation system. Martinez et al. (2016) stressed how evaluation systems become the tools to determine struggling and high-performing teachers. For example, some of the components of the teacher evaluation system include classroom observations to assess instructional practices, student surveys to assess students' perceptions of instruction and classroom climate, standardized tests to evaluate subject matter or pedagogical knowledge, and finally, the incorporation of VAMs to evaluate student learning and growth (Martinez et al., 2016).

Teachers' personalities are connected with teaching efficacy and motivation. In Kim et al.'s (2019) meta-analysis on the effects of personality on teacher effectiveness and burnout, the results indicated that openness, conscientiousness, extraversion, agreeableness, and emotional stability are the big five domains linked to teaching efficacy. Khalilzadeh and Khodi's (2018) findings were similar; they determined that teachers' personality traits, such as honesty and fairness, positively impacted students' intrinsic motivation and learning, while teachers' extraversion personality traits negatively affected students' intrinsic motivation. In other words, the more the teacher was active, warm, and sociable, the less the students desired to learn and explore (Khalilzadeh & Khodi, 2018). A quantitative study conducted in one of the school districts in Texas revealed that teacher personality, particularly grit, predicted teacher retention, as teachers who have grit in teaching were 17% less likely to leave the teaching profession in the

district (Robertson-Kraft & Zhang, 2018). Further, the results indicated that the lowest retention rate of 75.9% was among novice teachers compared to veteran teachers, with a retention rate of 85.2%, meaning that novice teachers had higher turnover rates than experienced teachers (Robertson-Kraft & Zhang, 2018). Contrary to personality traits that contributed to teacher effectiveness, Zayac et al.'s (2021) research examined the qualities and behaviors of ineffective teachers. It determined that students and faculty members perceived disrespect as the first quality of ineffective teachers and that a weak rapport signaled ineffective educators.

The Evaluator

One of the most valuable individuals for the success of the school organization is the school administrator. In addition to the instructional leadership roles, administrators are promoters of professional development and are influential in implementing the evaluation system (Donaldson & Mavrogordato, 2018; Kraft & Gilmour, 2016). School principals conduct evaluations through classroom observations formally or informally throughout the year and provide feedback for refinement and growth. Since principals evaluate teachers, several states require school evaluators to take and successfully pass the evaluator proficiency exam (Dodson, 2018). Since the new evaluation systems were based on Charlotte Danielson's Framework for Teaching (FFT), principals provided some insights about the evaluation system. The results suggested that most principals were somewhat satisfied with the FFT but desired to overhaul the method used to input the teacher evaluations, eliminate the use of value-added measures, reduce the time and paperwork required, and want more training on the FFT (Dodson, 2017). Research showed that administrators embraced a legal and moral obligation to all school staff and ensured to assist low-performing teachers in enhancing instructional practices that include cognitive, relational, and organizational aspects (Donaldson & Mavrogordato, 2018; Rigby et al., 2017).

Ensuring that evaluators promote professional development, previous research on teachers' perceptions of implementing the evaluation and whether or not the administrators are supportive and encouraging findings suggested that administrator ratings decreased the likelihood of teachers leaving the school (Robertson-Kraft & Zhang, 2018).

Principals embraced more responsibilities concerning the expectations of the teacher evaluation. While principals exhibited confidence in conducting classroom observations, administrators expressed less certainty in distinguishing teaching behaviors connected with student academic mastery (Lavigne & Olson, 2020). Further, principals' comprehensive judgments were more robustly influenced by teachers' instructional practices and other significant contributions to the school than by teachers' influence on student academic growth (Briggs & Dadey, 2017).

The Evaluation Process

The dissatisfaction with implementing the teacher evaluation system prompted policymakers to redesign it. In quantitative and qualitative analyses of the teachers' and administrators' attitudes and beliefs of the teacher evaluation system in New Jersey, teachers highly perceived communication/collaboration (23.53%) regarding the evaluation process was the most helpful, while the evaluation method not susceptible to practical application (21.15%) was the least beneficial (Reddy et al., 2018). Even after the reform, teachers perceived devalued or lack of control over their value-added measures (VAMs) scores and were unmotivated to learn about the VAMs (Presley et al., 2018). Moreover, teachers became the subject of undermining intentions and exacerbating inequities (Hewitt, 2015). Some educators lambasted the evaluation system for administrator subjectivity, in which principals rated female, experienced, and elementary educators more highly than male, novice, and high school teachers, and the

implementation process contradicted the primary goals to enhance teacher performance (Williams & Hebert, 2020; Wind et al., 2019).

School administrators have a crucial responsibility in evaluating teachers. Research findings have illustrated that school administrators perceive that conducting evaluations takes time, such as spending long hours on the job and delegating more tasks to other members of the leadership team and school staff (Derrington & Campbell, 2018; Lavigne & Chamberlain, 2017; Lavigne & Olson, 2020), an increase of workload, and a lack of training and time for implementation, which prompted principals to seek assistance from others about clerical and administrative tasks (Dodson, 2015; Flores & Derrington, 2015; Lochmiller & Mancinelli, 2019). School leaders perceived that mandated policy on teacher evaluation systems created tensions within the relationship between teachers and administrators. They caused frustrations among principals, such as teachers' hesitations to invite administrators into the classroom, an absence of informal communication, and trust issues (Flores & Derrington, 2015; Neumerski et al., 2018). Distrust in the evaluator's credibility to conduct observations could lead to teachers discrediting the entire teacher evaluation process (Ford et al., 2017). A study in Illinois suggested that administrators expressed confidence in their abilities to engage in formal observations and provide feedback to improve instruction; however, principals did not value student achievement data as part of the teacher evaluation (Lavigne & Olson, 2020). Neverthess, most administrators perceive student data as essential measures of teacher effectiveness (Lavigne & Chamberlain, 2017), but it is through classroom observations that school principals emphasize most valuable processes in assessing teachers' instructional effectiveness (Lavigne & Olson, 2020).

Conversely, Paufler and Sloat (2020) found that some administrators believed in a satisfactory alignment between ideal and actual teacher evaluation intentions. Supported by

Flores and Derrington (2015), the evaluation system's objective is to improved all teachers' professional growth and development. The evaluation process promotes teacher reflection as it is grounded in the evidence of objectivity and transparency (Lewis et al., 2020). Evidence suggested that teachers believed that the school principals were transparent and sincerely communicated effectively before and after the process, particularly in providing feedback and thoroughly explaining the evaluation process to all staff (Reid, 2020).

The Feedback Received After the Evaluation

Teachers formally received feedback from the evaluator as part of the evaluation process. The feedback is intended to provide insight into the strengths and areas for refinement (Williams & Hebert, 2020) that could continuously develop teachers' practices, especially in ensuring safety, providing challenges, and increasing engagement (Mireles-Rios et al., 2019) for all students. Evaluators providing feedback could communicate with teachers and administrators, leading to a trusting relationship. Carreiro (2020) noted that relationship matters to experience success with the teacher evaluation process and that feedback and support could allow teachers to embrace and meet high expectations. Teachers fully appreciate feedback if an evaluator has knowledge of the content, instructional skills, and the evaluation system as a whole (Liu et al., 2019). Research scholars argued that either formally or informally, constructive and individualized educator feedback from classroom observations is a promising teacher growth technique (Stronge & Tucker, 2017). Teachers who regarded principals as available and willing to discourse challenges in a supportive manner were more likely to buy into the evaluation system, engage in the process, and find overall job satisfaction (Ford et al., 2018; Liu et al., 2019; Robertson-Kraft & Zhang, 2018).

Although several studies highlighted the benefits of providing feedback to teachers, research showed how teachers negatively perceived the input from the evaluators. For instance, Ridge and Lavigne (2020) supported the assumptions made by Flores and Derrington (2017) that 55% of the administrators felt that time was limited and failed to provide rich feedback to teachers because they spent most of the time writing the evaluation. One of Kraft and Gilmour's (2016) study findings supported the assumption that tasking principals with primary responsibilities for conducting evaluations could result in an unintended consequence that undermines the quality of teacher feedback. Administrators struggled to provide content-based feedback after evaluating teachers whose instructional content expertise differs from evaluators' expertise (Ridge & Lavigne, 2020). Some teachers did not perceive the feedback received as constructive in helping refine instruction (Liu et al., 2019). Similarly, a current study in Louisiana indicated that transparent and actionable feedback was often missing as part of the evaluation protocols (Ford et al., 2017). In addition, while the principals felt confident that feedback could improve instructions, administrators felt uncertain about providing constructive criticism to low-performing and novice teachers (Lavigne & Olson, 2020).

The Evaluation Context

The teacher evaluation contexts were developed based on the amount of time spent in the evaluation process, the district's purpose and policy regarding evaluations, and the evaluation curriculum standards used. Shubert (2018) asserted that evaluation should be a central thread of developing, implementing, and designing a curriculum. Further, Williams and Hebert (2020) stressed the significance of a well-rounded evaluation system comprised of evaluation components that produce effective teaching. In Tennessee, the features of the evaluation system involved observation rubrics used during observations. Part of the classroom observation rubrics

ensures teachers provide instruction based on the standards specified in curriculum maps. Rubrics could lend an air of transparency since the evidence consists of observable actions of the teacher and objectivity because the evaluators' conclusions do not rely on personal opinions but are grounded in the evidence (Lewis et al., 2020). However, the primary concerns of teachers are not on the contents or standards used during instruction but on the complexities and ambiguities of the rubric and a lack of confidence in the evaluators' ability to interpret the criteria used for scoring teachers (Campbell & Derrington, 2019) and used during the evaluation. In Kraft and Gilmour's (2017) study, the findings reported that nearly 70% of the sampled participants perceived that the evaluation rubric was a significant and positive improvement to the evaluation system. Furthermore, the administrators felt that the rubric ratings based on teacher observable practices could help educators understand why feedback was provided (Kraft & Gilmour, 2017).

Perceptions of Teacher Evaluation System

Since implementing the teacher evaluation system, there have been ongoing discourses on whether the evaluation system is significant. The teacher evaluation systems vary from state to state, often involving school administrators' judgments on the effectiveness level (Smith et al., 2020). One criterion of teachers in the evaluation process is the extent of the evaluators' experiences and expertise to carry out the evaluation system. Teachers believe the teacher evaluation system is correctly enforced when evaluators attend and participate in teacher evaluation training and comprehensively comprehend the teachers' strengths and weaknesses (Reid, 2020). Previous research had shown that the evaluators had not been trained in observing classrooms and failed to provide constructive feedback in a leadership context (de Lima et al., 2018), which was paralleled to Reddy et al.'s (2018 study that teachers identified collaborative communication and evaluation feedback as the most beneficial aspects of the evaluation process.

In Tuytens and Devos's (2017) research, findings indicated that teachers who received constructive feedback and actionable feedback from the school administrators were more likely to embrace instructional changes, which was supported by Derrington and Martinez's (2019) study results that 87% of teachers agreed that feedback would be utilized to change instructional practices.

Studies measuring teachers' quality through required classroom observations have documented nervousness and anxiety among teachers (de Lima et al., 2018) since the mandated classroom observation could relate to the issue of teacher autonomy and control, which is one of the components of self-determination theory. The evaluation process will be associated with skepticism (Cherry et al., 2017). For example, in Derrington and Martinez's (2019) study on teachers' perceptions about the effectiveness of Tennessee's evaluation system after five years of implementation, the findings showed a high percentage (68%) of teachers who disagreed with the evaluation process, specifically when using the observation rubric (Garret & Steinberg, 2015; Derrington & Martinez, 2019). Several research findings indicated that 46.2% of teachers perceived that the evaluation system negatively affected teachers' professional practice (Paufler & Sloat, 2020) and undermined their hopes that the new evaluation systems could help them develop as professionals (Kraft & Gilmour, 2017), and that school administrators rated female, veteran elementary teachers more highly than male, novice secondary teachers (Wind et al., 2019). Teachers freely expressed that the evaluation systems created stressful and burdensome experiences (de Lima et al., 2018).

Impact on Morale and Wellbeing

Promoting educators' wellbeing is vital in many aspects, as it positively correlates to students' learning outcomes and the teachers' decision to retain or leave the profession (Anderson

et al., 2019). With the new teacher evaluation demands, one of the inquiries associated with the evaluation system is whether the system supports teacher morale and wellbeing. Advocates of the new teacher evaluation systems asserted that better differentiation of performance and aligning consequences directly with outcomes would improve teacher efficacy and motivation (Robertson-Kraft & Zhang, 2018). Further, an evaluation system that aligns performance and rewards will entice talented individuals and positively impact the labor workforce (Robertson-Kraft & Zhang, 2018).

Due to the demands of holding teachers accountable for student learning success, Anderson et al. (2019) found that educators negatively perceived the teacher evaluation process, which could lead to unhealthy competition among teachers. Administrators strongly stressed that the evaluation system had resulted in a decline in morale, diminished autonomy, and decreased value (Paufler, 2018). Teachers perceived that the primary concern of teacher autonomy and control is not the observations but rather the reality that classroom observations are mandated and based on external criteria for practice and performance (de Lima et al., 2018). The evaluation system fundamentally focuses on teachers' competency and quality; it does not foster genuine professional learning and contributes to demoralization (Bradford & Braaten, 2018). Further, accountability and urgency for improvement remain substantial contributing factors to teacher demoralization (Bradford & Braaten, 2018).

Additionally, the evaluation process urges harm or a loss of trusting relationships with teachers and principals (Derrington & Martinez, 2019; Neumerski et al., 2018), perhaps due to subjectivity and biases, unfairness, and inaccuracy (Hewitt, 2015) since teachers have a little power of autonomy. The teacher evaluation impedes teachers' freedom (Anderson et al., 2019) to become creative and competent as a notion of self-determination theory. Teachers consider

observation feedback to foster self-efficacy and motivation to change instructional strategies for professional growth. However, Paufler and Clark (2019) mentioned that school principals reportedly struggle to furnish valuable feedback to teachers due to time, a lack of expertise in all areas, and insufficient training to improve teachers' instructional practices (Kraft & Gilmour, 2017). Most of the teachers suffered from increased anxiety and decreased teacher motivation, especially among educators who do not meet the expectations established by the school administration (Cuevas et al., 2018).

Pathways to Professional Development (PD)

One of the desired purposes of teacher evaluation is to open pathways for teachers to professional growth through a series of professional developments (PDs). Teacher quality assurance and professional development are the two central goals of the teacher evaluation system (Robertson-Kraft & Zhang, 2018). Through professional development, teachers have significant growth in formative instructional practices, preserve collaborative feedback, and assist students in monitoring academic progress (Liang et al., 2020). Research findings indicated that principals desired teachers to apply the knowledge and skills acquired from professional development consistently, and those professional developments should focus on exposing classroom educators to actual practice (Brown & Militello, 2016; Koonce et al., 2019).

The school principals become the proponents of professional development. Research confirmed the role of administrators in influencing student academic performance through professional development commitments (Kennedy, 2016). Additionally, school leaders who were knowledgeable, active, and engaged were more likely to accurately distinguish productive and high-quality professional developments, which comprised mentoring and individualization that could build relationships with teachers (Koonce et al., 2019). Further, teachers were self-

determined to engage in different workshops with various levels of content understanding to better serve students in achieving academic excellence (Gupta & Lee, 2020). Study results suggested a remarkable growth in productive instructional pedagogies when teachers engage in collaborative professional learning communities (Liang et al., 2020). Although principals aspired to provide relevant and purposeful professional development, there were limitations and barriers to creating meaningful PDs, including time and money, substitute teacher availability and costs, and poor communication, planning, and organization (Koonce et al., 2019).

Job Satisfaction

Educational stakeholders viewed teachers as direct promoters and advocates for student learning and academic success, and fulfilling such duties and responsibilities requires job satisfaction. In California, a study explored the perceptions of male teachers about career satisfaction; 96.7% of respondents strongly agreed that teaching elementary schools is a fulfilling career for men/men of color (Gaza, 2022). Research showed that 90% of school administrators reported utilizing student data about teacher recruitments, teacher evaluations, and teacher assignments (Cohen-Bogel et al., 2019). With the current teacher evaluation systems, several studies were conducted on job retention for classroom teachers. Texas, for instance, was one of the states that highlighted concerns about teacher retention in some school districts (Robertson-Kraft & Zhang, 2018). Following the implementation of the evaluation systems, research findings suggested that the current teacher evaluation system did not significantly impact teacher retention and that individual and school factors contributed to the variation of retention patterns (Robertson-Kraft & Zhang, 2018). High-performing teachers, measured by classroom observations and student test scores, are less likely to leave schools with effective principals. In contrast, low-performing teachers who received low classroom observation scores are likelier to

leave the school regardless of whether the teachers have low or high value-added data (Hanushek et al., 2016). Research suggested that more effective administrators can retain teachers because of positive climates and reinforce teachers with beneficial opportunities for professional growth that could hone teaching practices that lead to greater job satisfaction (Grissom & Bartanen, 2019; Kraft et al., 2016).

Validity of the Teacher Evaluation System

There was an increase in concern regarding the validity of the current teacher evaluation system. Almost all states mandated multiple measures (Martinez et al., 2016) to assess teachers' performance, such as standardized classroom observations, student achievement data, and teacher portfolios (Xu et al., 2016). Some apprehensions about the validity and reliability of particular measures used to assess teachers' performance have tremendously affected school administrators' views and practices (Paufler & Clark, 2019). Although there are tools involved in assessing teacher competence, the evaluators' interpretations of the criteria used in the evaluation are critical. An in-depth understanding of the criteria used for evaluation is crucial because these factors affect the validity and reliability of the evaluation results (Williams & Hebert, 2020). Further, teacher evaluations' inclusion of student achievement data frequently results in inaccurate and inconsistent evaluation ratings (Paufler & Clark, 2019), which paralleled Kraft and Gilmour's (2017) study findings that among 24 states that adopted the evaluation system, the percentage of teachers rated unsatisfactory varies, such that some of the states like Hawaii (1%,) New Mexico (28.7%), Georgia (6%), Massachusetts (9%) rated teachers below proficient, while in Tennessee, 62% rated above proficient. While teacher evaluation matters, Xu et al. (2016) noted that if an evaluation system does not exemplify validity, its worth for accountability, professional growth, and any allowable purposes is hugely compromised.

Educators' participation in the implementation of the teacher evaluation systems is essential to determine how teachers perceive the overall process of the evaluation system. In a study conducted by Pizmony-Levy and Woolsey (2017) about teachers' engagement in high-stakes teacher accountability policies in the state of New Jersey, 81.8% of the teachers disagreed with the statement that the rollout of the evaluation system was done at a comfortable pace for effective implementation, 77.9% disagreed that the new laws and changes transpired in teacher evaluation system are effectively communicated by the state's department of education, and 95% endorsed the notion that the evaluation system should be developed with significant input from educators. Furthermore, other countries such as Korea implemented three systems, namely, the performance rating for promotion, the teacher performance-based pay, and the teacher evaluation for professional development, but the evaluation results obtained from the three systems have often been inconsistent, leading to stakeholders' urgency to develop a single comprehensive teacher evaluation system in the country (Choi & Park, 2016).

Value-Added Measures (VAMs)

As part of incentivizing public schools, policymakers demand value-added measures (VAMs) in the teacher evaluation system. The goal of incorporating VAMs into teacher evaluation is to hold teachers accountable for student learning based on student performance standards (Williams & Hebert, 2020). However, VAMs can be unreliable, invalid, biased, not transparent, and unfair (Close et al., 2020). Most states established the educator evaluation system based on student test scores. In 2015, due to the negative impact of using VAMs, the new wave of federal accountability policy, the Every Student Succeeds Act (ESSA), was passed to eliminate the mandate that the teacher evaluation system used student test scores (Paufler & Sloat, 2020). However, many states integrated VAMs into the teacher evaluation system and

mired debates on assessing teachers' performance. While educational accountability policy rests heavily on the assessments used to influence teaching, learning, and school improvement (Brewer et al., 2015), teachers expressed anxiety and hope for high students' test scores (Presley et al., 2018). Smith and Kubacka (2017) explored how student test scores are increasingly utilized to assess teachers' performance and the basis of teachers' employment.

The linking of VAMs to teacher performance to hold teachers accountable for student learning raised opposition in several states. Geiger et al. (2020 provided an overview of 15 teacher evaluation lawsuits throughout the U.S. (Florida = 2, Louisiana = 1, Nevada = 1, New Mexico = 4, New York = 3, Tennessee = 3, and Texas = 1). Paige (2020) also outlined research methods and a law and policy framework to review lawsuits concerning VAMs to refine key principles for state and local policymakers. However, after ESSA, a study has shown a decrease in states using VAMs as part of the teacher evaluation system. Using Collins and Amrein-Beardsley's 2014 to 2018 comparison map in 51 states, Close et al. (2020) found that the number of states using VAMs decreased from 21 to 15 (41% to 29%), states that did not incorporate VAMs increased from 7 to 22 (14% to 43%), states that use VAMs formatively increased from 0 to 3 (0% to 6%). States reporting local control increased from 3 to 10 (6% to 20%). In a study conducted by Lee (2018) examining the relationship between teacher quality and teacher effectiveness (measured by VAM), the results indicated that students whom high-performing and qualified teachers had taught tended to have a positive relationship with students. Advocates of VAM found that high-value-added teachers provided a more equitable means of measuring growth and improved student achievement three times more than reducing the class size and tended to have a positive influence on students' behavior, attendance, and graduation (Nye et al., 2004; Sanders, 2000).

Teacher Evaluation: The Role of Gender, Years of Experience, and School-Level Taught

The role of gender in the educational system is a significant factor in maintaining gender inequality in society. Gender is a concept that emphasizes the psychological and sociological aspects of being a man and a woman (Dökmen, 2017). As reported by Jerrim and Sims (2019), women hold 66% and 57% of the teaching jobs in primary and secondary education worldwide. Male elementary teachers viewed teaching as equally suited for both men and women, and both could be role models and make a difference for students (Meader & Larwin, 2022). In an educational setting, gender bias refers to an attitude either in favor of or against a particular gender, which unconsciously influences the evaluation rating of teachers (Dökmen, 2017).

Several studies investigated the existence of gender bias in a school setting in a context of teacher evaluation, and researchers found that men teachers, particularly teachers of color, received statistically lower observation ratings than women and white teachers (Campbell & Ronfeldt, 2018; Drake et al., 2019; Jiang & Sporte, 2016) even after adjusting the teachers' instructional practices for teacher quality. Moreover, male teachers rated below proficient yearly (Bocala et al., 2016). Farinde-Wu and Fitchett's (2018) study suggested that black female teachers teaching in urban, non-charter schools expressed job satisfaction and acknowledged that administrators' support was evident. However, teachers with low observation ratings were found to exit the teaching profession across the public school setting (Drake et al., 2019). Further, even when both black and white teachers were similarly effective in the classrooms, evaluators favorably rated white female teachers more than black female teachers (Campbell, 2020), which negatively influenced black women educators' intentions to remain in the teaching profession (Farinde-Wu et al., 2016).

As new teachers enter the profession, educators seek to develop to become highly effective (Ladd & Sorensen, 2017). Several researchers defined novice teachers as certified educators who have taught for three or fewer years, while veterans or experienced teachers are certified educators who have taught for more than three years (Donahue & Vogel, 2018; Smith et al., 2020). Research has indicated that novice teachers are less effective than veteran teachers (Conley et al., 2019; Drake et al., 2018; Redding & Henry, 2019). A study in Colorado highlighted the usefulness of the state's teacher evaluation system by providing novice teachers with a guide to the expectations of good teaching. In contrast, veteran teachers perceived that the state's evaluation system provided a framework to hang the instruction (Donahue & Vogel, 2018). While research has shown how novice teachers look for instruction improvement (Ladd & Sorenson, 2017), several researchers asserted that high turnover rates occurred in the first three years of teaching (Papay et al., 2017; Redding & Henry, 2019).

The working environment is essential for teachers to retain in the teaching profession. Elementary teachers perceived that losing control of the student achievement data and a lack of information regarding the value-added measures led to feelings of hopelessness for future scores (Pressley et al., 2018). Middle school teachers, on the other hand, expressed concerns about the evaluation rubrics used during evaluation because the results could impact the relationships with the principals (Derrington, 2019). However, secondary teachers perceived the importance of the principals' capacity to evaluate teachers on time and how effective principals as evaluators of teaching (Frasier, 2021). Further, high school teachers implied that the reformed evaluation system did not effectively provide learning opportunities for secondary teachers who have previously been evaluated as competent (Derrington, 2019).

Teacher Evaluation System Implemented in Tennessee

With an excellent multi-million-dollar federal grant to overhaul teacher evaluation within the public school setting, the state of Tennessee was one of the proponents of evaluation reform. Tennessee piloted its evaluation model after ESSA reduced student growth and achievement requirements and no longer required states to implement a teacher evaluation model. Inspired by the Race to the Top (RTTT) grant, the Tennessee Department of Education approved a teacher evaluation framework (Tennessee Department of Education, 2021), accompanied by high-stakes consequences. One of the most notable changes was to tenure laws to allow for the more flexible firing of ineffective teachers and piloted recruitment and retention bonuses programs to attract and retain effective teachers in low-performing settings (Springer et al., 2015; Swain et al., 2019). The Tennessee State Board of Education approved four teacher evaluation models: the Tennessee Educator Acceleration Model (TEAM), Project Coach (COACH), Teacher Effectiveness Measure (TEM), and Teacher Instructional Growth for Effectiveness and Results (TIGER); however, TEAM is the default evaluation model across the state (Department of Education, 2021). Tennessee's performance-based teacher evaluation model requires 50% of the evaluation to be comprised of student achievement data that includes 35% based on student growth measure represented by the Tennessee Value-Added Assessment (TVAAS) and 15% based upon the student achievement selected by the teacher (Tennessee Department of Education, 2021). Under the new evaluation system, all school personnel were evaluated by school leaders and were responsible for ensuring the fidelity of the teacher evaluation implementation (Campbell & Derrington, 2019).

Overview of Teacher Evaluation Accelerator Model (TEAM)

The Teacher Evaluation Accelerator Model (TEAM) is one of the four models that were

utilized by the state to measure teacher effectiveness. TEAM aims to allow administrators and educators to collaborate to ensure quality education occurs in the classroom daily and to support teachers in honing instructional practices through frequent evaluation, constructive feedback, student data, and professional development (Tennessee Department of Education, 2021). The evaluation scores could be used for high-stakes decision-making such as hiring, tenure, promotion, and firing of teachers (Moran, 2017). Table 1 provides specifics of TEAM implemented by the state.

Table 1

Specifics on the Teacher Evaluation Accelerator Model

Evaluation Scores- 50% Qualitatively Based	*50% Quantitatively Based*
Classroom evaluations based on the following broad criteria: • Instruction • Planning • Environment • Professionalism	Teacher evaluation scores based on the following student data: • 35% value added data • For K–2 teachers, this is a schoolwide average of 3 years of scores.
Teachers scored on a scale of 1–5 based on the following: • Motivation of students • Presentation of instructional content • Lesson structure and pacing • Activities and materials • Questioning • Academic feedback • Grouping of students • Teacher content knowledge • Teacher knowledge of students • Problem solving	Remaining 15% of quantitative score based on one of the following: • Professional development score • Value added data • Other form of student assessment (could include DIBELS scores, running records, or other forms of assessment)
Teacher evaluation scores used for high stakes decisions such as: • Promotion • Tenure • Hiring • Firing	

Note: Specifics on the Teacher Evaluation Acceleration Model implemented. From “The Impact of a High Stakes Teacher Evaluation System: Educator Perspectives on Accountability” by R.M.R. Moran, 2017, *Educational Studies*, p. 7 (https://doi.org/10.1080/00131946.2017.1283319).

In 2011, Tennessee piloted TEAM, requiring school administrators to pass an inter-rater reliability assessment in which participants viewed videos of lessons being delivered by teachers and assessed teachers using the mandated rubrics (Tennessee Department of Education, 2021). The school principals must use classroom rubrics to rate teachers' instructional practice during classroom observations, using indicators 1, 2, 3, 4, and 5 for scoring (Tennessee Department of Education, 2021). Depending on the number of observations a teacher may have annually, the final observation score is the average of all observations conducted throughout the year (Tennessee Department of Education, 2021). At the end of the school year, teachers received the Level of Overall Effectiveness (LOE) rating ranging from Level 1 to Level 5, indicating Level 1 as significantly below expectations and Level 5 as significantly above expectations (Tennessee Department of Education, 2021). The qualitative, student growth, and student achievement data were combined to create a range score between 100 to 500 (Tennessee Department of Education, 2021). Table 2 provides the score range with the respective LOE scales.

Table 2

Tennessee Level of Overall Effectiveness Scale

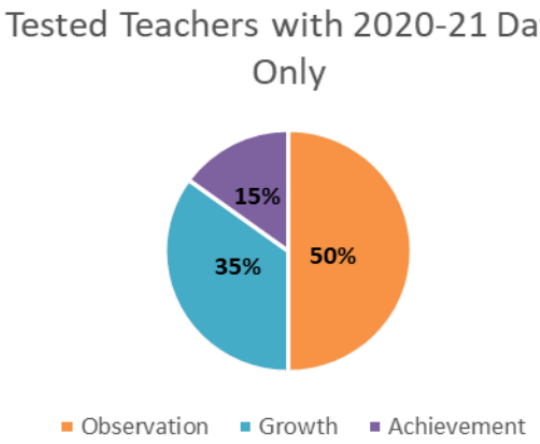
Score Range	Level of Overall Effectiveness (LOE) Scale
<200	1
200-274.99	2
275-349.99	3
350-424.99	4
425-500	5

Note: 1 = significantly below expectations; 2 = below expectations; 3 = meets expectations; 4 = above expectations; 5 = significantly above expectations. From “Tennessee Teacher Evaluation” by Tennessee Department of Education, 2021, p. 1 (https://team-tn.org/teacher-evaluation-2-2/).

Due to legislative changes made during the 2013 legislative sessions, the determination of the LOE depends on whether a classroom teacher has an individual growth score or a school- or system-wide growth score (Tennessee Department of Education, 2021). The LOE scores were utilized for merits and professional support since LOE scores became the basis of incentivizing teachers for performance and retention bonuses. In Springer et al.'s (2016) study, 20% of Level 5 teachers in Tennessee who received retention bonuses were more likely to remain in high-priority schools. Under the TEAM model, classroom observation composes 50% of the evaluation, student growth is 35 %, and student achievement is 15% (Tennessee Department of Education, 2021). Figure 1 provides a visual representation of the components of the Tennessee teacher evaluation system for teachers teaching tested subject areas.

Figure 2

Components of the Tennessee Teacher Evaluation System Accountability



Note: The 2020-2021 Level of Effectiveness (LOE) calculations for teachers with data only. The LOE combines measures of qualitative (observation), growth, and achievement. From "Tennessee Teacher Evaluation System Accountability" by Tennessee Department of Education, 2021, p. 1 (https://www.tn.gov/content/dam/tn/education/2020-21-leg-session/Accountability%20Teacher%20Evaluation%20Guidance%20-%20508%20(1).pdf).

Classroom observation is a qualitatively based measure comprised of four domains: instruction, planning, environment, and professionalism. Classroom teachers vary several observations, either announced or unannounced, based on the LOE scores in a previous year. All school evaluators utilized a universal TEAM rubric to score teachers during observations. The rubric entails clear expectations of high-quality instruction demonstrated in the classroom with a rating ranging from 1 to 5. The TEAM required conferences before and after the observation except for unannounced observations, which pre-conferences are unnecessary. The meetings aim to ensure evaluators and teachers collaborate on the outcomes and discuss the areas of refinement and growth. A teacher who receives an unsatisfactory score of less than a 3 rating could receive support through professional development and peer mentoring in the areas that need improvement.

Student growth refers to the Tennessee Value-Added Assessment System (TVAAS), measured by how much gain or progress a student or group makes over time (Tennessee Department of Education, 2021) based on mandated state assessments. All educators who teach core subjects such as mathematics and English and Language Arts are expected to receive individualized TVAAS scores. In contrast, teachers who do not teach tested subject areas can choose school-wide or a district-wide score for numeracy, literacy, or both (Tennessee Department of Education, 2021). Student achievement refers to proficiency or mastery of grade-level standards (Tennessee Department of Education, 2021). Early in the school year, educators are to select an achievement measure that is most closely aligned with the teaching subject assignment. Like growth measures, teachers teaching core subjects are most likely to receive proficiency or mastery student data. In contrast, non-tested subject teachers can select between numeracy and literacy.

Advantages and Disadvantages of Teacher Evaluation Accelerator Model (TEAM)

Enhancing teachers' instructional pedagogies and improving student learning are the primary goals of TEAM. The Tennessee Education Research Alliance (TERA) gathered information annually through the Tennessee educator survey (Tennessee Department of Education, 2021) to determine teachers' experiences with the state's evaluation process. Since TEAM launched in 2012, there has been a significant increase in the percentage of teachers, indicating that the evaluation has led to instructional improvements and student learning. In 2012, 38% of educators said that the evaluation has improved teaching, and 28% expressed that the evaluation has improved student learning. In 2019, the results of the educator survey doubled the percentage of the benefits of the evaluation (Tennessee Department of Education, 2021). Amidst the positive results of the advantages of TEAM, the results of the educator survey in 2018 indicated that 50% of the participants expressed that the evaluation was a significant burden due to both time and resource constraints and anxiety (Tennessee Department of Education, 2021). The teachers expressed how the implementation of TEAM impacted how teachers plan the lessons, and educators focused more on higher-order thinking skills instead of allowing students to master the pre-requisite skills (Tennessee Department of Education, 2021).

The Tennessee Value-Added Assessment System (TVASS)

The Tennessee Value-added Assessment System (TVAAS) is used to measure student academic growth and academic achievement based on the results of state standardized tests, the Tennessee Comprehensive Assessment Program (TCAP) in grades three through eight and end-of-course (EOC) exams in high school. The TVAAS utilized the previous and current student's achievements to determine academic growth, allowing teachers to reflect on instructional

practices. The TVAAS incorporates 35% of the teacher LOE score rating annually, and one of the focal points of TVAAS is to increase teacher's accountability for student learning. (Tennessee Department of Education, 2021).

Overview of the Tennessee Value-Added Assessment System (TVAAS)

William Sanders developed the Tennessee Value-Added Assessment System (TVAAS) in the early 1990s (Colson et al., 2018; Davis et al., 2016) as a component of the Tennessee evaluation system. TVAAS utilized statistical mixed-method model theory and methodology to enable multivariable and longitudinal analysis (Sanders & Horn, 1998) of student academic gains from the Tennessee Comprehensive Assessment Program (TCAP) in grades three through eight and end-of-course (EOC) exams in high school subjects in five areas: mathematics, science, social studies, reading, and language arts. TVAAS was designed to ascertain the effectiveness of the school districts and teachers in producing academic achievement for all students in Tennessee (Sanders & Horn, 1998). TVAAS is currently available commercially from the SAS statistical software company (Kupermintz, 2003) to provide reports annually of students' gains for each subject and grade.

Advantages of The Tennessee Value-Added Assessment System (TVAAS)

The proponents of TVAAS highlighted some of the advantages of incorporating TVAAS into teacher evaluation. Sanders and Horn (1998) mentioned that the primary objective of TVAAS was to deliver information of summative assessment about the effectiveness of the school districts, the schools, teachers, and students over time. The TVAAS reports allowed educational leaders to determine the grade and subject failures and successes and allocate immediate efforts and resources accordingly (Colson et al., 2018; Sanders & Horn, 1998). Proponents of TVAAS, such as Stone (1999), asserted that TVAAS could answer inquiries about

teacher effectiveness and robust accountability. Tracking students' academic achievement could provide accurate and trustworthy measures of individual students' needs (Sanders, 2000). Additionally, TVAAS allowed educational leaders to offer classroom educators some professional development needs, showcase students' growth and achievement, and ensure all students are challenged to grow and learn (Tennessee Department of Education, 2021).

Disadvantages of The Tennessee Value-Added Assessment System (TVAAS)

While there were advantages of TVAAS, criticisms have surfaced, indicating concerns over using students' test scores for teacher evaluation. Some researchers have argued that the TVAAS model failed to control for socioeconomic status and demographic factors (Ballou et al., 2004; Kupermintz, 2003). Further, Ballou et al. (2004) stressed that students were not randomly assigned to teachers and schools. If disadvantaged students were systematically assigned to less effective schools and teachers, including SES as control could cover genuine differences in school and teacher quality. While proponents of TVAAS called the model a robust measure for accountability, Warren and Ward's (2018) analysis of videos during school board meetings indicated that teachers felt the pressure of responsibility and experienced a disconnect between the promise of policy and what transpired and delivered. Additionally, the use of TVAAS to determine teacher tenure pay, retention, bonuses, and continuing employment is under scrutiny (Konstantopoulos, 2014). Further, Hunter and Rodriguez's (2021) study findings backed up on how the demands of teacher evaluation affected the work of school administrators, and results suggested that evaluators allocated a set amount of time for observing teachers.

Summary

Over the past ten years, policymakers have spurred discussions to provide information on

what makes a quality teacher. The transformation of the traditional evaluation system into the new teacher evaluation system has been a heated topic worldwide. In the United States, policymakers urge educational stakeholders to utilize school districts' evaluation systems to measure teacher quality, retain highly effective, and let go of ineffective teachers. Several researchers undermined the underrated theoretical underpinnings of teacher evaluation systems and how they might intersect theoretically with self-determination theory, in which classroom educators have the autonomy to deliver valuable instruction to students. The high-stakes evaluation system incorporated value-added measures to assess teachers' performance, and teachers have been subject to merits and penalties based on the evaluation results (Xu et al., 2016). While managing emotions and social networks is essential in retaining and sustaining teachers, the literature review has shown how the teacher evaluation system negatively impacts teachers' morale and well-being (Anderson et al., 2019; Cuevas et al., 2018). Some of the negative factors of teacher evaluation supported by the literature are increased stress, unhealthy competition (Paulfer, 2018), and administrators' biases and inconsistencies.

The validity and reliability of the evaluation process are crucial in measuring teachers' effectiveness. Research has provided faculty teaching perceptions of the multi-source evaluation methods (Lyde et al., 2016), questioning the process's validity and reliability. State assessment data tied to the teacher evaluation could potentially improve school (Brewer et al., 2015), but are not necessarily valuable for teaching and learning. While research has investigated the teacher evaluation system, a gap in the literature exists regarding the teachers' voices of the overall evaluation system, especially in Tennessee. Additionally, the gap is not only about the body of literature on teachers' views of the system but a time gap since the teachers' perception concerning the teacher evaluation in Tennessee existed in 2014. Additionally, there is a gap in

research instrumentation; the Teacher Evaluation Profile (TEP) incorporated in the study was primarily used in the theses and dissertations.

Since the role of gender is a critical factor in maintaining gender inequity and knowing that women hold more than sixty percent of the teaching jobs in primary and secondary education worldwide (Jerrim & Sims, 2019), teacher perception regarding the evaluation system is crucial in improving or redesigning the evaluation process. Allowing educators' voices to be heard could impact teachers' professional growth because those teachers who may have low well-being will, or teachers rated low, such as male and novice teachers (Campbell & Ronfieldth, 2018; Drake et al., 2019; Jiang & Sporte, 2016) most likely leave the teaching profession. The success of policy implementation is contingent on buy-in from stakeholders. Such willingness and acceptance should be evidenced by transparent support of the teacher evaluation system, feedback, process, and the teachers' drive to change and improve after the observation (Reddy et al., 2018). Policymakers and other stakeholders need to acknowledge and listen to teachers' voices because they are directly affected by the evaluation system's implementation and directly impact students' academic achievement and growth.

CHAPTER THREE: METHODS

Overview

This quantitative, predictive correlational study aimed to determine if a predictive relationship exists between teacher perceptions of Tennessee's teacher evaluation and the linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high). This chapter introduced the study's design, including full definitions of the variables. The research question and null hypothesis follow. The participants, setting, instrumentation, procedures, and data analysis plans were presented.

Design

This quantitative, predictive, correlational study aimed to ascertain how accurately teacher perception of Tennessee's teacher evaluation can be predicted from a linear combination of teacher gender, years of experience, and school-level taught (elementary, middle, and high) for Tennessee educators. It used a predictive correlational design to predict relationships, not causality, among the variables (Dean et al., 2017; Gall et al., 2007; Ott & Longnecker, 2017). However, it was worth mentioning that the proposed design has a prominent limitation: the lack of being used to conclude causal relationships among the measured variables (Warner, 2013). The design was suitable because the variables were identified and defined, data collection was possible, and it featured only one group of participants with no interventions needed before, during, or after data collection; data were collected at one point with a focus on a linear relationship between variables (Leddy & Omrod, 2018; Ott & Longnecker, 2017; Warner, 2013).

The predictive correlational design implemented in the study endeavors to predict teacher perceptions of the current Tennessee teacher evaluation system and the linear combination of teacher age, years of experience, and school-level taught. Previous peer-reviewed literature had

focused on teacher evaluation as a process of gathering data to measure a teacher's performance (Holloway, 2019) and determine the quality of teaching. A teacher's quality is an educator's characteristic that influences student achievement and has been studied using a variety of designs (Kolman, 2017). Student achievement also has a long, well-established history of designs focusing on the general linear model, targeting proficiency or mastery outcomes of student learning and academic performance across tested subject areas, such as English and Language Arts, mathematics, and science (Alexander et al., 2017). The Tennessee Educator Accelerator Model (TEAM), for instance, is one of the tools used to measure teacher performance annually (Tennessee Department of Education, 2021), and correlational research has incorporated the Tennessee Value-Added Assessment (TVAAS) that measures students' performance through standardized test scores (Tennessee Department of Education, 2021).

The predictor variables were gender, years of experience, and school level taught—elementary, middle, high. The criterion variable was the teacher perceptions of Tennessee's teacher evaluation system. For this study, gender was defined as a variable with a possible response of females and males (Lindqvist et al., 2020). Teacher years of experience was defined as a variable that pertains to the number of years the participant teaches (less than three years are considered beginning teachers), four to five years are considered transitioning teachers, and more than five years are considered experienced teachers) (Campbell, 2020). The school level was a variable that referred to the respondents' instructional level, either in elementary (first to fifth grades), middle (sixth to eighth grades), or high (ninth to twelfth grades) (Dogan et al., 2021). The teacher evaluation system is trustworthy when stakeholders, such as classroom educators, become engaged and are professional resources rather than unvalued individuals (Paufler & Sloat, 2020). Since low-performing teachers are most likely to be the subject of public scrutiny

(Jewell, 2017), teacher evaluation outcomes allow teachers to promote reflective discourses for educators' pedagogical improvement (Williams & Herbert, 2020).

Research Question

The research question for this study was:

RQ: How accurately can teacher perception of Tennessee's teacher evaluation system be predicted from a linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high) for Tennessee educators?

Hypothesis

The null hypothesis for this study was:

H_0: There is no significant predictive relationship between the criterion variable (teachers' perceptions of the Tennessee evaluation system) and the linear combination of predictor variables (teacher gender, years of experience, and school-level taught (elementary, middle, high) for Tennessee educators.

Participants and Setting

The teacher population and sample demographics were discussed in this section. Specifically, the following three sections presented the population, the participants, and the study's setting from which the population and sample emerge. The reader would find a methodological warrant for the sample size and a justification for using a convenience sample.

Population

Elementary, middle, and high school teachers were invited to participate from one of the largest school districts in southwestern Tennessee. The school district was invited to participate because of the district's accessibility and size, comprising more than 6000 teachers serving more than 100,000 students (Tennessee Department of Education, 2021). For this current study, only

teachers in elementary, middle, and high who have a valid license were invited to participate. Additionally, the participants included general and special education teachers of all subjects, while administrators, professional coaches, and instructional leaders were excluded from the present study.

Participants

The study participants were a convenience sample from a population of elementary, middle, and high school teachers in one of the largest school districts in southwestern Tennessee during the school year 2023-2024. Gall et al. (2007) indicated the reasons researchers employ convenience sampling as a method to select a sample: the location of the sample is close, where the researcher works, the researcher is familiar with the site, or the administrators who approve of collecting data are colleagues of the researcher. Convenience sampling was appropriate because the projected participants are easy to find, and the location of the study was accessible. For this study, the sample size required is the larger value of either $n > 50 + 8k$ or $n > 104 + k$, where k is the number of predictor variables. For example, If $k = 3$, the first calculation is $50 + 8k = 74$ (minimum $n = 75$). The second calculation is $n > 104 + k = 107$ (minimum $n = 108$). Therefore, the minimum sample size should be 108 (Warner, 2013).

Participants included 67 (62%) females, and 41 males (38%) in the sample. There were 28 (25.9%) of the respondents have less than three years of teaching experience, 6 (5.6%) have four to five years experience, and 74 (68.5%) with 6 or more years of experience. Among the participants, there were 13 (12%) of the teachers in the study taught at the elementary school (first to fifth grade), 13 (12%) middle school (sixth through eighth grade), and 82 (76%) high school (ninth through twelfth grade).

Setting

Data were collected through an online survey during the spring semester of the 2023 – 2024 school year. Educators accessed the survey via a hyperlink provided through the invitation e-mail. Although the survey should take about 10 – 15 minutes, participants could complete it at their own pace. Due to the online nature of the survey, participants were able to complete the survey when and where they choose.

Instrumentation

After obtaining permission (See Appendix A) from the developer of the instrument, the study incorporated the use of the modified Teacher Evaluation Profile (TEP) questionnaires (See Appendix B) designed by Stiggins and Duke and was initially developed by Northwest Regional Educational Laboratory (NREL) but was introduced in 1988 (Stiggins & Duke, 1988). This instrument examined teachers' perceptions of teacher evaluation experiences (Stiggins & Duke, 1988). It was developed to validate the list of critical attributes of teacher evaluation by determining if these features were related to recognizing growth outcomes of evaluation (Stiggins & Duke, 1988). Since the responses to the TEP questionnaires delivered an exploration of the relationships among components of the evaluation process and the relationships between key attributes and the various perceived outcomes of the evaluation, TEP was a reliable instrument to measure teachers' perceptions of the evaluation system (Stiggins & Duke, 1988). No empirical research incorporated the TEP instrument despite its consistent reliability. Up to date, the TEP is primarily used in theses and dissertations. The instrument was used in numerous dissertations (e.g., Cumpston, 2018; Thomas, 2009).

Stiggins and Nickel (1989) performed a factor analysis on the TEP subscales using varimax rotation. The analysis indicated that TEP was a helpful instrument that could help

school districts design a teacher evaluation process. The TEP delivered the highest reliability and internal consistency reliability of .93 (Stiggins & Nickel, 1989). The TEP comprises 48 questions and wass divided into three sections; the first section asked about the the participants' demographics, the second section asked about teachers' perceptions of the overall quality of the teacher evaluation system, and the third section involved the critical attributes of the teacher evaluation that foster the professional growth of teachers. The third section was grouped into five key attributes: the self as a teacher, the evaluator, the evaluation process, feedback, and context. The first section included three (3) questions about participants' demographics, the second section comprised of three (3) questions, and there were 42 questions in the third section, further disaggregated by sections including teacher attribute (8), evaluator attribute (11), evaluation process (11), feedback attribute (8), and context (4).

The participants responded by rating each question from 1 to 5. Each item and response is worded so that the responses can be congruent with the phrase. For example, a phrase in which participants rated the overall quality of the teacher evaluation system. The responses were set from "Poor Quality (1)" to "High Quality (5)". A rating of 5 indicates that participants perceived the high quality of the teacher evaluation system, while a rating of 1 indicates that the teacher evaluation has a poor overall quality. Each section incorporates a five-point Likert scale survey. According to Morgan and Hammon (2001), a Likert scale survey instrument was established to measure perceptions of particular groups or concepts. The teachers' perception score ranges from 1.00 – 1.80 (Very Low), 1.81 – 2.60 (Low), 2.61 – 3.40 (Moderately High), 3.41 – 4.20 (High), and 4.21 – 5.00 (Very High) (Morgan & Hammond, 2001). The range of 1.00 to 1.80 indicates a meager rating of the overall quality of the teacher evaluation system, and the range of 4.21 to 5.00 means a very high rating of the overall quality of the teacher evaluation system.

The TEP was administered to teachers online using Microsoft Forms. The presentations of questions will be arranged through sections. The TEP questionnaires could be completed in approximately 10 – 15 minutes. Data will be collected via Microsoft Forms, then converted into an Excel spreadsheet and uploaded to SPSS 29.0 for analysis. Permission to use the modified TEP was granted (See Appendix A for permission to use the instrument).

Procedures

Before conducting research, the study proposal was submitted to the Liberty University Institutional Review Board (IRB) for approval since human beings were surveyed. Upon securing IRB's approval (Appendix C), an application was completed to the target school district's research and evaluation department requesting permission to conduct a study on elementary, middle, and high school teachers. The attached informed consent form (See Appendix D) to that application, indicating that the participants signified their consent to participate in the study upon submission of the completed survey. After authorization was granted, the researcher sent a recruitment email initially to school administrators, informing them that the teachers would be surveyed and explained the study's intent, including (a) an introductory email explaining the study, (b) the two-week window for the study, (c) request for participation, and (d) a hyperlink and QR Code to access the survey. Responses was monitored within days (5) days. If insufficient surveys were collected, the researcher sent a reminder email to school leaders one week from the initial email as a follow-up reminder until a sufficient sample size is reached.

The study's survey was presented as continuous pages. Microsoft Forms platform allowed respondents to navigate a hyperlink for the survey page. By clicking the attached link, the teachers anonymously completed 48 questions. By clicking "Submit," a "Thank you for your

Participation" appeared once the survey was submitted. The survey could be completed within 10 to 15 uninterrupted minutes.

The participants' responses was kept confidential and secured. To ensure that the participants' privacy and responses were securely stored, the survey was self-reported and administered via Microsoft Forms, a web-based survey software that allows participants to respond anonymously. According to Gal et al. (2007), anonymity is crucial to protect participants' identities and encourage direct participation in the survey. Within the survey time frame, access to Microsoft Forms will be limited by the researcher. Following collection, the data were downloaded and stored on the researcher's password-protected computer for no more than five years. The researcher inspected for missing entries and incomplete surveys and assigned participants a participant number for organizational purposes. Once data were downloaded and secured, all web-based data was deleted. The collected survey responses were scored and analyzed using the Statistical Package for the Social Sciences (SPSS) Software. Warner (2013) mentioned that SPSS could make reasonable default decisions on grouping scores, interval widths, and the number of intervals to use. Additionally, Ong and Puteh (2017) stressed that SPSS allows the researcher to check for the test assumptions, and descriptive statistics were essential in organizing and summarizing data mathematically (Gall et al., 2007).

Data Analysis

For this study, a multiple regression analysis was utilized to support the investigation of the relationship between two or more variables (Gall et al., 2007). Multiple regression analysis allowed the researcher to determine the significance of the change in the criterion variable for an incremental increase of a predictor variable (Gall et al., 2007). Multiple regression was used to measure the relationship between the criterion variables, teacher perceptions, and a combination

of the predictor variables, teacher gender, years of experience, and school-level taught. Multiple regression was chosen because of its general applicability to educational research and its ability to show a depth of information about the relationship between two or more predictor variables on criterion variables (Gall et al., 2007).

Multiple Regression Data Screening and Assumptions

Data screening was conducted through a visual screening of the data set to check for missing data points and inaccuracies suggested by Barthlow et al. (n.d.), Ott and Longnecker (2017), and Warner (2020). Also, casewise diagnostics were administered to check for extreme outliers (Ott & Longnecker, 2017). Since gender, years of experience, and school-level are categorical variables, dummy variables were assigned to levels to make categorical variables meaningful. To this extent, 1 was assigned for females and 2 for males. Additionally, 1 was assigned to less than 3 years (beginning), 2 was assigned to 4 – 5 years (transitioning), and 3 was assigned to 6 or more years (experienced). Similarly, 1 was assigned to elementary, 3 was assigned to middle, and 2 was assigned to high school.

A statistic or graph was not used to test the first two assumptions; the assumptions were methodologically determined by the criterion variable measured at the continuous level, and by the predictor variables measured at categorical level, as categorical predictors are allowed in multiple regression analysis (Field, 2018; Hardy, 1993; Lewis-Beck, 1980). Assumptions three through eight are calculated statistically. Further, independent observations was assessed using the Durbin-Watson statistic, which ranges from 0 to 4, with a value of approximately 2, indicating no correlation between residuals. Linearity between the dependent and independent variables collectively and between the dependent variable and each independent variable was evaluated by viewing a scatterplot of the residuals against the predicted values. The researcher

then determined if linear relationships exist using partial regression plots between each independent and dependent variable. The subsequent overall assumptions and homoscedasticity of residuals—equal error variances—were tested by visual inspection of unstandardized or standardized residual scatterplots against the predicted values. If homoscedasticity existed, the residuals were identical across the standardized fitted values. The plot points exhibited no pattern and were approximately constantly spread across the fitted values. The absence of multicollinearity was the following assumption.

Correlation coefficients and Tolerance and Variance Inflation Factor (VIF) values were used to test for multicollinearity. None of the independent variables should have correlations greater than 0.7; VIF values should be lower than 10. The penultimate assumption related to data screening – no significant outliers. The researcher used Casewise diagnostics to highlight data points for which the standardized residual is more remarkable than three standard deviations. If the researcher detected outliers, one of four acceptable actions was appropriate: transforming the outlier, winzorizing it, applying a robust estimation method, or trimming the outlier or outliers (Field, 2018). Finally, multiple regression assumed a normal distribution of residuals – errors. A normal distribution assumed if the points are aligned with the line of fit (Barthlow et al., 2004). To test for normality, P-P Plots or Q-Q plots will be used.

Multiple Regression Data Analysis

In a multiple regression analysis, three pivotal tables emerged. Initially, the model summary delineated the model's explanatory power to fit the data, signifying this through the coefficient of determination (R^2). This same coefficient was also employed as the study's effect size indicator. Subsequently, the ANOVA table provided insight into whether the model’s explanatory power was statistically significant. In the third output, the coefficients table

elucidated which specific independent variables, if present, have a statistically significant predictive capacity for the dependent outcome (Field, 2018; Warner, 2013). The chosen alpha (α) of .05 represented the likelihood of a Type I error in the hypothesis test (Dean et al., 2017). The coefficient of determination (R^2) effect size used Cohen's (1988) conventions of .01 = small; .06 = medium; and .14 = large.

CHAPTER FOUR: FINDINGS

Overview

The purpose of this quantitative, predictive correlational study was to determine if teacher gender, years of experience, and school level taught (elementary, middle, high) could predict teachers' perceptions of Tennessee's teacher evaluation system. The predictor variables were gender, years of experience, and school-level taught (elementary, middle, high). The criterion variable was teachers' perceptions of Tennessee's teacher evaluation system. A multiple linear regression was used to test the hypothesis. The Results section includes the research question, null hypothesis, data screening, descriptive statistics, assumption testing, and results.

Research Question

RQ: How accurately can teacher perception of Tennessee's teacher evaluation system be predicted from a linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high) for Tennessee educators?

Null Hypothesis

H_0: There is no significant predictive relationship between the criterion variable (teachers' perceptions of the Tennessee) and the linear combination of predictor variables (gender, years of experience, school-level taught (elementary, middle, high)) for Tennessee educators.

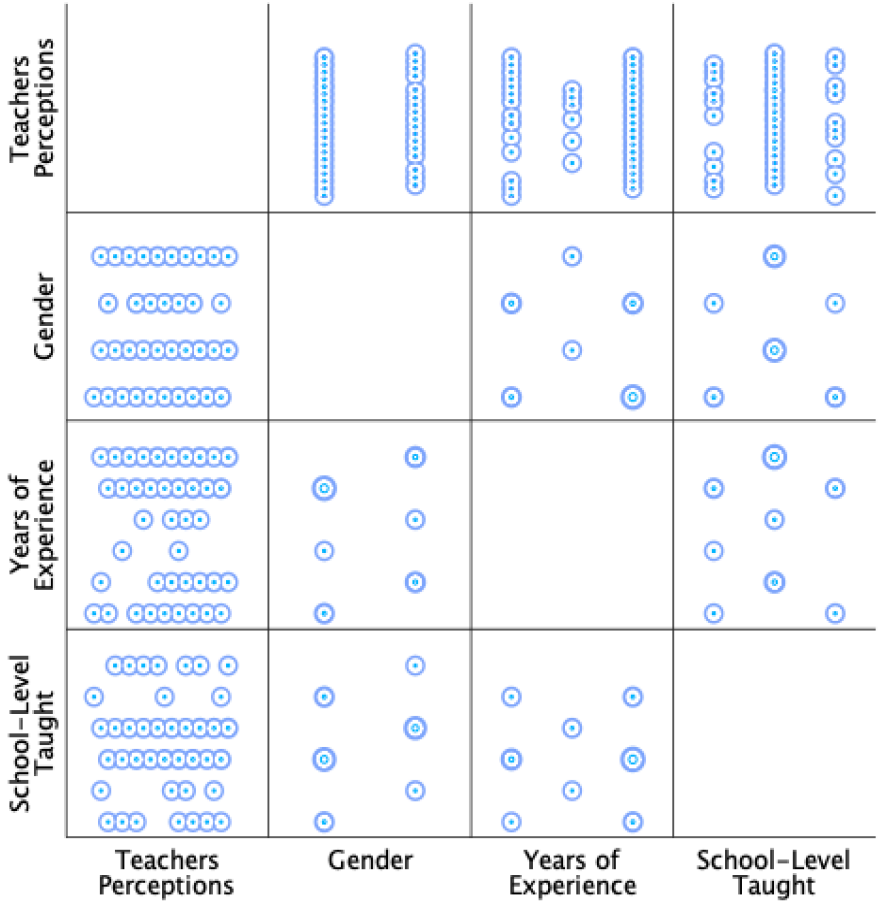
Data Screening

The researcher sorted the data and scanned for inconsistencies in each variable. The study consisted of 111 Tennessee educators. Three data errors or inconsistencies were identified and were removed from the data set, resulting to 108 respondents. The data were entered into SPSS. When complete, all entries were checked for accuracy. A matrix scatter plot was used to detect

bivariate outliers between the predictor variables and the criterion variable. No bivariate outliers were identified. See Figure 3 for the matrix scatter plot.

Figure 3

Matrix Scatter Plot



Descriptive Statistics

Descriptive statistics were obtained on each of the variables. The sample consisted of 108 participants. Teachers' perceptions were measured using the modified Teacher Evaluation Profile, with a 5-point Likert scale. Possible scores on the questionnaires ranged from 1.00 to 5.00. A score of 4.21 to 5.00 indicates a very high perception, while a range of 1.00 to 1.80 indicates a meager rating of how teachers perceived (Morgan & Hammond, 2001). Table 3 provides the descriptive statistics for each variable.

Table 3

Descriptive Statistics

	n	Min.	Max.	*M*	*SD*
Gender	108	1	2	1.38	.49
Years of Experience	108	1	3	2.43	.88
School-Level Taught	108	1	3	2.00	.49
Teachers' Perceptions	108	2.79	5.00	4.14	.60
Valid *n* (listwise)	108				

A frequency analysis was conducted for the categorical predictor variable. Table 4 provides the results of the frequency analysis for gender, years of experience, and school-level taught.

Table 4

Frequency Table for Predictor Variables

	Frequency	Percent
Gender		
Female	67	62.0
Male	41	38.0
Years of Experience		
0 – 3	28	25.9
4 – 5	6	5.6
6 or more years	74	68.5

	Frequency	Percent
School-Level Taught		
Elementary (1 – 5)	13	12.0
High (9 – 12)	82	75.9
Middle (6 – 8)	13	12.0

Note. $n = 108$.

Assumptions Testing

The first assumption for running a multiple linear regression is that the criterion variable is continuous, which was true for this study. The second assumption for running a multiple linear regression is that there are two or more predictor variables that are either continuous or nominal. All predictor variables for this study were measured on a nominal scale; therefore, this assumption was also met.

Independence of Observations

The assumption of independence of observations was tested using the Durbin-Watson statistic. The Durbin-Watson statistic can range from 0 to 4, with a value near 2 indicating that the assumption of independence observation is tenable and that the residuals are uncorrelated. If the Durbin-Watson value is greater than 2, then it indicates negative correlation. For this study, a value of 2.34 was reported, which was greater than 2; however, using Field's (2018) heuristic that considers values less than 1 or greater than 3 problematic, the assumption of independence was met.

Assumption of Linearity

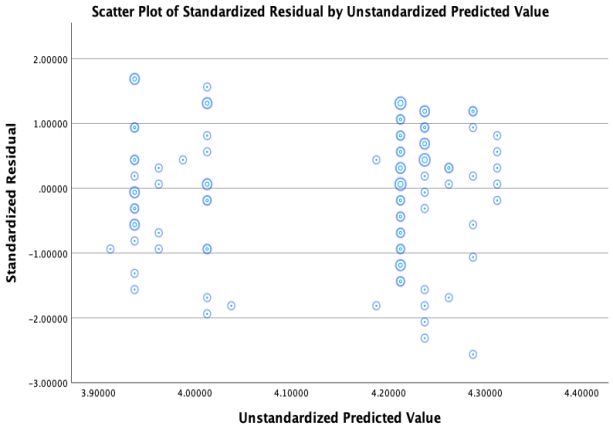
Multiple linear regression requires that the assumption of linearity be met. This was done in two parts. First, linearity was assessed between the criterion variable and the predictor variables collectively by plotting the standardized residuals against the unstandardized predicted values. Next, the linear relationship between the criterion variable and each of the predictor variables was assessed using partial regression plots. Visual inspection of the scatter plot indicated a nearly normal distribution for all. Next, the partial regression plots were created between each predictor variable and the criterion variable. Visual inspection revealed a nearly linear relationship. The assumption of linearity was tenable. See Figure 3 for the matrix scatter plot.

Assumption of Homoscedasticity

The assumption of homoscedasticity was checked using the scatter plot created when assessing the assumption of linearity by plotting the studentized residuals against the unstandardized predicted values. The scatter plot showed that the residuals were evenly spread; thus, the assumption of homoscedasticity was tenable. See Figure 4 for the scatter plot.

Figure 4

Scatter Plot of Residuals



Assumption of Multicollinearity

To test this assumption, Pearson's correlation coefficients and variance inflation factor (VIF) tests were conducted. If the VIF is greater than 10, then the predictor variables are highly correlated. Examination of Pearson's correlation coefficients for all predictor variables revealed all correlations are below the threshold of $r = .7$, as shown in Table 5.

Table 5

Pearson Correlation (r)

Variable		1	2	3
1	Gender	---	.004	.424
2	Years of Experience	.004	---	.505
3	School-Level Taught	.424	.505	---

Note. $n = 108$.

Table 6 below shows the collinearity statistics. The assumption of the absence of multicollinearity between the predictor variables was met.

Table 6

Collinearity Statistics

		Collinearity Statistics	
Model		Tolerance	VIF
1	Gender	.92	1.09
2	Years of Experience	.92	1.10
3	School-Level Taught	.99	1.01

Note. Dependent Variable: Teachers' Perceptions.

Assumption of No Significant Outliers

Casewise diagnostics were used to examine the data for the assumption of no significant outliers. The casewise diagnostics for standardized residual should not be greater than ±3 standard deviations. There were three cases that were identified for which the standardized residual exceeded three standard deviations, so the three cases were removed from the data set.

Assumption of Normal Distribution of Residuals

A P-P plot was created to determine if the data were normally distributed. The normality assumption is met when points fall predominantly on the line of best fit. A visual examination of the P-P plot found that the points were aligned along the line of best fit, indicating that the residuals were normal, as seen in Figure 5. Thus, the assumption was met.

Figure 5

P-P Plot of Residuals

Results

Multiple linear regression was conducted to determine whether a predictive relationship existed between teachers' perceptions and the linear combination of teacher gender, years of experience, and school-level taught. The criterion variable was teachers' perceptions and the predictor variables were gender, years of experience, and school-level taught. The researcher failed to reject the null hypothesis at the 95% confidence level where $F(3, 104) = 1.86$, $p = .14$. There was no significant relationship between the predictor variables, taken together as a model, and the criterion variables. Table 7 provides the regression model results.

Table 7

Regression Model Results

Model		*SS*	*df*	*MS*	*F*	Sig.
1	Regression	1.96	3	.65	1.86	.14[b]
	Residual	36.42	104	.35		
	Total	38.38	107			

[a]Dependent Variable: Teachers' Perceptions.

[b]Predictors: (Constant), Gender, Years of Experience, School-Level Taught

The model's effect size was small, where $R = .23$. Furthermore, $R^2 = .05$, indicates that approximately 5% of the variance of the criterion variable can be explained by the linear combination of predictor variables. However, a more accurate measure of the shared variance would be from the adjusted $R^2 = .02$, indicating that about 2% of the variation in teachers' perceptions can be explained by gender, years of experience, and school-level taught after adjusting for the number of predictors. Table 8 provides a summary of the model.

Table 8

Model Summary

Model	*R*	R^2	Adjusted R^2	*SEM*
1	.23[a]	.05	.02	.59

[a]Predictors: (Constant), Gender, Years of Experience, School-Level Taught.

Because the researcher failed to reject the null hypothesis, analysis of the coefficients was not required. However, based on the coefficients, it was found that gender was the best predictor of teachers' perceptions, where $p = .02$. At the same time, years of experience and school-level taught were not significant predictors within the model, where $p = .58$ and $p = .84$ respectively. Table 9 below provides the coefficients.

Table 9

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	*SE*	*β*	*t*	Sig.
1	(Constant)	*4.67*	.38		12.32	<.001
	Gender	*-.29*	.12	-.24	-2.36	.02
	Years of Experience	*-.04*	.07	-.06	-.56	.58
	School-Level Taught	*-.02*	.12	-.02	-.20	.84

Note. Dependent Variables: Teachers' Perceptions.

CHAPTER FIVE: CONCLUSIONS

Overview

Chapter Five discusses the results of the study in light of whether they support or contradict prior research and literature on the relationship between teachers' perceptions of the Tennessee teacher evaluation system and the linear combination of teacher gender, years of experience, and school-level taught. This chapter also discusses the practical and theoretical implications of the findings as well as the limitations of the study. The chapter concludes with recommendations for further research.

Discussion

The purpose of this quantitative, predictive correlational study was to determine how accurately can teachers' perceptions of Tennessee's teacher evaluation system be predicted from a linear combination of teacher gender, years of experience, and school-level taught. The study's predictor variables were gender, years of experience, and school-level taught. The criterion variable was teachers' perceptions. The sample comprised 108 certified teachers in one of the school districts in Tennessee who completed an online survey. Data were analyzed using IBM SPSS Version 29. A multiple linear regression was used to answer the hypothesis:

H_0: There is no significant predictive relationship between the criterion variable (teachers' perceptions of the Tennessee evaluation system) and the linear combination of predictor variables (teacher gender, years of experience, and school-level taught (elementary, middle, high) for Tennessee educators.

The null hypothesis was that there was no significant predictive relationship between the criterion variable (teachers' perceptions) and the linear combination of predictor variables

(gender, years of experience, school-level taught) for Tennessee educators. The researcher failed to reject the null hypothesis at the 95% confidence level, $F(3, 104) = 1.86$, $p = .14$, $R^2 = .05$. The results for this study showed that gender ($p = .02$) was a significant predictor of teachers' perceptions, however, years of experience ($p = .58$) and school-level ($p = .84$) taught did not have a predictive relationship with teachers' perceptions. The model indicated a small effect size (Warner, 2021), where the linear combination of the three predictor variables explained 5% of the variance in the total score of teachers' perceptions. However, the adjusted R^2 value of .02 is a more accurate measure of the shared variance indicating that 2% of the variability in teachers' perceptions can be explained by gender, years of experience, and school-level taught after adjusting for the number of predictors.

One of the study's findings is the frequency of males and females as participants. The results indicated that 62% of females and 38% of males participated in the study. Such results are consistent with the prior research reported by Jerrim and Sims (2019)'s study, in which 66% and 57% of women hold teaching jobs in elementary and secondary education worldwide. The current research suggests that gender is a significant factor in the educational system. Additionally, the influence of gender on teachers' perspectives of the state's teacher evaluation system strongly implies that both male and female educators consider the teacher evaluation system crucial for teacher accountability and growth, despite potential differences in their views. It could also mean a potential change in the labor force in gender equity, attracting more male educators in the classroom setting.

While the study's results show no significant relationship between teachers' perceptions and teacher years of experience, it does not necessarily suggest that novice teachers perceive the teacher evaluation system as the same as transitioning and veteran teachers. The findings of the

current study do not necessarily support the previous research that the 75.9% lowest retention rate among novice teachers compared to 85.2% for experienced teachers (Robertson-Kraft & Zhang, 2018), and higher turnover rates occurred in the first three years of teaching because novice teachers are less effective than experienced teachers as asserted by several researchers (Conley et al., 2019; Drake et al., 2018; Papay et al., 2017; Redding & Henry, 2019). This notation may be attributed to the sample size of novice teachers who participated in the present study (n = 28). Further, the TEP questionnaire does not include questions allowing teachers to perceive if the teacher evaluation system is the reason for retention or leaving the profession. Further, the study's results cannot support Donahue and Vogel's (2018) study that veteran or experienced teachers perceived that the evaluation system provided a framework for becoming proficient in the instruction.

Moreover, the current study shows no significant relationship between teachers' perceptions of the teacher evaluation system and the school level taught (elementary, middle, high). To date, no empirical study has been conducted on how teachers perceive the state's evaluation system across the school levels. However, previous research indicated that middle and high school teachers appraised the teacher evaluation system as neither positive nor negative (Reddy et al., 2017). In general, previous research indicated that some educators criticized the teacher evaluation system as subjective, in which female, experienced, and elementary school teachers rated higher than male, novice, and high school teachers by the evaluators (Williams & Herbert, 2020; Wind et al., 2019). On the contrary, as it relates to job satisfaction, male teachers agreed that teaching at the elementary school level is a fulfilling career and viewed that teachers could be role models and make a difference for students regardless of gender (Gaza, 2022; Meader & Larwin, 2022).

The present study supported the theory used to frame it. Self-determination theory (SDT) asserts the fundamental psychological needs in motivation, autonomy, relatedness, wellness, and competence (Ryan & Deci, 2017). In this study, self-determination positively explained educators' perceptions of the teacher evaluation system, indicating that regardless of gender, years of experience, and school-level taught, teachers embraced self-determination to grow professionally and are open to change that could potentially become competent and contribute a significant impact to students' academic achievement and growth (Franklin-Hayslip, 2021). Such assertion supports prior research that competence plays a pivotal role in strengthening a person's capabilities in hopes of future accomplishments (Ford, 2018). Further, previous research suggested that incorporating self-determination theory provided evidence of job satisfaction among educators, regardless of gender, novice, transitioning, or experienced teachers, and whether educators teach in elementary, middle, or high (Ford et al., 2018). Contrary to the prior studies (Ford, 2018; Ford et al., 2018; Ryan & Deci, 2017) that self-determination theory provided affirmation that SDT increased job satisfaction, teacher evaluation based on student performance can negatively affect teacher well-being, which has several potential implications for applied practice (Cuevas et al., 2018).

Implications

Researchers have studied the teacher evaluation system in Tennessee (Kupermintz, 2003; Konstrantopoulos, 2014; Tennessee Department of Education, 2021), but there is no study found in the literature that has investigated the predictive relationship between teachers' perception of the Tennessee evaluation system and the linear combination of gender, years of experience, and school-level taught for Tennessee educators. The findings revealed that even though there was no significant predictive relationship between teachers' perceptions and teacher gender, years of

experience, and school level taught, the predictor gender was found to be a significant factor in how male and female teachers perceived the state's evaluation system. Thus, the current study contributes to the existing body of literature on how women make up a higher percentage of the classroom population compared to men. The study provides additional context on how gender influences the way teachers perceive the Tennessee teacher evaluation system.

Due to a lack of empirical research about the teachers' perceptions of the Tennessee teacher evaluation system, this study has provided empirical data and made a significant contribution. The study's results also added to the lens of theoretical literature as it holds the tenet of self-determination theory. This theory supports a positive influence on teachers' ability to perceive the evaluation system with openness, autonomy, relatedness, and competence.

A significant practical implication of the study is for the state of Tennessee to provide a teacher evaluation system that is fair for all teachers and not on the repertoire of one-size-fits-all. An evaluation system will be incorporated to gauge teachers' efficacy without putting teachers' competence at stake in the student's academic test scores. Moreover, stakeholders and Tennessee lawmakers will consider teachers' voices should the current teacher evaluation system be revamped, and teachers' beliefs about the teacher evaluation system are a road to potential teacher buy-in to an evaluation system that could aim to promote development and growth. Researchers suggested that the teachers' perceptions regarding evaluation should be included, such that disregarding these perceptions could be a long and arduous process on the road from policy creation to enactment of change in the classroom, and if the state seeks to develop or redesign teacher evaluation must purposefully involve teachers in the development and validation of the evaluation process (Moran, 2017; Warren & Ward, 2018; Warsame & Valles, 2018).

Limitations

This study had several limitations that could impact its internal and external validity. First, while the minimum sample size requirement for the study was met, the findings may not represent Tennessee's broader population of teachers and limit the generalizability of the study's findings. Second, since the survey was administered at one point in time, it is outside the scope of the study to determine if teachers' perceptions have changed over time. A third limitation was using the Teacher Evaluation Profile (TEP) instrument. It could result in social desirability bias or response where respondents tend to present themselves favorably rather than reveal their fundamental beliefs (Gall et al., 2007). Teachers who participated in the study may have reported a more desirable view of themselves or provided responses they believed were expected, resulting in inaccurate data. A fourth threat to validity is the ratio of elementary, middle, and high school participants. The study included 26 combined elementary and middle school teachers and 85 high school educators. While school-level teaching was not a significant factor, a more balanced distribution of participants would have been desirable. Lastly, using regression design analysis was another limitation of the study. Correlational studies may indicate the strength and direction of relationships between variables and do not provide evidence of cause-and-effect relationships (Gall et al., 2007; Warner, 2021).

Recommendations for Future Research

Based on this study's findings and limitations, the following are recommendations for future research to expand the body of knowledge in teaches' perceptions of the Tennessee teacher evaluation system.

1. It would be helpful to repeat this study with a bigger sample size to determine if the size of the sample had an effect on the results.

2. It would be beneficial to replicate this study separately in elementary, middle, or high school teachers within the district to make connections and comparisons easily.
3. Conduct another study of all certified teachers across the state of Tennessee and compare the results between districts and the state as a whole to increase the generalizability of the findings to other contexts.
4. Include other variables, such as age, in a regression model to see if age is a significant predictor of how educators perceive the teacher evaluation system.
5. Replicate this study with another instrument, such as the Teacher Job Satisfaction Questionnaire, to determine if the teacher evaluation system has something to do with job satisfaction. This could assist policymakers in understanding why there is a shortage of teachers within the state.
6. Since the current study was correlational, future research could utilize a different quantitative research design. For example, a longitudinal study to determine if the state's teacher evaluation system has improved over time.
7. Another study using a qualitative design incorporating teacher interviews will be conducted to uncover teachers' experiences and classroom decisions regarding the Tennessee teacher evaluation system.

REFERENCES

Alexander, N.A., Jang, S.T., & Kankane, S. (2017). The performance cycle: The association between student achievement and state policies tying together teacher performance, student achievement, and accountability. *American Journal of Education*, *123*(3), 413 – 446. https://doi.org/10.1080/691229

Alford, A. (1965). The Elementary and Secondary Education Act of 1965: What to anticipate? *The Phi Delta Kappan, 46*(10), 483 – 488. https://www.jstor.org/stable/20343436

Amrein-Beardsley, A. (2020). Policies and practices of promises in teacher evaluation: The introduction to the special issue. *Educational Policy Analysis Archives*, *28*(55).

Anderson, D., Cameron-Standerford, A., Bergh., & Bohjanen, S., (2019). Teacher evaluation and its impact on wellbeing: Perceptions of Michigan teachers and administrators. *Education*, *139*(3), 139 – 150.

Ballou, D., & Springer, M.G. (2015). Using student test scores to measure teacher performance: Some problem in the design and implementation of evaluation systems. *Educational Researcher*, *44*(2), 77 – 86.

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

Barthlow, D. L., Graham, J. R., Ben-Porath, Y. S., & McNulty, J. L. (2004). Construct validity of the MMPI-2 College Maladjustment (Mt) Scale. *Assessment*, *11*(3), 251–262. http://www.sagepub.com/journal.aspx?pid=339

Barthlow, M., Jensen, R, Lunde, R., Fong, B., Savage, J., & Foster, L. (n.d.). *Quantitative research statistics resource guide* [unpublished manuscript]. Liberty University.

Bocala, C., Shakman, K., & Zweig, J. (2016). *Teacher demographics: A descriptive study in a large urban district*. https://ies.ed.gov

Bradford, C., & Braaten, M. (2018). Teacher evaluation and the demoralization of teachers. *Teaching and Teacher Education*, *75*, 49 – 59. https://doi.org/10.1016/j.tate.2018.05.017

Brewer, C., Knoeppel, R. C., & Lindle, J. C. (2015). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy, 29*(5), 711 – 745. https://doi.org/10.1177/0895904813518099

Briggs, D.C., & Dadey, N. (2017). Principal holistic judgments and high-stakes evaluations of teacher. *Educational Assessment, Evaluation and Accountability*, *29*(2), 155 – 178. https://doi.org/10.1007/s11092-016-9256-7

Brown, C., & Mlitello, M. (2016). Principal's perceptions of effective professional development in schools. *Journal of Educational Administration*, *54* (6), 703 – 726. https://doi.org/10.1108/JEA-09-2014-0109

Campbell, S.L. (2020). Ratings in black and white: A quantcrit examination of race and gender in teacher evaluation reform. Race, Ethnicity and Education, ahead-of-print (ahead- of-print), 1 – 19. https://doi.org/10.1080/13613324.2020.1842345

Campbell, J.W., & Derrington, M.L. (2019). Principal's perceptions of teacher evaluation reform from structural and human resource. *Journal of Educational Supervision*, *2*(1), 58 – 77. https://doi.org/10.31045/jes.2.1.4

Campbell, S.L., & Ronfeldt, M. (2018). Observational evaluation of teachers: Measuring more than we bargained for? *American Educational Research Journal*, *55*(6), 1233 – 1267. https://doi.org/10.3102/0002831218776216

Carreiro, D. (2020). Making a shift in educator evaluation. *Journal of Educational Supervision*, *3*(2), 97 – 104. https://doi.org/10.31045/jes.3.2.6

Cherry, B. D., Grasse, N., Kapla, D., & Hamel, B. (2017). Analysis of academic administrators' attitudes: Annual evaluations and factors that improve teaching. *Journal of Higher Education Policy and Management, 39*(3), 296 – 306. https://doi.org/10.1080/1360080X.2017.1298201

Choi, H.J., & Park, J.H. (2016). An analysis of critical issues in Korean teacher evaluation systems. *Center for Educational Policy Studies Journal*, *6*(2), 151 – 171.

Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of states' teacher evaluation systems post–Every Student Succeeds Act. *Education Policy Analysis Archives, 28*(58). https:/doi.org/10.14507/epaa.28.5252

Cohen, J. (1988). *Statistical power analysis for the social sciences* (2nd ed.). Routledge

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum.

Cohen-Vogel, L., Little, M., & Fierro, C. (2019). Evidenced based staffing in high schools: Using student achievement data in teacher hiring evaluation and assignment. *Leadership and Policy in Schools*, *18*(1), 1 – 34. https:/doi.org/10.1080/15700763.2017.1326146

Colson, T.L., & Satterfield, C. (2018). The effects of strategic compensation on teacher retention. *Power and Education*, *10*(1), 92 – 104. https:/doi.org/10.1177/1757743818658782

Conley, S.F., Mainz, E.A., & Wellington, L.D. (2019). Providing teachers with a choice in evaluation: A case study of veteran teachers' views. *Differentiated Teacher Evaluation and Professional Learning*. https:/doi.org/10.1007/978-3-030-16454-6_3

Creswell, J.W., & Guetterman, T.C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.

Cuevas, R., Ntoumanis, N., Fernandez-Bustos, J. G., & Bartholomew, K. (2018). Does teacher evaluation based on student performance predict motivation, well-being, and ill-being? *Journal of School Psychology, 68*, 154 – 162. https://doi.org/10.1016/j.jsp.2018.03.005

Curtis, E. A., Comiskey, C., & Dempsey, O. (2016). Importance and use of correlational research. *Nurse Researcher, 23*(6), 20-25. https://doi.org/10.7748/nr.2016.e1382

Davis, J., Lampley, J.H., & Foley, V. (2016). The relationship between growth scores and the overall ratings for teachers in public school system in Tennessee. *Journal of Learning in Higher Education*, *12*(1), 45 – 50.

Dean, A., Voss, D., & Draguljic, D. (2017). *Design and analysis of experiments*. Springer.

de Lima, J. Á., & Silva, M. J. T. (2018). Resistance to classroom observation in the context of teacher evaluation: Teachers' and department heads' experiences and perspectives. *Educational Assessment, Evaluation and Accountability, 30*(1), 7 – 26. https://doi.org/10.1007/s11092-017-9261-5

Derrington, M.L., & Campbell, J.W. (2018). Teacher evaluation policy tools: Principals' selective use in instructional leadership. *Leadership and Policy in Schools*, *17*(4), 569 – 590.

Derrington, M.L., & Martinez, J.A. (2019). Exploring teachers' evaluation perceptions: A snapshot. *NASSP Bulletin*, *103*(1), 32 – 50.

Derrington, M. L., Jackson, T., & Campbell, J. W. (2021). Do principals really know what their teachers believe about evaluation? exploring principals' reactions to teachers' beliefs in

the United States. *Educational Management, Administration & Leadership.* https://doi.org/174114322110519. https://doi.org/10.1177/17411432211051915

Dodson, R. L. (2015). Kentucky principal perceptions of the state's new teacher evaluation system: A survey analysis. *Educational Research Quarterly, 39*(2), 53.

Dodson, R.L. (2017). An analysis of principals' perceptions of the primary teaching evaluation system used in eight U.S. states. *International Journal of Education Policy and Leadership*, *12*(5).

Dodson, R.L. (2018). An analysis of principal perceptions of required evaluator proficiency exam used in the primary teaching evaluation in seven U.S. states. *Educational Research Quarterly*, *41*(4), 3 – 29.

Dökmen, Z. Y. (2017). Toplumsal cinsiyet sosyal psikolojik açıklamalar (society gender psychological explanations). İstanbul: Remzi Kitabevi.

Donaldson, M.L., & Papay, J.P. (2015). An idea whose time had come: Negotiating teacher evaluation reform in New Haven, Connecticut. *American Journal of Education*, *122*(1), 39 – 70. https://doi-org.ezproxy.liberty.edu/10.1086/683291

Donaldson, M., & Mavrogordato, M. (2018). Principal and teacher evaluation. *Journal of Educational Administration*, *56*(6), 586 – 601. https://doi.org/10.1108/JEA-08-2017-0100

Drake, S.A., Auletto, A., & Cowen, J.M. (2019). Grading teachers: Race and gender differences in low evaluation ratings and teacher employment outcomes. *American Educational Research Journal*, *56*(5), 1800 – 1833. https://doi.org/10.3102/0002831219835776

Duke, D., & Stiggins, R.J. (1986). *Teacher evaluation: Five keys to growth*. National Education Association.

Farinde-Wu, A., & Fitchett, P.G. (2018). Searching for satisfaction: Black female teachers' workplace climate and job satisfaction. *Urban Education*, 53, 112 – 186. https://doi.org/10.1080/13613324.2020

Farnsworth, V., Kleanthous, I., & Wenger-Trayner, E. (2016). Communities of practices as a social theory of learning: A conversation with Etienne Wenger. *British Journal of Educational Studies*, *64*(2), 139 – 160. https://doi.org/10.1080/00071005.2015.1133799

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.

Finster, M., & Milanowski, A. (2018). Teacher perceptions of a new performance evaluation system and their influence on practice: A within – and between – school level analysis. *Education Policy Archives*, *26*(41). https://doi.org/10.14507/epaa.26.3500

Flores, M. A., & Derrington, M. L. (2015). School principals' views of teacher evaluation policy: Lessons learned from two empirical studies. *International Journal of Leadership in Education, 20*(4), 416 – 431. https://doi.org/10.1080/13603124.2015.1094144

Ford, T.G. (2018). Pointing teachers in the wrong direction: Understanding Louisiana elementary teachers' use of compass high-stakes teacher evaluation data. *Educational Assessment, Evaluation & Accountability*, *30*(3), 251 – 283. https://doi.org/10.1007/s11092-018-9280-x

Ford, A. (2019). Examining and improving environments through the lens of self-determination theory. *Critical Questions in Education*, *10*(1), 65 – 77.

Ford, T.G., & Hewitt, K. (2020). Better integrating summative and formative goals in the design of next generation teacher evaluation systems. *Educational Policy Analysis Archives*, *28*(63). https://doi.org/10.14507/epaa.28.5024

Ford, T.G., & Ware, J.K. (2018). Teacher self-regulatory climate (TSRC): Conceptualizing an indicator of leader and support for the teachers' learning and development. *Leadership and Policy in Schools*, *17*(1), 27 – 51. https://doi.org/10.1080//15700763.2016.1197283

Ford, T.G., Urick, A., & Wilson, A.S.P. (2018). Exploring the effect of supportive teacher evaluation experience on U.S. teachers' job satisfaction. *Educational Policy Analysis Archives*, *26*(59). https://doi.org/10.14507/epaa.26.3559

Ford, T.G., Van Sickle, M.E., Clark, L.V., Fazio-Brunson, M., & Schween, D.C. (2017). Teacher self-efficacy, professional commitment and high-stakes teacher evaluation (HSTE) policy in Louisiana. *Educational Policy*, *31*(2), 202 – 248. https://doi.org/10.117/0895904815586855

Fradkin-Hayslip, A. (2021). Teacher autonomy, motivation, and job satisfaction: Perceptions of elementary school teachers according to self-determination theory. *Ilkogretim Online*, *20*(2), 198 – 205. https://doi.org/10.17051/ilkonline.2021.02.25

Frasier, A.S. (2021). Principal-led state evaluation policy: Teacher perceptions at four high schools. *A Qualitative Report*. https://doi.org/10.46743/2160-3715/2021.4696

Gall, M.D., Gall, J.P., & Borg, W.R. (2007). *Education research: An introduction* (8th ed.). Pearson Education, Inc.

Garrett, R., & Steinberg, M.P. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis*, *37*(2), 224 – 242.

Gaza, E. (2022). Male teachers in elementary setting: Perceptions of a teaching career. *International Journal on Social and Education Sciences* (*IJonSES*), *4*(1), 135 – 148. https://doi.org/10.10.46328/ijonses.291

Geiger, T.J., Amrein-Beardsley, A., & Holloway, J. (2020). Using test scores to evaluate and hold teachers accountable in New Mexico. *Educational Assessment, Evaluation and Accountability*, *32*(2), 187 – 235. https://doi.org/10.1007/511092.02009324-w

Gitomer, D. H. (2019) Evaluating instructional quality. *School Effectiveness and School Improvement*, *30*(1), 68-78, doi: 10.1080/09243453.2018.1539016

Glatthorn, A. A., Boschee, F., Whitehead, B. M., & Boschee, B. F. (2019). *Curriculum leadership: Strategies for development and implementation* (5th ed.). SAGE.

Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession*. Doubleday.

Grissom, J.A., & Youngs, P. (Eds.). (2016). *Improving teacher evaluation systems: Making the most of multiple measures*. New York, NY: Teacher College Press.

Gupta, A., & Lee, G.L. (2020). The effects of site-based teacher professional development program on student learning. *International Electronic Journal of Elementary Education*, *12*(5), 417 – 428. https://doi.org/10.26822/iejee.2020562132

Hanushek, E.A., Rivkin, S.G., & Schimar, T.C. (2016). Dynamic effects of teacher turnover on the quality of instruction. *Economic Education Review*, *55*, 132 – 148.

Hardy, M. A. (1993). *Regression with dummy variables*. Sage Publications.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J. (2023). Visible learning: The sequel. *A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.

Hazi, H.M., & Rucinski, D.A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, *17*(5), 1 – 19.

Heise, M. (2017). From no child left behind to every student succeeds: Back to a future for education federalism. *Columbia Law Review, 117*(7), 1859-1896.

Hess, F. M., & Noguera, P. A. (2021). *A search for common ground: Conversations about the toughest questions in K–12 education*. Teachers College Press.

Hewitt, K. K. (2015). Educator evaluation policy that incorporates EVAAS value-added measures: Undermined intentions and exacerbated inequities. *Education Policy Analysis Archives, 23*(76). http://dx.doi.org/10.14507/epaa.v23.1968

Holloway, J. (2019). Risky teachers: Mitigating risk through high-stakes teacher evaluation in the U.S.A. Discourse: *Studies in Cultural Politics of Education*, *40*(3), 399 – 411. https://doi.org/10.1080/01596306.2017.1322938

Hunter, S.B., & Rodriguez, L.A. (2021). Examining the demands of teacher evaluation: Time use, strain and turnover among Tennessee school administrators. *Journal of Educational Administration*, *59*(6), 739 – 758. https://doi.org/10.1108/JEA-07-2020-0165

Jewell, J. W. (2017). From inspection, supervision, and observation to value-added evaluation: A brief history of U.S. teacher performance evaluations. *Drake Law Review, 65*(2), 363.

Jerrim, S., & Sims, S. (2019). *The teaching and learning international survey* (*TALIS*). Department of Education: London.

Jiang, J.Y., & Sporte, S.E. (2016). Teacher evaluation in Chicago: Differences in observation and value-added scores by teacher, student, and school characteristics. *Research Report. University of Chicago Consortum on School Research*.

Kennedy, M.M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*, 945 – 980.

Khaldi, K. (2017). Quantitative, qualitative, or mixed research: Which research paradigm to use? *Journal of Educational and Social Research*, *7*(2), 15.

Khalilzadeh, S., & Khodi, A. (2018). Teachers' personality traits and students' motivation: A structural equation modeling analysis. *Current Psychology*, 1 – 16. https://doi.org/10.1007/s12144-018-0064-8

Kim, L.E., Jörg, V., & Klassen, R.M. (2019). A meta-analysis of the effects of teacher personality on teacher effectiveness and burnout. *Educational Psychology Review*, *31*(1), 163 – 195. https://doi.org/10.1007/s10648-018-9458-2

King, K. M., & Paufler, N. A. (2020). Excavating theory in teacher evaluation: Evaluation frameworks as Wengerian boundary objects. *Education Policy Analysis Archives, 28*(57). https://doi.org/10.14507/epaa.28.5020

Kirschner, P.A, & Hendrick, C. (2020). *How learning happens: Seminal works in educational psychology and what they mean in practice*. Routledge.

Koedel, C., Li, J., Springer, M.G., & Tan, L. (2017). The impact of ratings on job satisfaction for public school teachers. *American Educational Research Journal*, *54*(2), 241 – 278, https://doi.org/10.3102/0002831216687531

Kolman, J.S. (2017). Situated teacher quality: A case study of an experienced elementary school teacher. *Pedagogy, Culture and Society*, *25*(4), 501 – 514. https://doi.org/10.1080/14681366.2017.1285344

Konstantopoulos, S. (2014). Teacher effects, value-added models, and accountability. *Teacher College Record*, *116*(1), 1 – 21. https://doi.org/10.1177/016146811411600109

Koonce, M., Pijanowski, J.C., Bengston, E., & Lasater, K. (2019). Principal engagement in professional development process. *NASSP Bulletin*, *103* (3), 229 – 252. https://doi.org/10.1177/0192636519871614

Kraft, M.A., & Gilmour, A.F. (2016). Can principals promote teacher development as evaluators? A case study of principal views and experiences. *Educational Administration Quarterly*, *52*(5), 711 – 753. https://doi.org/10.1177/001361X16653445

Kraft, M.A., & Gilmour, A.F. (2017). Revisiting the widget effect: Teacher evaluation reform and the distribution of teacher effectiveness. *Educational Researcher*, *46*(5), 234 – 249. https://doi.org/10.3102/0013189X17718797

Kraft, M.A., Brunner, E.J., Dougherty, S.M., & Schwegman, D.J. (2020). Teacher accountability reform and the supply and quality of new teachers. *Journal of Public Economics*, 188. https://doi.org/10.1016/j/jpubeco.2020.104212

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Education and Policy Analysis*, *25*(3), 287 – 298. https://doi.org/10.3102/01623737025003287

Ladd, H.F., & Sorensen, L.C. (2017). Returns to teacher exoerience: Student achievement and motivation in middle school. *Education Finance and Policy*, *12*(2), 241 – 279. https://doi.org/10.1162/edfp_a_00194

Laerd Statistics (n.d.). Friedman Test in SPSS. https://statistics.laerd.com/spss-tutorials/friedman-test-using-spss-statistics.php

Lavigne, A. L., & Chamberlain, R. W. (2017). Teacher evaluation in Illinois: School leaders' perceptions and practices. *Educational Assessment, Evaluation and Accountability, 29*(2), 179 – 209. https://doi.org/10.1007/s11092-016-9250-0

Lavigne, A.L., & Olson, A.M. (2020). Teacher evaluation through the eyes of the principal: How individual and school context shape perceptions of practice? *Leadership and Policy in Schools*, *19*(4), 575 – 596. https://doi.org/10.1080/15700763.2019.1637899

Lee, S.W. (2018). Pulling back the curtain: Revealing the cumulative importance of high-performing, highly qualified teachers on students' educator outcome. *Educational Evaluation and Policy Analysis*, *40*(3), 359 – 381. https://doi.org/10.103102/0162373718769379

Lewis-Beck, M. (1980). *Applied regression: An introduction*. Sage Publications.

Lewis, J.M., Reid, D.B., Bell, C.A., Jones, N., & Qi, Y. (2020). The mantle of agency: Principals' use of teacher evaluation policy. *Leadership and Policy in Schools*, 1 – 17. https://doi.org/10.1080/15700763.2020.1770802

Liang, X., Collins, L.J., Lenhart, L., & Ressa, V. (2020). Instructional change following instructional practices professional development. *Teacher Development*, *24*(1), 108 – 125.

Lillejord, S., & Børte, K. (2020) Trapped between accountability and professional learning? School leaders and teacher evaluation. *Professional Development in Education*, *46*(2), 274-291. https://doi.org/10.1080/19415257.2019.1585384

Lindqvist, A., Sendén, M. G., & Renström, E. A. (2020). What is gender, anyway: A review of the options for operationalizing gender. *Psychology and Sexuality,* 1-13. https://doi.org/10.1080/19419899.2020.1729844

Liu, Y., Visone, J., Mongillo, M.B., & Lisi, P. (2019). What matters to teachers if evaluation is meant to help them improve? *Studies in Educational Evaluation*, *61*, 41 – 54. https://doi.org/10.1016/j.stueduc.2019.01.006

Lochmiller, C.R., & Mancinelli, J.L. (2019). Principals' instructional leadership under statewide teacher evaluation reform. *International Journal of Educational Management*, *33*(4), 629 – 643. https://doi.org/10.1108/IJEM-06-2017-0151

Lyde, A. R., Grieshaber, D. C., & Byrns, G. (2016). Faculty teaching performance: Perceptions of a multi-source method for evaluation. *Journal of the Scholarship of Teaching and Learning, 16*(3), 82 – 94. https://doi.org/10.14434/josotl.v16i3.18145

Martinez, J.F., Schweig, J., & Goldschmidth, P. (2016). Approaches for combining multiple measures of teacher performance: Reliability, validity, and implications for evaluation policy. *Educational Evaluation and Policy Analysis*, 38(4), 738 – 756. https://doi.org/10.3102/0162373716666166

Meader, K., & Larwin, K. H. (2022). Exploring the perceptions of male teachers about the shortage of male elementary school teachers: A Q Methodology. *Journal of Education*, *202*(4), 416-435. https://doi.org/10.1177/0022057421996243

Mertens, D.M. (2001). Inclusivity and transformation: Evaluation in 2010. *American Journal of Evaluation*, *22*(3), 367. https://doi.org/10.1177/109821400102200313

Mills, J.I. (2008). A legislative overview of No Child Left Behind. *New Directions for Evaluation*, *117*, 9 – 20. https://doi.org/10.1002/ev.248

Mireles-Rios, R., Becchio, J.A., & Roshandel, S. (2019). Teacher evaluations and contextualized self-efficacy: Classroom management, instructional strategies and student engagement. *Journal of School Administration Research and Development*, *4*(1), 6 – 7.

Moran, R.M.R. (2017). The impact of high stakes teacher evaluation system: Educator perspective on accountability. *Educational Studies*, *53*(2), 178 – 193. https://doi.org/10.1080/00131946.2017.1283319

Morgan, G.A., & Harmon, R.J. (2001). Data collection techniques. *Journal of the American Academy of Child and Adolescent Psychiatry*, 1 – 4.

Neumerski, C. M., Grissom, J. A., Goldring, E., Rubin, M., Cannata, M., Schuermann, P., & Drake, T. A. (2018). Restructuring instructional leadership: How multiple-measure teacher evaluation systems are redefining the role of the school principal. *The Elementary School Journal, 119*(2), 270 – 297. https://doi.org/10.1086/700597

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*, 237 – 257.

Ong, M.H.A., & Puteh, F. (2017). Quantitative data analysis: Choosing between SPSS, PLS, and AMOS in social science research. *International Interdisciplinary Journal of Scientific Research*, *3*(1), 14 – 25.

Ott, R. L., & Longnecker, M. (2017). *Statistical methods and data analysis: An introduction* (7th ed.). Cengage Learning.

Paige, M. (2020). Moving forward while looking back: Lessons learned from teacher evaluation litigation concerning value-added models (VAMs). *Education Policy Analysis Archives*, *28*, 64. https://doi.org/10.14507/epaa.285067

Papay, J.P., Bacher-Hicks, A., Page, L.C., & Marinell, W.H. (2017). The challenge of teacher retention in urban schools: Evidence of variation from cross-site analysis. Educational Research, 46(8), 434 – 448. https://doi.org/10.3102/0013189X17735812

Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high-stakes teacher evaluation system. *International Journal of Education Policy and Leadership, 13*(8). https//doi.org/10.22230/ijepl.2018v13n8a813

Paufler, N.A., & Clark, C. (2019). Reframing conversations about teacher quality: School and district administrators' perceptions of the validity, reliability, and justifiability of a new teacher evaluation system. *Educational Assessment Evaluation Accountability 31*, 33 – 60. https://doi-org.ezproxy.liberty.edu/10.1007/s11092-019-09292-w

Paufler, N.A., & Sloat, E.F. (2020). Using standards to evaluate accountability policy in context: School administrators and teachers' perceptions of a teacher evaluation system. *Studies in Educational Evaluation*, *64*. https://doi.org/10.1016/j.stueduc.2019.07.007

Peralta, M. (2013). *Design of experiments*. Texas Instruments.

Pizmony-Levy, O., & Woolsey, A. (2017). Politics of education and teacher support for high-stakes teacher accountability policies. *Education Policy Analysis Archives*, 25(87). https://dx.doi.org/10.14507/epaa.25.2892

Pogodzinski, B., Umpstead, R., & Witt, J. (2015). Teacher evaluation reform implementation and labor relations. *Journal of Education Policy: Policy Actors, 30*(4), 540 – 561. https://doi.org/10.1080/02680939.2014.999827

Pressley, T., Roehrig, A. D., & Turner, J. E. (2018). Elementary teachers' perceptions of a reformed teacher-evaluation system. *The Teacher Educator, 53*(1), 21 – 43. https://doi.org/10.1080/08878730.2017.1391362

Ravitch, D. (2016). Solving the mystery of the schools. *New York Review of Books*, *63*(5), 34 – 36.

Redding, C., & Henry, G.T. (2019). Leaving shool early! An examination of teachers' within- and end-of-year turnover. *American Educational Research Journal*, *56*, 204 – 236. https://doi.org/10.3102/0002831218790542

Reddy, L. A., Dudek, C. M., Peters, S., & Kurz, A. (2018). Teachers' and school administrators' attitudes and beliefs of teacher evaluation: A preliminary investigation of high poverty school districts. *Educational Assessment, Evaluation and Accountability, 30*(1), 47 – 70. https:/doi.org/10.1007/s11092-017-9263-3

Reid, D. B. (2020). Teachers' perceptions of how principal use new teacher evaluation system. *Teachers and Teaching, Theory and Practice*, *26*(1), 129 – 144. htpps://doi.org/10.1080/13540602.2020.1740199

Ridge, B.L., & Lavigne, A.L. (2020). Improving instructional practice through peer observation and feedback. *Educational Policy Analysis Archives*, *28*(61). https://doi.org/10.1014507/epaa.28.5023

Rigby, J.G., Larbi-Cherif, A., Rosenquist, B.A., Sharpe, C.J., Cobb, P., & Smith, T. (2017). Administrator observation and feedback: Does it lead toward improvement in inquiry-oriented math instruction? *Educational Administration Quarterly*, *53*(3), 475 – 516. https://doi.org/10.1177/0013161X16687006

Robertson-Kraft, C., & Zhang, R.S. (2018). Keeping great teachers: A case study on the impact and implementation of a pilot teacher evaluation system. *Educational Policy (Los Altos, Calif.)*, *32*(3), 363 – 394. https://doi.org/10.1177/0895904816637685

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. United Kingdom: Guilford Publications.

Saeki, E., Segool, N., Pendergast, L., & Embse, N. (2020). The influence of test-based accountability policies on early elementary teachers: School climate, environmental stress, and teacher stress. *Psychology in the Schools*, *55*(4), 391 – 403. https://doi.org/10.1002/pits.22112

Sanders, W.L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education*, *14*(4), 329 – 339. https://doi.org/10.1023/A:1013008006096

Sanders, W.L., & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, *12*(3), 247 – 256. https://doi.org/10.1023/A:1008067210518

Sartain, L., & Steinberg, M.P. (2016). Teachers' labor market responses to performance evaluation reform: Experimental evidence from Chicago public schools. *Journal of Human Resources*, *51*(3), 615 – 655. https://doi.org/10.3368/jhr.51.3.0514-6390R1

Schuls, J.V., & Flores, J.M. (2020). Improving teacher retention through support and development. *Journal of Educational Leadership and Policy Studies*, *4*(1), 1 – 19.

Schunk, D. (2020). *Learning theories: An educational perspective* (8th ed.). Hoboken, NJ: Pearson, Inc.

Shubert, W. (2018). Respective on evaluation from curricular context. *Educational Policy Analysis Archives*, *26*(47). https://doi.org/10.14507/epaa.26.3812

Skedsmo, G., & Huber, S.G. (2018). Teacher evaluation: The need for valid measures and increased teacher involvement. *Educational Assessment, Evaluation and Accountability*, *30*(1), 1 – 5. https://doi.org/10.1007/s11092-018-9273-9

Smith, W.C., & Holloway, J. (2020). School testing culture and teacher satisfaction. *Educational Assessment, Evaluation and Accountability*, *32*(4), 461 – 479. https://doi.org/10.1007/s11092-020-09342-8

Smith, E. C., Starratt, G. K., McCrink, C. L., & Whitford, H. (2020). Teacher evaluation feedback and instructional practice self-fffficacy in secondary school teachers. *Educational Administration Quarterly*, *56*(4), 671-701. https://doi.org/10.1177/0013161X19888568

Smith, W.C., & Kubacka, K. (2017). The emphasis of student test scores in teacher appraisal systems. *Educational Policy Analysis Archives*, 25(86). https://doi.org/10.14507/epaa.25.2889

Springer, M.G., Swain, W.A., & Rodriguez, A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, *38*(2), 199 – 221. https://doi.org/10.3102/0162373715609687

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

Stiggins, R.J., & Nickel, P. (1989). The teacher evaluation profile: A technical analysis. *Journal of Personnel Evaluation in Education*, *2*, 151 – 156.

Stone, J.E. (1999). Value-added assessment: An accountability revolution. *Better Teachers, Better Schools*, 239 – 249.

Stronge, J.H. (1995). Educating homeless students. *Journal for a Just & Caring Education*, *1*(2), 128 – 141.

Stronge, J.H., & Tucker, P. D. (2013). *Handbook on teacher evaluation: Assessing and improving performance*. Routledge.

Swain, W.A., Rodriguez, L.A., & Springer, M.G. (2019). Selective retention bonuses for highly effective teachers in high poverty schools: Evidence from Tennessee. *Economics of Education Review*, *68*, 148 – 160. https://doi.org/10.1016/j.econedurev.2018.12.008

Tennessee Department of Education (TDOE). (2021). *Tennessee value-added assessment*. https://team-tn.org/achievement-measures/

Tennessee Department of Education (TDOE). (2021). *Teacher evaluation*. https://team-tn.org/teacher-evaluation-2-2/

Tuytens, M., & Devos, G. (2017). The role of feedback from the school leader during teacher evaluation for teacher and school improvement. *Teachers and Teaching*, *23*(1), 6 – 24. htpps://doi.org/10.1080/13540602.2016.1203770

U. S. Department of Education. (2002). *No Child Left Behind Act of 2001*. https://www2.ed.gov/

Warner, C. K. (2016). Contested definitions of excellent teaching: An analysis of the discourse of quality. *Journal of Thought, 50*(1-2), 20-36.

Warner, R.M. (2013). *Applied statistics: From bivariate through multivariate techniques* (2nd ed.). Pearson.

Warner, R.M. (2020). *Applied statistics I: From bivariate through multivariate techniques* (3rd ed.). Sage Publications.

Warren, A.N., & Ward, N.A. (2018). This is my new normal: Teachers' accounts of evaluation policy at local school board meetings. *Journal of Education Policy*, *33*(6), 840 – 860. https://doi.org/10.1080/02680939.2017.1390164

Warsame, K., & Valles, J. (2018). An analysis of effective support structure for novice teachers. *Journal of Teacher Education and Educators*, *7*(1), 17 – 42.

Watt, H.M.G., & Richardson, P.W. (2015). Teacher motivation. *International Encyclopedia of the Social and Behavioral Sciences*, 64 – 71. https://doi.org/10.1016/B978-0-08-097086-8.268082-0

Williams, K., & Hebert, D. (2020). Teacher evaluation systems: A literature review on issues and impact. *Research Issues in Contemporary Education*, 5(1), 42 – 50.

Wieczorek, D., Aguilar, I., & Mette, I. (2022). System-level leaders' local control of teacher supervision and evaluation under Every Student Succeeds Act. *AASA Journal of Scholarship & Practice*, *19*(3), 10.

Wind, S.A., Jones, E., Bergin, C., & Jensen, K. (2019). Exploring patterns of principal judgments in teacher evaluation related to reported gender and years of experience. *Studies in Educational Evaluation*, *61*, 150 – 158. https://doi.org/10.1016/j.stueduc.2019.03.011

Woolfolk, A., & Usher, E. L. (2022). *Educational psychology* (15th ed.). Pearson.

Xu, X., Grant, L. W., & Ward, T. J. (2016). Validation of a statewide teacher evaluation system: Relationship between scores from evaluation and student academic progress. *National Association of Secondary School Principals. NASSP Bulletin, 100*(4), 203 – 222. https://dx.doi.org.ezproxy.liberty.edu/10.1177/01926365166832

Young, M.D., Winn, K.M., & Reedy, M.A. (2017). The Every Student Succeeds Acts: Strengthening the focus on educational leadership. *Educational Administration Quarterly*, *53*(5), 705 – 726.

Zayac, R.M., Poole, B.D., Gray, C., Sargent, M., Paulk, A., & Haynes, E. (2021). No disrespect: Student and faculty perceptions of the qualities of ineffective teachers. *Teaching of Psychology*, *48*(1), 55 – 62. https://doi.org/10.1177/0098628320959978

APPENDIX A: Permission to Use the Modified Teacher Evaluation Profile (TEP)

DL

Duke, Daniel [redacted]

To: Bell, Leonida

Sun 2/28/2021 5:25 AM

[EXTERNAL EMAIL: Do not click any links or open attachments unless you know the sender and trust the content.]

Dear Leonida:

Thank you for your inquiry. You have my permission to use the Teacher Evaluation Profile for your dissertation research. I wish you all the best as you pursue the final stage of your doctoral work. Please give my regards to Lisa. You have a very talented advisor, as I'm sure you realize.

Daniel L. Duke
Professor Emeritus
University of Virginia

APPENDIX B: Modified Teacher Evaluation Profile (TEP)

Participants Demographics

1. Including the current year, how many total years have you taught? *

- ◯ 0 - 3 years
- ◯ 4 - 5 years
- ◯ 6 or more years

2. Your current teaching assignment grade level (select the answer that best describes your current position) *

- ◯ Elementary (1-5)
- ◯ Middle (6 - 8)
- ◯ High (9 - 12)

3. You gender *

- ◯ Female
- ◯ Male

Overall Rating

4. Rate the overall quality of the evaluation: *

1	2	3	4	5

Very poor quality — Very high quality

5. Rate the overall impact of the evaluation on your professional practices. (Note: A rating of 5 would reflect a strong impact leading to profound changes in your teaching practices, attitudes about teaching, and/or understanding of the teaching profession. A rating of 1 would reflect no impact at all and not changes in your practices, attitudes, and/or understanding) *

1	2	3	4	5

No impact — Strong impact

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

1	2	3	4	5

No impact — Strong impact

Rating Attributes of Evaluation

A. Describe yourself in relation to the following attributes:

7. The strength of your professional expectations of yourself *

1	2	3	4	5

I demand little — I demand a great deal

8. Orientation to risk taking *

1	2	3	4	5

I avoid risks — I take risks

9. Orientation to change *

1	2	3	4	5

I am relatively slow to change — I am relatively flexible

10. Orientation to experimentation in your classroom *

1	2	3	4	5

I don't experiment — I experiment frequently

11. Openness to criticism *

1	2	3	4	5

I am relatively closed — I am relatively open

12. Knowledge of technical aspects of teaching *

1	2	3	4	5

I know a little — I know a great deal

13. Knowledge of curriculum content *

1	2	3	4	5

I know a little — I know a great deal

14. Experience with teacher evaluation prior to most recent experience *

1	2	3	4	5

Waste of time — Very helpful

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

15. Credibility as a source of feedback *

1	2	3	4	5

Not credible — Very credible

16. Working relationships with you *

1	2	3	4	5

Adversary — Helper

17. Level of trust *

1	2	3	4	5

Not trustworthy — Trustworthy

18. Interpersonal manner *

1	2	3	4	5

Threatening — Not threatening

19. Temperament *

1	2	3	4	5

Impatient — Patient

20. Flexibility *

1	2	3	4	5

Rigid Flexible

21. Knowledge of technical teaching *

1	2	3	4	5

Not knowledgeable Very knowledgeable

22. Capacity to model or demonstrate needed improvement *

1	2	3	4	5

Low High

23. Familiarity with your particular teaching assignment *

1	2	3	4	5

Unfamiliar Very familiar

24. Usefulness of suggestions for improvement *

1	2	3	4	5

Useless Very useful

25. Persuasiveness of rationale for suggestions *

1	2	3	4	5

Not persuasive Very persuasive

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

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

26. Were standards communicated to you? *

1	2	3	4	5
Not at all				In a great detail

27. Were standards clear to you? *

1	2	3	4	5
Vague				Very clear

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

1	2	3	4	5
Not endorsed				Highly endorsed

29. Were the standards.....

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

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

30. Observation of your classroom performance *

1	2	3	4	5

Not considered Used extensively

31. Meetings with evaluator *

1	2	3	4	5

Not considered Used extensively

32. Examination of student performance *

1	2	3	4	5

Not considered Used extensively

33. Student evaluations *

1	2	3	4	5

Not considered Used extensively

34. Self-evaluations *

1	2	3	4	5

Not considered Used extensively

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

35. Number of formal observations per year *

- ◯ 0 observations
- ◯ 1 observation
- ◯ 2 observations
- ◯ 3 observations
- ◯ 4 observations

36. Approximate frequency of informal observations per year *

- ◯ 0 observation
- ◯ 1 observation
- ◯ 2 observations
- ◯ 3 observations
- ◯ 4 observations

Describe the attributes of the feedback you received during your last evaluation experience.

37. Amount of information received *

1	2	3	4	5

None Great deal

38. Frequency of formal feedback *

1	2	3	4	5

Infrequent Frequent

39. Frequency of informal feedback *

1	2	3	4	5

Infrequent Frequent

40. Depth of informal feedback *

1	2	3	4	5

Shallow In-dept

41. Quality of ideas and suggestions contained in the feedback *

1	2	3	4	5

Low High

42. Nature of information provided *

1	2	3	4	5

Judgmental Descriptive

43. Timing of feedback *

1	2	3	4	5

Delayed Immediate

44. Feedback focused on standards *

1	2	3	4	5

I ignored the standards Reflected the teaching standards

Please describe these attributes of the evaluation context. Resources available for evaluation:

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

1	2	3	4	5

None — Great deal

46. Time allotted during the semester for professional development *

1	2	3	4	5

None — Great deal

District values and policies in evaluation

47. Clarity of policy statements regarding purpose of evaluation *

1	2	3	4	5

Vague — Very clear

48. Intended role of evaluation *

1	2	3	4	5

Teacher accountability — Teacher growth

APPENDIX C: IRB Approval

LIBERTY UNIVERSITY.
INSTITUTIONAL REVIEW BOARD

February 6, 2024

Leonida Bell
Jeffrey Savage

Re: IRB Exemption - IRB-FY23-24-1096 Predicting Educator Perceptions of Tennessee's Teacher Evaluation System: The Role of Gender, Years of Experience, and School Level Taught

Dear Leonida Bell, Jeffrey Savage,

The Liberty University Institutional Review Board (IRB) has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and no further IRB oversight is required.

Your study falls under the following exemption category, which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:104(d):

Category 2.(i). Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met:
The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;

For a PDF of your exemption letter, click on your study number in the My Studies card on your Cayuse dashboard. Next, click the Submissions bar beside the Study Details bar on the Study details page. Finally, click Initial under Submission Type and choose the Letters tab toward the bottom of the Submission Details page. Your information sheet and final versions of your study documents can also be found on the same page under the Attachments tab.

Please note that this exemption only applies to your current research application, and any modifications to your protocol must be reported to the Liberty University IRB for verification of continued exemption status. You may report these changes by completing a modification submission through your Cayuse IRB account.

If you have any questions about this exemption or need assistance in determining whether possible modifications to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,
G. Michele Baker, PhD, CIP
Administrative Chair
Research Ethics Office

APPENDIX D: Participant Consent Form

Information Sheet

Title of the Project: Predicting Educator Perceptions of Tennessee's Teacher Evaluation System: The Role of Teacher Gender, Years of Experience, and School Level Taught
Principal Investigator: Leonida Bell, Doctoral Candidate, Doctor of Education, Liberty University

Invitation to be Part of a Research Study

You are invited to participate in a research study. To participate, you must be a licensed teacher in Tennessee. Taking part in this research project is voluntary.

Please take time to read this entire form and ask questions before deciding whether to take part in this research.

What is the study about and why is it being done?

The purpose of the study is to determine if a predictive relationship exists between teacher perceptions of Tennessee's teacher evaluation system and the linear combination of teacher gender, years of experience, and school level taught (elementary, middle, high).

What will happen if you take part in this study?

If you agree to be in this study, I will ask you to do the following:

1. Participate in an online survey that will take approximately 10-15 minutes.
2. Please click the survey link provided in the attached information sheet in my recruitment e-mail.

How could you or others benefit from this study?

Participants should not expect to receive a direct benefit from taking part in this study.

Benefits to society include teachers' influence to engage in the revamping process should policymakers can seize the opportunity to redesign the teacher evaluation system.

What risks might you experience from being in this study?

The expected risks from participating in this study are minimal, which means they are equal to the risks you would encounter in everyday life.

How will personal information be protected?

The records of this study will be kept private. Research records will be stored securely, and only the researcher will have access to the records. Participant responses will be anonymous. Data