2019

Teacher Perspectives on Performance Pay in One Southeastern State

Wendy B. Firtell

Teacher Perspectives on Performance Pay in One Southeastern State

by
Wendy Firtell

An Applied Dissertation Submitted to the
Abraham S. Fischler College of Education
and School of Criminal Justice in partial
Fulfillment of the Requirements for the
Degree of Doctor of Education

Nova Southeastern University
2019

Approval Page

This applied dissertation was submitted by Wendy Firtell under the direction of the persons listed below. It was submitted to the Abraham S. Fischler College of Education and School of Criminal Justice and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Nova Southeastern University.

Sandra Duncan, EdD
Committee Chair

Roberta Schomburg, EdD
Committee Member

Kimberly Durham, PsyD
Dean

Statement of Original Work

I declare the following:

I have read the Code of Student Conduct and Academic Responsibility as described in the *Student Handbook* of Nova Southeastern University. This applied dissertation represents my original work, except where I have acknowledged the ideas, words, or material of other authors.

Where another author's ideas have been presented in this applied dissertation, I have acknowledged the author's ideas by citing them in the required style.

Where another author's words have been presented in this applied dissertation, I have acknowledged the author's words by using appropriate quotation devices and citations in the required style.

I have obtained permission from the author or publisher—in accordance with the required guidelines—to include any copyrighted material (e.g., tables, figures, survey instruments, large portions of text) in this applied dissertation manuscript.

Wendy Firtell
Name

September 5, 2019
Date

Acknowledgments

This journey would not have been possible without the support of my family, friends, professors and mentors. To my family, thank you for encouraging me in all of my aspirations and inspiring me to follow my dreams. Thank you to my sister Heidi who has always been my role model and best friend. Thank you to father and mother, who both left this earth during this venture and who encouraged me to complete my degree when I thought I could not continue. I want to especially thank my mom for just being the most important person in my life.

I would especially like to thank the professors of the Fischler School of Education and School of Criminal Justice at Nova Southeastern University for educating me over the last seven years and showing me what it means to be a dedicated student, each in their own unique way. Dr. David Gorman, Dr. Sandra Duncan, and Dr. Aliza Berger thank you for listening, offering advice and guiding me through this arduous process of phone calls, emails, texts and editing advice and for that extra push that I sometimes needed.

Thank you to my son Justin for always being by my side and for making me proud to be his mom. A special thank you goes out to my partner in life, Jesstina. You have enriched my life in more ways that you can ever imagine both spiritually and emotionally and showed me what it is like to have complete dedication in life.

Abstract

Teacher Perspectives on Performance Pay in One Southeastern State. Wendy Firtell, 2019: Applied Dissertation, Nova Southeastern University, Abraham S. Fischler College of Education and School of Criminal Justice. Keywords: value-added model, performance pay, expectancy, developmental growth practice target, perceptions, florida standards assessment, stanford achievement test, english language learners, english as a second language

This study investigated teachers' perceptions of a teacher performance-pay initiative. Despite implementing a performance pay initiative at the beginning of the 2015-16 school year, the target school district had not investigated teachers' perceptions of the initiative. The researcher included a convenience sample of 54 teachers who worked at a Title 1 elementary school located in the southeastern United States.

This applied dissertation used a descriptive survey research design. The researcher found Pre-K, kindergarten, and Grade 1 teachers were slightly more supportive of performance pay initiatives than Grade 2, Grade 3, Grade 4, and Grade 5 teachers (TPPP difference of 0.17 between the two groups); however, the difference between the two groups was not significant and the effect size was small. The researcher also found that non-tenure teachers were slightly more supportive of performance pay initiatives than tenured teachers (TPPP difference of 0.10 between the two groups); however, the difference between the groups was not statistically significant. Further analysis of the findings revealed a negative correlation between years of teaching experience and perceptions of performance pay as the number of teaching years increased teachers' perceptions of performance pay decreased.

The use of convenience sampling procedures and the data collection and analysis procedures were limitations. Future research should replicate this study in other settings across the target state and use random sampling procedures. Future research should also use a qualitative approach as a methodology to investigate teachers' perceptions of performance pay.

Table of Contents

Appendices

Tables

Chapter 1: Introduction

Statement of the Problem

Effective teaching is a critical component in improving student achievement (M. Jones, 2013; Vacca, 2016; Yuan et al., 2013). Recognizing the profound impact that effective teaching has on student achievement, the Florida legislature established the Florida School Recognition Program (FSRP) to provide monetary awards to public schools as well as charter schools, state university systems, and developmental research K-12 schools that received a grade of "A" or improved one letter grade over the previous academic year (Florida Department of Education, 2014). For the 2011-14 school years, schools in Florida received an overall competency grade of A through F based on their Florida Comprehensive Assessment Test (FCAT) scores. For the 2016 school year, these schools received an overall competency grade of A through F based on their Florida Standards Assessment (FSA) scores (Florida Department of Education, 2016).

The topic. The No Child Left Behind Act, The Florida Student Recognition Program, Every Student Succeeds Act (ESSA) and Florida's Race to the Top initiatives are national and statewide school reforms that have evolved in response to an increase in public and government demand for more accountability in schools (Carlon, 2015; Yuan et al., 2013). Each of these reforms has provided a performance or merit-based financial incentive for school employees that states have adopted (Jones, 2013; Rice, Malen, Jackson, & Hoyer, 2015; Yuan et al., 2013). These incentive plans compensate teachers based on their demonstrated ability to improve students' standardized test scores, and researchers found these incentive plans prompted considerable debate among education policymakers (Carlon, 2015; Rice et al., 2015).

Background and Justification

Although the state of Florida has allocated millions of dollars to implement teacher performance pay, the statute creating the Florida School Recognition Program (FSRP) initiative did not include a provision for evaluating either its effectiveness or the effectiveness of teacher performance (Florida Department of Education, 2008). Prior to implementing performance pay in public schools, the state of Florida assigned a competency letter grade to schools based on their FCAT and FSA scores. To be eligible for the FSRP initiative, schools must meet their state-projected Annual Measurable Objectives (AMOs), which replaced the Adequate Yearly Progress (AYP) measurement (Pressley, 2015). Each school has state-projected AMOs which translate into the assigned competency letter grade. Schools that meet their designated AMOs each year receive financial incentives (Laliberte, 2015)). During the 2014-2015 school year, schools across the state of Florida factored in students' learning gains on statewide exams into the teacher's value-added model (VAM) overall score.

The state of Florida established the FSRP to provide financial awards to K-12 public and charter schools and publicly funded developmental research K-12 schools that received a grade of A or made a year-over-year improvement of at least one letter grade under Florida's A++ plan (Laliberte, 2015). Despite earmarking millions of dollars to implement teacher performance pay, the FSRP initiative did not include a provision for determining the effectiveness of either FSRP or Florida's K-12 teachers (Florida Department of Education, 2008). Prior to implementing FSRP in its public schools, the state of Florida assigned a competency letter grade to schools based on their students' FCAT scores, with A being the highest grade and F being the lowest grade schools

received (Rouse, Hannaway, Goldhaber, & Figlio, 2013). To be eligible for the FSRP initiative, schools must meet their state-projected Annual Measurable Objectives (AMOs). The AMOs replaced the adequate yearly progress measurement associated with NCLB. Each school has state-projected AMOs that correspond to an assigned competency letter grade, with A representing the highest grade a school can receive and F representing the lowest. School meeting their designated annual AMOs receive financial incentives (Cocke, 2014).

The research problem. The problem was that despite implementing teacher performance pay in 2015, the target school district had not investigated teachers' perceptions of the teacher performance pay initiative (Francilus, 2015). According to the Assessment, Research, and Data Analysis Division (2014) for the target school district, the target elementary school's FCAT English and math scores decreased from 567 in 2012 to 513 in 2013. In 2014, another decrease occurred as the target school's FCAT English and mathematics score was 504 (Assessment, Research, and Data Analysis Division, 2014). In 2014, the state of Florida stopped administering the FCATs and piloted the FSA. In 2015, the target school's FSA English and mathematics scores were 388. In the summer of 2015, the target school implemented a teacher performance pay incentive (Francilus, 2015). Determining teachers' perceptions of the teacher performance pay initiative may have provided a foundation for conducting further research.

The target elementary school received either an "A" or a "B" during those five school years. For the years in which the target school received an A, the target school's Educational Excellence School Advisory Council (EESAC) selected a committee to

determine an equitable distribution of the financial award. The committee recommended that the target school's teachers receive 100% of the award money. This allocation of award money to teachers represented a form of performance pay. Proponents of performance pay argue that incentivizing teacher pay will improve not only teacher quality but also student achievement (Marsh, 2014).

Value-added model. Prior to the FSRP initiative, principals were primarily responsible for evaluating and observing Florida teachers, and they did not use students' FCAT scores to determine teachers' performance. Beginning in 2011, Florida's legislature passed the Student Success Act (SSA), also referred to as Senate Bill 736, mandating that schools base at least 50% of educators' evaluations on students' performance on state standardized tests (Harrison & Cohen-Vogel, 2012; Vacca, 2016). In response to this mandate, the state of Florida developed a Value Added Model (Cocke, 2014), which is a statistical model that determines the extent to which teachers affect student learning by determining the difference between students' predicted scores at the beginning of the school year and their actual FCAT scores at the end of the school year (Haertel, 2013; Vacca, 2016). One of the distinguishing characteristics of the Value-Added Model is its ability to measure teacher impact while controlling for sociodemographic factors such as income level, race, ethnicity, language, family background, and prior educational history (Pivovarova, Amrein-Beardsley, & Broatch, 2016).

During the 2016-2017 school year, the VAM ratio was changed to account for 34% of each teacher's year-end evaluation while teacher observations and the Deliberate Growth Practice Target to account for 50% and 16% according to Senate Bill 736. Both

the Florida's SSA in 2011 and Florida's Race to the Top initiative have linked performance pay awards to the overall teacher ratings as determined by the VAM, the results of teacher observations, and DGPTs (Vacca, 2016). Under the current teacher evaluation model, teachers with a rating of 92.4-100 are highly effective, those with a rating of 62.5-92.3 are effective, those with a rating of 50-62.4 need improvement (developing if they are in their first three years of teaching), and those with a rating below 50 are unsatisfactory. The target elementary school teachers understand their students' FSA scores significantly impact their final performance evaluation and have expressed their concerns regarding the development of the VAM and its relationship to their students' FCAT and FSA scores. One teacher at the target school was concerned about the amount of performance pay she would receive if her students' test scores accounted for 50% of yearly evaluation. She expressed concern that many of her students lacked the ability to be successful on the FSA (M. Mejia, personal communication, April 4, 2014).

During the 2017-18 school year, when full implementation occurred, the target school's teachers who received a rating of highly effective received more performance pay than the target school's teachers who received a rating of either effective or developing. Although the target district's school board and the teacher's union only recently negotiated the salary portion of the annual contract, the target school district's leaders believed financial incentives would compel teachers to increase their effectiveness. According to Carlon (2015), the resultant goal of teacher performance pay is to increase student achievement. Some critics such as Rice et al. (2015) and colleagues have noted the shortsighted nature of a model that rewards teachers based on their ability

to increase student achievement. These critics asserted that a performance-based model does not consider student factors beyond a teacher's scope of influence yet profoundly impact students' achievement (Carlon, 2015; Rice et al., 2015). Examples of factors outside the scope of a teacher include students' home lives, cognitive capabilities, and socioeconomic status (Yuan et al., 2013).

Researchers have identified and discussed several logistical issues pertaining to the implementation of performance pay for teachers (Marsh, 2014; Pivovarova et al., 2016; Vacca, 2016; Yuan et al., 2013). Marsh (2014) and Yuan et al. (2013), for example, believed using standardized test scores as the primary criteria for determining bonus pay was shortsighted because inherent variables such as socioeconomic status and mobility impacted student achievement. Pivovarova et al. (2016) discussed VAM's reliability issues in determining teacher effectiveness, noting the VAM did not consider causative factors, which resulted in an imprecise measure of teacher effect. Finally, Marsh (2014) concluded that a performance-pay plan would promote competition among teachers and thereby impact professional collegiality within the school building.

Deficiencies in the evidence. In an era characterized by increased accountability for teachers, Viscardi (2014) found there was limited research investigating teachers' perceptions of performance pay. Marsh (2014) recommended expanding the body of knowledge regarding teachers' perceptions of performance pay by investigating the perceptions of teachers who work in urban settings. Stephens (2015) and Viscardi (2014) stressed that performance pay initiatives would not succeed unless educational policymakers ascertained a precise understanding of the impact of these initiatives on teacher motivation. Their research contributed to the overall body of knowledge by

investigating urban elementary school teachers' perceptions of performance pay and providing local stakeholders at both the school and district level with a precise understanding how performance pay impacts teacher motivation.

Audience

The audience for this applied dissertation was teachers, school- and district-level administrators from the target school district as well as state educational policymakers. As of 2018, there was a lack of knowledge across the school district regarding teachers' perceptions of performance pay (Robertson-Kraft, 2014). Conducting a study of this nature could enable educational stakeholders at both the district- and school-level to develop a clear understanding of the extent to which one elementary school's teachers perceived that performance pay impacted their job performance across a range of dimensions. Given the amount of money the state earmarks for teacher performance pay, state educational policymakers should consider teachers' perceptions of the state's performance pay initiative. Conducting a study of this nature could prompt the target school's district-level administrators to expand the scope of this study by investigating teachers' perceptions of performance pay across the entire school district. Similarly, educational stakeholders could investigate teachers' perceptions across the state of Florida.

Definition of Terms

For this proposed applied dissertation, the following terms were defined to give the reader an understanding of the context in which words were used or their usual or unrestricted meaning within the context of the study:

Accountability. Accountability is the willingness of individuals or organizations

to accept responsibility for their actions. Within the framework of public education, accountability is the willingness of public-school districts and their respective schools to be held responsible for the delivery of educational services to and the academic achievement of their students (K. Jones, 2014).

Assessment. An assessment is a process wherein educators incorporate various methods to assess both student learning and teacher effectiveness. The methods can be summative, diagnostic, or formative (Carlon, 2015).

Expectancy. The construct of expectancy refers to the extent to which individuals perceive their efforts will produce the desired results; hence, effort and performance are inextricably linked (Vroom, 1964).

Expectancy Theory. Vroom's (1964) expectancy theory is an outcome-based theoretical model that recognizes the inextricable link between the level of effort individuals put forth to complete a task and their level of motivation to complete the task successfully. Galvanizing this link are the variables of expectancy, instrumentality, and valence, which work synergistically to influence individuals' beliefs that their efforts will produce the desired results.

Florida Comprehensive Assessment Test. The FCAT is a standardized test administered in the state of Florida to students in grades 3-11. The FCAT consists of criterion-referenced items in mathematics, reading, science, and writing. The FCAT measures student progress towards the Sunshine State Standards (Florida Department of Education, 2012).

Florida Student Recognition Program. Created by the Florida Legislature in 1997, the FSRP provides public recognition and financial awards to schools that sustain

high student performance or schools that demonstrate substantial improvement in student performance (Florida Department of Education, 2012).

Incentive pay. Researchers defined incentive pay as a plan to pay teachers based on their demonstrated competence in teaching (Pivovarova et al., 2016; Vacca, 2016). Throughout this study, the researcher used incentive pay interchangeably with both pay for merit and pay for performance.

Instrumentality. The construct of instrumentality refers to the likelihood that individuals believe their rewards are commensurate with their performance (Parijat & Bagga, 2014). Within education, teachers believe they will receive additional pay or a promotion if they successfully complete a task (Marsh, 2014).

Merit pay. Gius (2013) defined merit pay as an incentive plan wherein teachers receive additional pay for meeting performance criteria.

Motivation Theory. In the context of the teaching profession, motivation theory is a framework for understanding the factors influencing teachers' desire to enter the teaching profession and their willingness to remain in their initial teacher position, and the extent to which they contribute to their professional growth and the overall teaching profession (Han & Yin, 2016).

No Child Left Behind (NCLB). Signed into law by Congress in 2002, NCLB required schools to have highly qualified teacher in the core subjects in every classroom and held schools accountable for students' results, gave states and districts flexibility regarding how they spent federal money, required schools to use research-based approaches to guide classroom practices, and compelled schools to involve parents in their children's education through regular progress-monitoring measures and

communication initiatives (Florida Department of Education, 2012).

Race to the Top. A significant part of the American Recovery and Reinvestment Act of 2009, the Race to the Top program is a competitive grant program where in the U.S. Department of Education rewards states for implementing significant reforms in the following areas: (a) enhancing standards and assessments; (b) improving the collection and use of data; (c) increasing teacher effectiveness; (d) achieving equity in teacher distribution; and, (e) turning around struggling schools (Meier & Rutherford, 2016).

Valence. The construct of valence is the extent to which individuals desire a goal or outcome. Valence is positive when individuals desire to achieve a goal and negative when they do not (Vroom, 1964). Valence among teachers is positive when teachers value the rewards received for meeting their desired goals **(**Marsh, 2014).

Value Added Model. The VAM is a statistical tool school districts use to determine the extent to which teacher instruction impacts student learning over the course of a particular time period (Pivovarova et al., 2016).

Purpose of the Study

The purpose of this quantitative study was to investigate teachers' perceptions of a teacher performance-pay initiative. The participants taught at a Title 1 elementary school that is part of an urban public-school district in Florida. Despite implementing a teacher performance-pay initiative, the target school district had not assessed teachers' perceptions of the initiative (Francilus, 2015). Prior studies conducted in states such as Texas, Colorado, and Arizona found teachers had negative perceptions of performance-pay initiatives that used students' test scores as the primary criteria for determining teacher effectiveness (Davis, 2014; Munroe, 2017; Toch, 2016).

Chapter 2: Literature Review

The literature in this study describes several commonly accepted theoretical traditions that support the appropriateness of performance pay. These traditions include the relative success of performance pay in other fields will transfer to education; performance pay as a key incentive and retainer of quality teachers and an identifier of low-quality instructors; performance pay for educators and the affirmative effect on student achievement.

Theoretical Framework

The theoretical framework for this quantitative study was Vroom's (1964) expectancy theory. Vroom's based his theory on the behaviorist belief that employees react favorably to carefully calibrated pay incentives (Marsh, 2014). Vroom concluded that the prospect of receiving financial rewards motivated employees to improve their job performance (Liang & Akiba, 2015). Vroom identified and described three foundational variables in his expectancy theory: expectancy, instrumentality, and valence. Although Vroom noted that each variable had a distinct influence on individuals' motivation and performance, he concluded the combined effect of the variables was synergistic (Gemeda, 2015). According to Parijat and Bagga (2014), expectancy is individuals' beliefs that their efforts will enable them to attain their desired goals. Instrumentality is the belief that if an individual reaches the performance expectation, he or she will receive the reward. Valence refers to the value individuals affix to rewards (Marsh, 2014; Vroom, 1964).

Expectancy. Yuan et al. (2013) defined expectancy as the strength of the relationship between the level of effort teachers put forth to achieve student outcomes

and the extent to which they perceive their efforts will increase student achievement. Expectancy values range from 0 to 1, with 0 being the lowest level of expectancy and 1 being the highest (Gemeda, 2015; Rice et al., 2015). Educators are unlikely to put forth the requisite effort to increase student achievement if they believe they will not be rewarded; in contrast, educators are likely to put forth the necessary effort if they believe they will receive financial incentives for improving student achievement (Liang & Akiba, 2015; Parijat & Bagga, 2014). Researchers have noted that teachers who are unlikely to put forth the effort have low levels of expectancy while those who are likely to put forth the effort have high levels of expectancy (Marsh, 2014; Parijat & Bagga, 2014). Teachers exhibiting low levels of expectancy may perceive that factors outside of their locus of control negatively impact their ability to improve student achievement; conversely, teachers exhibiting high levels of expectancy may be undeterred by factors out of their control and perceive they possess the necessary skills to improve student achievement (Britton & Propper, 2016).

Low levels of self-efficacy, a lack of professional resources, and inadequate levels of administrative support are contributing factors to low levels of expectancy among educators (Yuan et al., 2013). Researchers found school districts could mitigate the potential impact of those factors by clearly defining expectations, by providing precise and consistent feedback, by valuing the efforts and contributions of teachers, and by rewarding teacher when they meet performance expectations (Wells, Combs, & Bustamante, 2013; Yuan et al., 2013). Giving teachers opportunities to acquire the requisite knowledge and skills they perceive as necessary is a mechanism for promoting high levels of expectancy (Britton & Propper, 2016; Lowe, 2013).

Instrumentality. The construct of instrumentality is a measure of individuals' belief that their performance will produce their desired outcomes (Rice et al., 2015; Yuan et al., 2013). The application of this construct in an educational setting is the extent to which teachers believe in their abilities to improve student achievement will result in them receiving financial incentives (Parijat & Bagga, 2014; Rice et al., 2015). Lowe (2013) stated the instrumentality "reflects the notion that acquiring the knowledge and skills valued by the school and positive consequences such as receiving a pay increase or seeing an increase in student learning are strongly connected" (p. 17). Similar to expectancy, instrumentality values range from 0 to 1, with 0 being the lowest level of instrumentality and 1 being the highest (Ozoemena, 2013; Rice et al., 2015). Undergirding this construct are the following three factors: (a) transparency throughout all phases of the process; (b) an unwavering trust in those who determine who receives what; and, (c) a clear understanding of both the standard of performance and the criteria by which leadership will determine financial rewards are the factors influencing individuals' level of instrumentality (Lowe, 2013; Yuan et al., 2013).

Having a clear understanding of the criteria by which school leadership determine teachers' performance and their potential for receiving financial awards is a strong determinant of instrumentality (Forand, 2012; Parijat & Bagga, 2014). Using a reliable instrument to measure teacher effectiveness, having instruments that account for factors such as children's socioeconomic status and cognitive abilities, and giving teachers a voice in the type of instrument are the hallmarks of a valid and reliable teacher evaluation model (Yuan et al., 2013). Likewise, using a reliable instrument gives credence to the idea of performance pay for teachers in that the system used to measure performance

must be both valid and reliable as well as take into account all of the factors influencing student achievement. Instrumentality may also have a direct effect on the third variable of Vroom's (1964) expectancy theory, which is valence. Because multiple outcomes occur from either meeting or not meeting the school's student achievement goals, there is a separate instrumentality belief for each outcome, and each outcome must have valence or a degree of desirability or undesirability to the teacher (Parijat & Bagga, 2014).

Valence. In order to be motivated by a reward, a person must perceive the reward as highly desirable (Deckers, 2014). Valence refers to whether or not teachers value the rewards associated with obtaining a desired goal, which for teachers is to improve students' test scores. An individual can experience varying levels of valence (Deckers, 2014; Najera, 2017). Najera (2017) stated, "An individual's perception of valence in achievement or task completion is a function or reflection of their own specific needs, goals, values, and preferences" (p. 26). If teachers believe their effort will help them achieve their desired goal and lead to a particular outcome, which in the case of teachers is either higher pay or other tangible rewards, then their motivation will have greater valence (Najera, 2017). Olcum and Titrek (2015) described motivation as a synthesis of the constructs of valence and expectancy. Deckers (2014) found that valence was a predictive factor in an individual's willingness to select a goal, with rewards that produce high levels of valence having a greater likelihood of being selected than those having low levels of valence.

Accountability in Public Education

In the 1800's, Horace Mann's mission was to ensure that every child received a basic education funded by local taxes (Messerli, 1972). As the father of the modern-day

public-school movement, Mann based his devotion on a belief of political stability and social harmony, both of which were dependent upon education. Approximately 200 years later, a free and appropriate public education became not only a necessity but also a world-wide expectation. From its development in 1867 as an agency within the Department of the Interior, the United States Department of Education (USDOE) has continuously advocated for education to be a right of all citizens (Messerli, 1972). However, Finnigan and Gross (2007) stressed that the teaching profession has gone largely unjudged for student performance because of complexities regarding the teacher evaluation process and political agendas.

The No Child Left Behind Act included over 1000 pages of mandates regarding teacher accountability, school funding, classroom instruction, and teacher quality. In 2009, the American Recovery and Reinvestment Act was announced by President Barack Obama and the Race to the Top Initiative soon followed (Lavania, Cohen-Vogel, & Lang, 2015). As part of this initiative and as a means to improve students' test scores and their overall achievement, the United States Department of Education encouraged states to devise performance plans for teachers (Lavania et al., 2015). The VAM is a mechanism for determining teachers' effect on students' test scores and their overall academic achievement (McCullough, English, Angus, & Gill, 2015).

School accountability in Florida. The school accountability era officially began for the state of Florida with the enactment of the Educational Accountability Act (Florida Educational Accountability Act, 1971). The implementation of a Statewide Assessment Program (SAP) was an integral component of the Educational Accountability Act of 1971 (Florida Educational Accountability Act, 1971). During the 1970's and 1980's, state

legislators the SAP program several times to address school accountability at the elementary and secondary levels. Statewide initiatives such as the FSRP reward teachers based on the average performance of students in the school (Buddhi, 2007).

Value-Added Model

The state of Florida developed Senate Bill (SB) 736 in an effort to improve the evaluation process, compensation system, and employment practices for educators. A component of this bill was written to revise and create a teacher evaluation system wherein 50% of the evaluation included student standardized test data as factors in determining teacher performance (Cocke, 2014). State legislators called the revision VAM (Guerere, 2013). Sanders developed VAM in 1992 to measure and improve both school and teacher performance (Wesson, Potts, & Hill, 2015). Schools used the VAM to disseminate data to the public and to appropriate educational stakeholders regarding the performance of school districts and their respective schools in the state of Tennessee (Owens, 2013). Schools in the state of Tennessee used the VAM as a mechanism to measure students' year-over-year growth and as a component of their teacher evaluation model (McCullough et al., 2015). School-level administrators can use the VAM to help identify teachers' areas of strength and weakness and to determine the types of professional development teachers need (McCullough et al., 2015; Wesson et al., 2015).

At the center of the educator accountability movement are the statistical models school districts across the United States use to measure teacher performance (McCullough et al., 2015; Perry, 2016). While statistical models vary from both district to district and state to state, the premise is the same: to hold teachers accountable for those aspects of student achievement they directly influence (Ballou & Springer, 2015;

McCullough et al., 2015). Proponents of the VAM believe it is a reliable measure of teacher effectiveness because of its ability to take into account the following factors: (a) students' prior level of achievement, (b) the number of courses within a subject area wherein students enroll, (c) students with disabilities (SWD) status, (d) English language learners (ELL) status, (e) gifted status, (f) school attendance, (g) school mobility (number of transitions), (h) grade retention (measured by difference from the most common age in grade), (i) class size, and (j) similarity of the test scores of students enrolled in the same class (Ballou & Springer, 2015; Darling-Hammond, 2015). These proponents also believe the VAM is a reliable measure for determining performance-based pay because it concurrently measures teachers' effect on students' academic performance in relation to where the students begin and controls for factors such as socioeconomic status, race, gender, class size, and school differences (Darling-Hammond, 2015; McCullough et al., 2015).

Researchers found there were benefits as well as disadvantages of using statistical models such as VAM to determine a teacher's value-added score (McCullough et al., 2015; Loeb, 2013; Perry, 2016). Loeb (2013) concluded that the VAM is an effective mechanism for improving both student and school achievement and a reliable tool for determining performance-based pay for teachers. Conversely, Collins and Amrein-Beardsley (2014) found varying levels of reliability across VAM models were predictive factor of variance in teacher ratings by as much as 50% from one school year to the next one. Collins and Amrein-Beardsley also noted that under these conditions, teachers could receive a high VAM score one year and a rating of highly effective and a low VAM score in the subsequent one and a rating of ineffective. Researchers underscored the importance

of developing a valid, reliable, and objective VAM across multiple states and having precise criteria for the extent to which students' test scores impact teachers' VAM scores and consequently their yearly evaluations (Darling-Hammond, 2015; McCullough et al., 2015; Moss & Haertel, 2016).

Value-Added Model: Target District

At the conclusion of the school year, the target school district's research department publishes a VAM Detail Report wherein the district classifies teachers into one of the VAM models (Assessment, Research, and Data Analysis Division, 2014). District administrators use the Florida VAM and one of the district's two VAM models when determining core teacher outcomes; however, they use reading and language arts assessments and one of the district's two VAM model assessments when determining non-core teacher outcomes (Research Services, Dade County Public Schools, 2017).

Core Teachers. The Florida VAM determines the teacher outcomes for the following: (a) reading and English/language arts (ELA) in grades 4 through 10 on the basis of students' FSA scores, (b) mathematics in grades four through eight using the FSA, and (c) Algebra 1 in grades eight and nine using only the end-of-course (EOC) assessment (Assessment, Research, and Data Analysis Division, 2014). Additionally, the target district administrators use the district covariance adjustment VAM to determine teacher outcomes for the following: (a) reading and mathematics in kindergarten through third grade, using the Stanford Achievement Test for kindergarten, first grade, and second grade and the FSA for third grade; (b) science in grades five and eight using the EOC, (c) civics in grade seven using the EOC, (d) geometry in grades 8 through 10 using the EOC, (e) algebra 2 in grades 9 through 12 using the EOC, (f) biology in grades 8

through 11 using the EOC, and (g) U.S. History in Grade 11 using the EOC and certain advanced placement courses and advanced placement exams. Lastly, the target district uses the district achievement VAM to determine teacher outcomes for the following: (a) other courses using advanced placement, International Baccalaureate, and Advanced International Certification Examination; (b) certain courses using industry certification exams, and (c) exceptional student education courses using the Florida Standards Alternate Assessment.

Non-Core Teachers. The non-core teachers are those who do not teach a core academic subject. These teachers' scores are based on their schools' composite scores on the following tests: (a) non-tested courses, (b) the Stanford Achievement Test in reading for grades kindergarten through second, (c) the FSA in ELA for grades 3 through 10, and (d) one of the following: the College Board Scholastic Aptitude Test, the American College Test, or the Post-Secondary Readiness Test for grades 11 and 12.

Outcomes

Outcomes are the expectations the Florida Department of Education expects teachers to achieve. The outcomes are also the criteria by which school districts determine teacher effect. The Florida Department of Education contractor reports teacher effect as either a positive or negative value-added score (Assessment, Research, and Data Analysis Division, 2014). Within this section, the researcher describes the three VAM models: the Florida VAM, the district covariance adjustment VAM, and the district achievement VAM.

Outcomes of the Florida VAM. The Florida Department of Education provides the three-year aggregated results from the Florida VAM in both ELA and mathematics.

Determining the three-year aggregated results for each teacher is a multi-step process wherein the Florida Department of Education's contractor conducts several calculations. The first step of the process involves determining the difference between the average performance of a teacher's students and the expected performance of academically and demographically comparable students across the state of Florida. The expected performance is a measurement of each student's prior achievement while adjusting for certain demographic, academic, and classroom characteristics (Florida Department of Education, 2013). The second step involves dividing by the difference in the mean scale scores between assessment results in two consecutive academic years for a specific grade level and subject area (average amount of academic growth).

This result can then be interpreted as a percentage of the difference between the average student performance and the expected performance makes of the average annual amount of academic growth. Finally, these results were aggregated across grade levels, subject areas and academic years. If the teacher receives a positive value outcome, then the average performance of his or her students exceeded the expected performance of academically and demographically comparable students across the state; conversely, if the receives a negative value outcome, then the average performance of his or her students was below the expected performance of academically and demographically comparable students across the state.

Outcomes of the district covariance adjustment VAM. Similar to the Florida VAM, the outcome for the district covariance adjustment VAM is the difference between the average performance of a teacher's students and the expected performance of academically and demographically similar students for only the targeted district. Like the

Florida VAM, the expected performance is a measurement of each student's prior achievement while adjusting for certain demographic, academic, and classroom characteristic. A positive value indicates the average performance of a teacher's students exceeds the expected performance of academically and demographically similar students across the school district whereas a negative value indicates the average performance of a teacher's students is below the expected performance of academically and demographically similar students. In contrast to the Florida VAM, outcomes of the district covariance adjustment VAM are reported as a scaled score (except for AP outcomes where the numbers represent the difference between the percentage of students who pass an AP exam [with scores of three to five] and the expected percentage) and the formula includes only student assessment data from the previous school year.

Outcomes of the district achievement VAM. The outcomes of the district achievement VAM differ from the previous two models discussed as outcomes are the difference between the passing rate of a teacher's students on all assessments combined and the passing rate of the target school district's students for a particular subject area such as mathematics or social science. The school district reports the percentages as decimals, with positive values indicating by how many percentage points the passing rate of a teacher's students exceeds the school district's average passing rate on a subject-area assessment and negative values indicating by how many percentage points the passing rate of a teacher's students are below the school district's average passing rate on a subject-area assessment. Similar to the district covariance adjustment VAM, the outcomes of the district achievement VAM are based on the student assessment data from the previous school year (Assessment, Research, and Data Analysis Division, 2014).

Benefits of VAMs. Researchers identified several benefits of VAMs across multiple school settings (Collins & Amrein-Beardsley, 2014; Darling-Hammond, 2015; Loeb, 2013; McCullough et al., 2015). McCullough et al. (2015) noted the dearth of empirical research investigating the impact of VAMs on multiple teacher outcomes. Hoping to contribute to the overall body of knowledge, McCullough et al. sought to determine the impact of VAMs on multiple outcomes among teachers who worked in eight Mid-Atlantic school districts. They conducted "interviews with district administrators, principals, teachers, and teachers' union representatives in the eight school districts" (p. 3). McCullough et al. found teachers and school- and district-level administrators perceived the VAMs were reliable indicators for assessing teacher effectiveness and appropriate mechanisms for determining teachers' eligibility for performance-based bonuses.

Loeb (2013) believed schools could use VAMs to assess the impact of teacher-based training programs on student achievement across both content areas and grade levels. Loeb also concluded that VAMs were effective measures for identifying ineffective teachers and those who required specialized professional development. Loeb further noted that schools could use VAMs as the criteria by which they either promote or fire teachers. Collins and Amrein-Beardsley (2014) concluded that VAMs provide a comprehensive perspective of teachers' effect on students' achievement over a period of years. The VAM is also a mechanism for helping teachers and schools identify what content areas require additional time and how educators can improve their pedagogical practices and promote student achievement (Darling-Hammond, 2015; Loeb, 2013).

Although VAMs are not perfect, they are more useful than other methods of

teacher evaluation because they incorporate multiple years of data and account for measurement error, which is referred to as the standard error (Collins & Amrein-Beardsley, 2014; Jensen, 2011). In essence, the VAM can account for where students are at the beginning of the school year and track their growth throughout the school year and thereby recognize the extent to which educators promote student achievement and schools meet the needs of their students. When schools use VAMs appropriately, they can foster collaboration among educators (Darling-Hammond, 2015).

Standard error. Standard error is the measure of uncertainty in the outcome as a result of factors that are outside of the teachers' scope of control (Doran, 2014). Although teachers in the Miami-Dade County School District may be instructing demographically and academically similar students in their classrooms, students' achievement on standardized tests can vary based on the test form and even the day of the week. The standard error is analogous to the margin of error, which is used when reporting the results of polls or elections. Doran (2014) suggested incorporating standard or measurement error into the point estimates when formulating the characteristics of teacher effect. Point estimates, such as medians or means of growth percentiles, indicate the impact on student achievement. Such point estimates are often the basis for ranking teachers when using a classification structure (Doran, 2014).

Issues pertaining to VAMs. Researchers described several issues regarding the use of VAMs to determine teacher effectiveness (Amrein-Beardsley, Pivovarova, & Geiger, 2016; Rouse et al., 2013; McCullough et al., 2015). Amrein-Beardsley et al. (2016) found teachers and administrators lacked a clear understanding of the intricacies of the VAM and were unable to use their value-added measures as a means to improve

their pedagogical practices and their students' achievement. Another issue Amrein-Beardsley et al. identified was the lack of empirical evidence supporting a relationship between any of the indicators associated with the VAM and "at least one concurrent measure of teacher effectiveness, such as supervisors observational assessment of teachers or students survey-based assessments" (p. 36). Rouse et al. (2013) made a similar conclusion regarding the lack of a discernible relationship between the VAM indicators and the measures schools used to measure teacher effect and therefore schools were not able to evaluate teachers accurately.

McCullough et al. (2015) identified several issues hindering the implementation of VAMs. These challenges included the cost-prohibitive nature of implementing the measures across school districts, the lack of evidence to support the reliability and the validity of the models, and the inability to ensure adherence to the evaluation model. Other challenges related to the implementation of VAM include the inability to account for differences in student demographic variables, the lack of support from teachers' unions, the inability to get a precise measure of a teachers' contributions to students' learning (Amrein-Beardsley et al., 2016; Rouse et al., 2013).

Other Issues With VAMs

Researchers identified several constructs that school districts must consider when implementing merit pay programs for teachers (Alger, 2014; Dean, 2015; Fulbeck, 2014; Yuan et al., 2013). Ritter and Jensen's (2010) four foundational constructs of attainability, transparency, substantiality, and sustainability is the focus of this section. In the subsequent section, the researcher describes the constructs and their application in an educational setting.

Attainability. The construct of attainability refers to teachers' beliefs that predetermined student benchmarks are achievable (Edenfield, 2014). For merit pay programs to motivate teachers and improve student outcomes, teachers must believe the reward is attainable (Edenfield, 2014). The degree to which individuals perceive the attainability of goals influences their levels of motivation. If individuals believe goals are attainable, they are more likely to exhibit high levels of motivation and to persist when challenges occur; conversely, if individuals believe goals are unattainable, they are less likely to exhibit high levels of motivation and more likely give up when challenges occur (Alger, 2014; Yuan et al., 2013).

When considering the construct of attainability, one of the hurdles many merit-based programs face is the likelihood of a group of students reaching predetermined benchmarks is low (Yuan et al., 2013). Teachers are acutely of the influence factors such as socioeconomic status and mobility have on student achievement. Teachers with either a disproportionately high percentage of students from low socioeconomic backgrounds or a high student mobility rate are less likely to believe their students will meet performance benchmarks (Marsh, 2014; Yuan et al., 2013); consequently, they are unlikely to put forth the requisite effort to achieve their performance benchmarks. Researchers note that recent attempts to account socioeconomic status, student mobility, and other factors beyond the scope of a teachers' control have done little to promote teachers' belief in their students' ability to meet achievement benchmarks and their willingness to expend the necessary energy to improve student achievement (Alger, 2014; Marsh, 2014).

Transparency. The construct of transparency refers to teachers' understanding of the formulas and the data school districts use to determine their VAM (Dean, 2015).

Providing a clear understanding of expectations and the criteria by which schools measure a teacher's effectiveness is paramount (Anderson, Hunt, Powell, & Dollar, 2013; Dean, 2015). Dean (2015) defined transparency "as the degree to which a VAM is understandable, replicable, or usable by researchers, educators, or the public" (p. 1). For teachers to be motivated, they must have a precise understanding of teacher performance expectations. These expectations must be a manner that educators believe the merit plan is transparent, comprehensible, and systematically efficient. Researchers found involving a consensus of teachers at the inception of a performance pay initiative was an effective means for promoting personal investment and the long-term sustainability of the initiative (Anderson et al., 2013).

Substantiality. The third foundational construct of Ritter and Jensen's 2010 is substantiality. Researchers stressed the importance of ensuring performance pay rewards were substantial enough to assure educators that school districts will compensate them for their talents and efforts in both the present and the future (Anderson et al., 2013; Edenfield, 2014). Financial rewards should be commensurate with what educators in other states receive. If teachers are to move out of their comfort zone and try innovative, research-based strategies, then the reward should outweigh the risks (Anderson et al., 2013).

Sustainability. The fourth and final foundational point is the construct of sustainability. If merit pay is going to be temporary, then teachers are unlikely to be motivated to adapt their instructional practices to improve student achievement. Researchers noted the sustainability of merit pay programs may be difficult in times of tax revenue shortfalls or budgeting issues (Fulbeck, 2014; Maranto, 2014). To ensure the

sustainability of their merit-pay programs, Colorado and Arizona earmarked funds for teacher incentive pay programs (Fulbeck, 2014).

Teacher Accountability

Throughout the course of history, the construct of teacher quality and effectiveness in America's schools has evolved in conjunction with societal norms. Similarly, models measuring teacher quality have changed to conform to paradigm shifts within society (Gamson & Hodge, 2016). Researchers found the construct of teacher evaluation changed in accordance with society's perceptions regarding the roles and responsibilities of educators, the manner in which students acquire knowledge, and the changing demographics of America's schools. One of the most recent models that schools systems across the United States have used is VAMs (McCullough et al., 2015). Researchers traced the construct of teacher evaluation back to the one-room schoolhouses wherein a hierarchical system was prevalent and compliance to procedures superseded professional growth (Farley, 2017). Prior to 1850, teacher-evaluation practices were a series of intermittent observations wherein evaluators ensured teachers conformed to the community's standards and terminated educators whom they deemed were ineffective (Gamson & Hodge, 2016). The evaluations of educators were more closely related to religious beliefs and societal norms within the community than to actual educational reforms (Farley, 2017; Gamson & Hodge, 2016).

Through the latter of the 19th century and the first half of the 20th century, the teaching profession evolved as did the idea of teacher accountability (Gamson & Hodge, 2016). It was during this period that male supervisors observed and evaluated teachers (Farley, 2017). These male supervisors observed and evaluated teachers because

policymakers held them in high esteem for their expertise regarding pedagogical practice and student learning (Gamson & Hodge, 2016). Furthermore, as schools began to grow and their curriculum expanded and included various core academic areas, qualified educators were in demand, thereby increasing the need for a reliable teacher evaluation model and increased accountability for students (Farley, 2017; Gamson & Hodge, 2016).

In the middle of the 20th century, schools adopted a checklist-style evaluation and school administrators and educators engaged in verbal dialogue concerning the evaluation process (Gamson & Hodge, 2016). This paradigm shift prompted researchers to investigate the relationship between educator behaviors and student outcomes (Farley, 2017). The dependence on local and state policy determined when and how often school administrators used these checklists to assess teachers; however, teacher evaluation became a fundamental component of both school district policy and collective bargaining agreements (Farley, 2017).

As the practice of using checklists increased, the construct of teacher accountability emerged and evolved. In the early 1970s, policymakers and educational stakeholders acrimoniously debated teacher evaluation policies and accountability (Yuan et al., 2013). Eventually, accountability within K-12 education encompassed teacher effectiveness, teacher practices, and school-wide success and standardized test scores became the primary criteria schools used to evaluate teachers. By the end of the 20th century, the link between teacher evaluation results and support services ushered in a new era of educator accountability wherein school district leaders felt pressure to use data from assessments in the teacher evaluation process and to enact professional development requirements for teachers (Yuan et al., 2013). This pressure compelled school districts to

meet state and federal mandates while safeguarding the belief that individual educator performance measures were reliable and valid (Kane & Staiger, 2012).

Merit Pay

Generally restricted to the private sector throughout the majority of the 20th century, merit-based pay initiatives emerged in federal employment agencies with the passage of the Civil Service Reform Act of 1978. By establishing merit-pay guidelines, legislators believed they had a mechanism for restoring responsiveness and efficiency within the federal sector (Nieberg, Pieper, & Trevor, 2016). The most noteworthy components of the Civil Service Reform Act of 1978 were comprehensive appraisal reforms, large cash awards for employees, merit pay and cash awards for specified managers, and the establishment of performance incentives for senior executive service employees (Nieberg et al., 2016).

Nieberg et al. (2016) defined merit pay as additional monetary compensation for employees who meet previous performance benchmarks. They found merit pay existed across various industries and under different compensation structures. Within the field of education, teachers receive merit pay for meeting predetermined performance criteria (Stephens, 2015). The ability to improve students' scores on standardized tests is one of the most prevalent criteria that must meet to receive merit pay (Gius, 2013; Stephens, 2015). Other criteria for earning merit pay include degree attainment and years of experience (Edenfield, 2014; Gius, 2013).

Edenfield (2014) investigated the perceptions of one southeastern state's teachers regarding merit pay initiative and found that more than 80% of the teachers opposed merit pay based on student achievement and preferred merit-based salary structure

wherein teachers receive additional compensation for degree attainment and for participation in professional growth activities. When Edenfield asked teachers to describe their perceptions regarding merit-pay initiatives on the basis of student achievement, teachers cited the inability of evaluation models to consider factors beyond their scope of influence (socioeconomic status, the extent of parental involvement in schools), the disproportionate emphasis on students' test scores in comparison to other components of the evaluation process, and the potential for a negative work environment.

Stephens (2015) also investigated teachers' perceptions of merit pay but in a different southeastern state. Stephens found teachers' overall perceptions of merit pay were indifferent. Stephens surmised the indifference among teachers regarding merit pay could have been the result of them perceiving that merit-pay initiatives aimed at improving instructional practices and thereby student achievement were ineffective. The other rationale Stephens provided for the indifference among teachers was the merit-pay program's decision to include additional benchmarks during the second year of implementation. Other data analysis conducted by Stephens included comparing teachers' perceptions of merit pay based on factors such as years of teaching experience (teachers with five or more years or experience vs. those with less than five years' experience), a school's socioeconomic status (low socioeconomic status vs. high socioeconomic status), and type of course (course with standardized test vs. course without a standardized test). Stephens found teachers' perceptions did not differ across the factors of years of teaching experience, socioeconomic status of school, and type of course. To improve teachers' perceptions of merit pay programs, Stephens recommended giving teacher input regarding the criteria for earning merit pay and closely aligning the criteria for merit pay

to the "the school district mission statement" (p. 93) as doing this would promote autonomy among stakeholders and clear rationale for improving student achievement.

Similar to Stephens (2015) and Edenfield (2014), Russ (2015) sought to determine teachers' perceptions of merit pay; however, Russ investigated the perceptions of over 250 teachers from two school districts within one Midwestern state. Russ used a quantitative approach and a descriptive design. Teachers completed a Likert-scale survey wherein they responded to a series of statements pertaining to various elements of teacher-performance models and the extent to which elements such as teacher evaluations, standardized test scores, and school- and district-level performance on assessments should be part of the criteria for determining merit pay for teachers. Russ also determined teacher overall perceptions of merit-pay initiatives. Teachers had negative perceptions regarding the use of standardized test scores and student performance on school- and district-level performance on assessments as the criteria for determining merit pay for teachers. Overall, teachers had negative perceptions of merit-pay initiatives as they perceived those initiatives were not effective mechanisms for rewarding teacher performance and would create a negative work environment.

In a study conducted in another Midwestern state, Routh (2014) investigated teachers' perceptions of merit-pay initiatives; however, in contrast to Russ's 2015 study, Routh's 2014 study was broader in scope as Routh included teachers from across the entire state. Routh found more than two thirds of the Midwestern state's teacher had negative perceptions regarding the use of administrators' evaluations as the basis for determining merit pay for teachers. Routh also concluded approximately three fourths of the teachers perceived that district-level leaders would foster a negative work

environment by implementing merit-pay initiatives.

Arizona. While various states have changed and modified the criteria by which they determine merit pay for teachers, Arizona is one of the only states to have permanent, uniform criteria for awarding merit pay. In a 2010 report, the Arizona Auditor General found 29 districts that received funding for merit pay for teachers were able to associate teacher performance pay to student learning gains (Buck & Greene, 2011).

At the beginning of the 2014-2015 school year, the state of Arizona passed statute ARS 15-977 and earmarked funds for teachers who met measurable achievement outcomes (Tucson Unified School District, 2017). ARS 15-977 has the following four elements: (a) the adoption of a performance-based compensation system by the governing school board during a public hearing, (b) the implementation of a teacher evaluation component as determined by the teacher's performance classification, (c) the use of precise instruments to measure students' progress toward meeting academic standards set forth by the Arizona State Board of Education, and (d) the approval of the performance-based compensation system by at least 70% of a school district's teachers who are eligible to participate in the system (Tucson Unified School District, 2017). Further analysis of the teacher evaluation component reveals that 40% percent of the money is earmarked for individual teacher performance and a teacher's performance classification accounts for 33% (Tucson Unified School District, 2017). Teachers are also able to earn performance-based compensation for earning certification from the National Board for Professional Teaching Standards, planning site-based professional development sessions, and meeting the requirements of the site-based plan, which consists of two components (Tucson Unified School District, 2017).

Colorado. Since 1993, the state of Colorado has been an innovator in performance pay systems. Hoping to reform its teacher compensation model, the state of Colorado rolled out a new plan at the beginning of the 2010-11 school year (Blazer, 2011). The new plan was part of Senate Bill 191 (Robles, 2015). Under this plan, schools assessed teachers annually using a comprehensive evaluation model that included the following components: (a) students' standardized test scores, (b) administrator observations, (c) measurements assessing teachers' aptitude to develop 21st century skills, (d) parent and student evaluations, and (e) the ability to meet yearly program goals (Blazer, 2011). Students' performance on standardized tests accounted for exactly 50% of a teacher's evaluation, and teachers who received positive evaluations were eligible for financial incentives; conversely, teachers who did not receive positive evaluations were not eligible for either financial incentives or yearly salary increases (Robles, 2015). Teacher unions across the state of Colorado vehemently opposed the plan, and since the 2015-2016 school year, teacher evaluations in Colorado have not included test scores and salary increases are based on meeting school performance goals or receiving a satisfactory personal evaluation (Robles, 2015).

Munroe (2017) conducted a study in one large urban Colorado school district wherein she compared teachers' perceptions of merit-based pay. Munroe used an experimental design and randomly assigned each participant to one of two groups. Each group of participants received a questionnaire that included a distinct speculative situation regarding merit pay, a series of Likert-scale items, and three short-answer questions. Munroe created Likert-scale items and addressed variables such as pedagogical practices, school climate, motivation, and commitment to the teaching

profession and open-ended questions regarding teachers' perceptions of merit-based compensation. Munroe assessed the first group of teachers' perceptions of a merit-based pay system wherein alternative measures of assessment were the criteria for determining student growth and the second group of teachers' perceptions of a merit-based pay system wherein standardized test scores were the criteria for determining student growth. Next, Munroe compared the two groups' perceptions across the following four variables: pedagogical practices, school climate, motivation, and commitment to the teaching profession.

Munroe (2017) found teachers perceived higher levels of motivation and commitment to the teaching profession under a merit-based pay system that used alternative measures to determine student growth than they did under a merit-based pay system that used standardized test scores to determine student growth. Munroe also analyzed participants' responses to the open-ended questions and found the recurrent theme of "fairness of implementation" emerging from both groups' data. Teachers perceived that schools promoted a fair merit-based pay system when they had uniform procedures regarding who assessed teacher performance and accounted for factors beyond teachers' locus of control.

Bruce Messinger, who is the Superintendent of the Boulder Public Schools, questioned the effectiveness of reforms aimed at monetarily reward teachers for improving student achievement performance (Robles, 2015). Messinger believed these reforms are simple solutions to a multi-faceted problem and making a connection between students' growth on standardized tests and teacher effectiveness is problematic as factors beyond the scope of a teacher's control impact students' achievement (Robles,

2015).

Although the new teacher evaluation plan is on hold across most of the school systems across the state of Colorado, several school systems have taken the initiative and implemented alternative teacher compensation models. Fulbeck (2014) described the Denver Public School's merit pay program "as one of the most prominent alternative teacher compensation reforms in the nation" (p. 67). Fulbeck found the ProComp program included various financial rewards for teachers who increased their overall pedagogical effectiveness as measured by their students' test scores. Plagued by high attrition rates among its teachers, the Denver Public Schools implemented the ProComp program in an effort to recruit and to retain effective teachers. To gain a clear understanding of ProComp's impact in the Denver Public schools, Fulbeck compared the attrition rates of teachers who received a financial incentive through ProComp and the attrition rates of teachers who did not. Fulbeck used a longitudinal design and collected attrition data over a nine-year period. Fulbeck found that teachers who received the ProComp financial incentive were less likely to leave the Denver Public Schools than teachers who did not receive the financial incentive.

North Carolina. The state of North Carolina has several districts involved in teacher compensation programs, with some being successful and others being both controversial and unsuccessful. These programs are continuous programs or programs set with a time-limit through state or national funding (Blazer, 2011). North Carolina's ABC program is one that was meant to reorganize the district around specific goals. These goals were based on a model whereas schools are evaluated based on student performance on standardized testing. The ABC's Accountability Model assigned one of

the following performance ratings to each school: (a) school of excellence, (b) school of distinction, (c) school of progress, or (d) low performing school (Blazer, 2011). Although the ABC program was discontinued in the 2008-2009 school year due to budget restraints, a new budget was approved during the 2012-2013 school year for an alternative teacher compensation plan with the most recent teacher-merit pay program in effect during the 2014-2015 school year (Lauen & Kozlowski, 2014).

Additionally, Guilford County Schools, which is the third most populous county in North Carolina, has initiated a performance-pay program over the past nine years that has been funded by two separate sources. The Mission Possible (MP) was a federally funded grant meant to recruit, retain, and reward highly qualified teachers through bonuses. While the $8 million federal grant has since expired, the MP is currently funded by a $22.8 million U.S. Department of Education Teacher Incentive Fund and due to expire in 2015 but deemed successful (Maranto, 2014).

Unlike Guilford County, the Charlotte-Mecklenburg School (CMS) district's performance pay plan was met with controversy. The state funded this plan through a 5-year grant and instituted the plan in the district's 20 highest-need schools, with teachers receiving additional pay for employing hard-to-staff schools (Blazer, 2011). The controversy ensued when public perceptions met with discourse on converting the entire state of North Carolina into teacher performance pay zones homogeneously and legislation built in to exclude teacher and district approval of the measures. The public's perception was that CMS added two public relations firms to help equalize the response as well as a donation given from the Gates Foundation of almost $250,000 to support the county's efforts on improving the performance-pay endeavor (Maranto, 2014).

Tennessee. The state of Tennessee instituted the VAM and called it the Tennessee Value-Added Assessment System. In 1995, Chattanooga became the first school district to institute an incentive-based pay system in the form of bonus pay for teachers who worked in its chronically low-performing schools or taught high-needs subject areas (Davis, 2014). Eventually, the state of Tennessee went away from using merit-based pay as an incentive for improving student achievement and adopted a rigorous teacher-evaluation model that used student achievement data as the primary criteria for determining teacher effectiveness (Davis, 2014; Toch, 2016). The state of Tennessee made this decision after it received a $500,000,000 as part of the U.S. Department of Education's Race to the Top Initiative (U.S. Department of Education, 2016). Disillusioned by the teacher-evaluation model's overemphasis on student achievement, teachers across the state of Tennessee voiced their displeasure to local and state politicians (Davis, 2014). The most poignant aspects of the teacher-evaluation model were the number of times administrators observed teachers each school year and the weighted percentage of student growth measures in teachers' yearly evaluations (Davis, 2014; Will, 2018).

Recently, the state of Tennessee adopted a new program wherein local teachers have more autonomy and are able to receive additional compensation for serving as instructional and data coaches and curriculum developers (Will, 2018). Tennessee is using Title II monies to fund the program (Will, 2018). Under the program, school districts must incorporate differentiated pay scales into their salary schedules, provide money awards for teachers who receive positive teacher evaluations, and give additional compensation to teachers who teach critical needs subject areas or work in high-needs

schools. Because this is a new program, there is no empirical evidence describing the effect of Tennessee's newer model's on students' test scores.

Texas. In 2008, the state of Texas used a voluntary merit pay program for teachers, and approximately close to one-fifth of Texas's public-school districts participated in the program (Blazer, 2011). The state of Texas funded the program through initiatives such as the Texas Educator Excellence Grant (TEEG) and the Governor's Educator Excellence Grant (GEEG) and teachers from almost 1000 high-poverty schools that the state designated as high-achieving schools participated in the program (Blazer, 2011; Stutz, 2013). In 2011, with funding cuts of over $360 million across the state of Texas, the voluntary merit pay program was on the brink of extinction (Blazer, 2011). Hoping to fund the program fully when revenue increased, Texas maintained a remnant of the voluntary merit pay program through the 2013-2014 school year, with a budget that was approximately 6% of the original budget.

By the conclusion of the 2013-14 school year, the state of Texas to end the voluntary merit pay program for teachers to replace it with the Educator Excellence Innovation Program (Stutz, 2013). Texas implemented the Educator Excellence Innovation Program at the beginning of the 2014-15 school year (Smith, 2018). The central aim of the program is to award grant money to schools with a disproportionately high percentage of students from low socioeconomic backgrounds and to prepare new teachers to be successful in the classroom and to provide veteran teachers with new career pathways (Stutz, 2013). During the 2014-15 school year, 40 schools participated in the Educator Excellence Innovation Program. As the Educator Excellence Innovation Program completes its fourth year, Smith (2018) noted the program is struggling to

implement a comprehensive teacher evaluation model in the midst of statewide budgetary constraints. The superintendent of the Dallas Public Schools, Michael Hinojosa, underscored the importance of implementing an innovative evaluation system and identifying additional resources to fund the program (Smith, 2018). In 2011, the state of Texas implemented another merit pay program for teachers titled the District Awards for Teacher Excellence (Stutz, 2013). Texas used state revenue to fund the District Awards for Teacher Excellence program, and teachers received merit pay for improving students' performance on standardized tests (Stutz, 2013).

Stutz (2013) investigated the impact that merit pay programs had on students' academic achievement. Stutz compared the achievement gains of students who attended Texas public schools with merit pay programs and of students who attended Texas public schools without merit pay. Stutz found that students who attended schools with merit pay had greater gains in standardized scores than those who attended schools without merit pay (Stutz, 2013).

Florida. In 2010, Florida was one of 12 states that received federal funding through a Race to the Top grant (U.S. Department of Education, 2012). The state of Florida used the funds from three-year grant to support school districts as they went through the modified and revised the teacher evaluation process. Florida earmarked almost $350 million of the $700 million in federal funding allocated under the Race to the Top grant to the implementation of the first three years of the pay for performance system, which span from 2011 to 2014 (U.S. Department of Education, 2012). In addition to the Race to the Top Grant, Florida's governor signed into law the SSA. The SSA mandates that school districts across the state of Florida link teacher pay to students'

academic performance (Haertel, 2013). Recently, President Trump backed mandates such as the SSA, stressing that it was time for merit pay for teachers (Strauss, 2017).

Under the SSA, school districts across the state of Florida must base at least 50% of teachers' performance evaluation on students' learning growth as determined by the FSA (Florida Department of Education, 2014). According to the Florida Department of Education (2014), school districts use a VAM as the formula to calculate student learning growth. Districts are able to choose one of three pre-approved frameworks for their teacher evaluation models. After choosing their preferred evaluation model, the department of education reviews and approves the evaluation system and monitors the implementation process so that school districts are in compliance with the law (Florida Department of Education, 2014). As the beginning of 2016-2017 school year, all Florida districts implemented a revised teacher evaluation based on the pre-approved models.

According to the latest census figures, Florida has 190,000 teachers working in over 4,200 public schools and the fourth largest student population in the United States with more than 2.6 million public school students (Florida Department of Education, 2013). When the state of Florida passes new education legislation, teachers across the state are impacted. The recent passage of the SSA had a profound effect on new teachers across the state of Florida. The effect on new teachers was so profound that the Florida Education Association (FEA) filed a lawsuit against the state challenging the constitutionality of the new law and the new teacher evaluation system (O'Connor, 2013). The 2013 case of *Cook et al, v. Pam Stewart, Florida Commissioner of Education, et al.,* called into question the equal protection rights of teachers whose evaluations were based on the performance of students whom they did not instruct and in subjects that they

did not teach (O'Connor, 2013). In 2014, the district courts ruled that there was a rational basis for Florida public schools to adopt policies requiring schools to base teacher evaluations on students' test scores, even when they did not teach either those students or those subject areas (Bauries, Sutherland, & Legare, 2014).

One of the most noteworthy controversies regarding teacher performance pay in Florida is the overarching purpose of the SSA: to increase student achievement (Postal, 2017). However, six years after the state signed the Students Success Act into law and the state's school districts implemented a teacher merit-pay model, there is limited evidence that student achievement has increased (Postal, 2017). Since the Florida Legislature signed the SSA into law, there has been no consistent improvement in student achievement (Carruthers, Figlio, & Sass, 2018).

Pay for performance for educators does not come without its challenges to public school districts throughout the country. According to Springer and Winters (2009), pay for performance had no impact on student achievement. This conclusion was based on a study conducted using a very large sample of elementary schools in New York City for both the experimental and control groups. Ritter and Jensen (2010) proposed specific foundational points can also pose as challenges in terms of educational performance pay designs and implementation to public school districts. These challenges include factors such as (a) attainability, (b) transparency, (c) substantiality and (d) sustainability.

Summary

Vroom's (1964) expectancy theory was the theoretical lens through which the researcher investigated the construct of performance pay for teachers. Vroom's theory is appropriate because of its inextricable link to behaviorism. Vroom's expectancy theory is

composed of three fundamental variables: expectancy, instrumentality, and valence. Though these variables are distinct, they work synergistically to influence individuals' motivation and performance (Gemeda, 2015). Expectancy is the strength of the relationship between the level of effort teachers put forth to achieve student outcomes and the degree to which they believe their efforts will increase student achievement (Yuan et al., 2013). Instrumentality is a measure of individuals' belief that their performance will produce their desired outcomes (Rice et al., 2015). Valence is the value teachers affix to a desired goal (Najera, 2017).

After the passage of NCLB and the Race to the Top Initiative, the USDOE encouraged states to implement comprehensive teacher evaluation models for teachers wherein teachers' effect on students' academic achievement was a significant factor in determining teacher effectiveness (Lavania et al., 2015). The VAM is a mechanism for determining teachers' effect on students' test scores and their overall academic achievement (McCullough et al., 2015). The Florida Department of Education provides the 3-year aggregated results from the Florida VAM in both ELA and mathematics. Researchers identified challenges related to the implementation of VAM including the inability to account for differences in student demographic variables, the lack of support from teachers' unions, the inability to get a precise measure of a teachers' contributions to students' learning (Amrein-Beardsley et al., 2016; Rouse et al., 2013). Researchers recommended that school districts consider the constructs of attainability, transparency, substantiality, and sustainability prior to implementing a VAM (Alger, 2014; Marsh, 2014; Yuan et al., 2013). A review of the related literature revealed that models measuring teacher quality changed to conform to paradigm shifts within society (Gamson

& Hodge, 2016). Current models assessing teacher quality use student outcomes on standardized testing as an integral component in determining teacher effectiveness (Yuan et al., 2013).

Hoping to motivate teachers to improve their instructional practices and promote student achievement, many states implemented teacher performance-pay initiatives (Fulbeck, 2014; Maranto, 2014; Munroe, 2017; Robles, 2015). Researchers found that many of these initiatives failed because of inadequate funding and ambiguous criteria for assessing teacher effectiveness (Fulbeck, 2014; Robles, 2015). Munroe (2017) found teachers preferred a merit-based pay system that used alternative measures to determine student growth than a merit-based pay system that used students' scores on standardized tests to determine student growth. School districts across the state of Florida base at least 50% of teachers' performance evaluation on students' growth as determined by the FSA (Florida Department of Education, 2014). Since signing the SSA into law and implementing a teacher performance-pay initiative, the state of Florida has not investigated teachers' perceptions of the initiative. Carruthers et al. (2018) noted that there has been no consistent improvement in student achievement since the Florida Legislature signed the SSA into law.

Research Questions

The researcher used a descriptive survey research design and a convenience sample of teachers who worked at a Title 1 elementary school in the southeastern United States The following research questions have been established to guide this proposed applied dissertation:

1. What are urban elementary teachers' perceptions of pay-for-performance initiatives?
2. How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to grade level?
3. How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to tenure and non-tenure status?
4. How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to their years of experience?

Chapter 3: Methodology

The purpose of this quantitative study was to investigate Title 1 elementary teachers' perceptions of a teacher performance pay initiative. Despite implementing teacher performance pay initiative during the 2015, the target school district had not investigated teachers' perceptions of the initiative (Francilus, 2015). Chapter 3 includes a description of the participants and the target setting, the instrument, the research design, and the related procedures for conducting the study and analyzing the data. The final section of chapter 3 includes a description of the limitations of this study.

Participants

The participants for this study were elementary school teachers who worked at a Title 1 elementary school that is part of a large urban school district. The target school employs approximately 60 professional staff, which includes classroom teachers, special education teachers, specialty area teachers, a school psychologist, and a guidance counselor. The setting for this applied dissertation study was an elementary school within an urban public-school district in Florida.

The following were the target elementary school's enrollment figures by race/ethnicity for the past five school years: (a) 2011-12 school year = Hispanic = 528 (57.6%), African American = 224 (24.5%), White = 122 (13.4%), and Other = 42 (4.5%); (b) 2012-13 school year = Hispanic = 550 (60.4%), African American = 208 (22.8%), White = 111 (12.2%), and Other = 33 (4.6%); (c) 2013-2014 school year = Hispanic 597 (62.9%) , African American = 186 (19.6%), White = 119 (12.5%), and Other = 39 (5%);

(d) 2014-2015 school year = Hispanic = 553 (62.2%), African American = 181 (20.4%), White = 117 (13.2%), and Other = 26 (4.2%); and (e) 2015-2016 school year = Hispanic =622 (68,7%), African American = 137 (15.1%), White = 109 (12%), and Other = 38 (4.2%). From 2011 to 2016, the Hispanic population had the largest population increase (approximately 14%) while the African American population had the largest population decrease (approximately 10%). The construction of new schools within the target school district, which began prior to the start of the 2011-12 school year, prompted school leaders to change attendance boundaries to enhance racial diversity.

Sample. The researcher used a convenience sampling procedure to select this study's participants. The sample for this study included 54 teachers who worked at a Title 1 elementary school in the southeastern United States. Convenience sampling is a non-probability sampling technique that is appropriate when researchers include participants who are easily accessible (Creswell, 2013; Suen, Huang, & Lee, 2014). One of the distinguishing characteristics of a non-probability sampling technique is that not all members of the target population have an equal chance of being part of a study. The inability to ensure that all members of the target population have an equal chance of being part of a study is referred to as sampling bias (Gurnsey, 2017). Sampling bias impacts researchers' ability to generalize findings to the broader population (Creswell, 2013; Suen et al., 2014).

Instruments

The researcher used a 15-question Likert scale survey to investigate elementary teachers' perceptions of a teacher performance pay initiative. Pemberton-Albright (2011) created the survey and used it in a study titled *The Merit of Merit Pay*. Pemberton-

Albright assessed teachers' and educational stakeholders' perceptions of merit-based pay. Pemberton-Albright's original survey includes the following types of question: (a) seven short-answer items pertaining to participant demographics, (b) 11 five-point Likert scale items, and (c) seven items requiring participants to either circle their responses or provide a short-answer response. Pemberton Albright conducted field testing to establish the reliability of the instrument. The field test included 20 elementary and middle school teachers from one Midwest state. The 20 teachers provided feedback regarding the wording of the questions and recommendations for enhancing the clarity of both the Likert-scale items and the short-answer responses. Pemberton Albright's dissertation committee also provided feedback regarding the wording of both the Likert-scale items and the short-answer responses. Forand (2012) used a modified version of Pemberton-Albright's survey and determined the Cronbach's alpha coefficient for the survey's Likert scale items. Forand calculated a Cronbach's coefficient of .60 for the Likert-scale items.

Although Pemberton-Albright's original survey included 25-questions, the researcher modified the original survey and included only 15 Likert-scale responses. The researcher also included four demographic items at the very beginning of the survey. The four demographic items addressed the following: (a) gender, (b) grade-level or specialty area, (c) overall number of years in teaching, and (d) tenure/non-tenure status. The researcher collected these demographic data to compare participants' perceptions across grade levels, years of experience, and tenure/non-tenure status.

Procedures

Design. The researcher used a descriptive survey research design. According to Creswell (2013), this type of design is appropriate when researchers administer a survey to

a group of participants and aim to determine and to describe the participants' perceptions, attitudes, or beliefs regarding a phenomenon. One of the distinguishing characteristics of the descriptive survey research design is the use of a pre-existing group to measure or to compare the perceptions and beliefs of its group members. The pre-existing group the researcher included in this study was teachers who worked at a Title 1 elementary school. Researchers using the descriptive survey design do not make causal inferences nor do they determine the strength of the relationships between variables; instead, they analyze measures of central tendency (Ali & Bhaskar, 2016). The descriptive survey research design is prevalent within various social science fields and education, and those conducting descriptive survey research often seek to expand the body of knowledge regarding a topic (Creswell, 2013; Warne, 2017). In regard to this study, the phenomenon under investigation was urban teachers' perceptions of teacher performance pay. The researcher used a survey instrument to determine teachers' perceptions and to compare participants' perceptions across several demographic categories.

Data collection. After completing all of Nova Southeastern University's IRB requirements and receiving permission from the university to conduct this study, the researcher completed all the target school district's related requirements for conducting research in its schools. The researcher submitted all related Nova IRB and school district documents to the target school district's research center. Upon receiving permission from the target school district's research center, the researcher met with the target school's principal and provided him with an overview of the study, the time commitment associated with participating in this study, and the related procedures for ensuring participant confidentiality, contacting potential participants, and collecting participant

data.

After receiving permission to conduct the study from the target school's principal, the researcher asked the target school's secretary to provide the school email addresses of the target school's teachers. The researcher asked the secretary to include each teacher's email address in an Excel spreadsheet. The spreadsheet included only the teacher's email address. The researcher distributed the survey through Survey Monkey and participants completed the survey online. Three days prior to receiving the actual survey, Survey Monkey sent out a pre-notification email describing the survey instrument and stating the amount of time it would take to complete the survey.

Survey Monkey sent out a second email to let participants know that they could complete the actual survey. The email included a link to the survey. Upon clicking on the link, participants first reviewed an informed consent document wherein the researcher stated that participation in this study was voluntary and that no identifying information would be recorded. By completing the survey, participants established their informed consent to participate. The survey was available for two weeks. Participants who did not complete the survey within the first the first week received a second email at the beginning of week two. The content in the second email was identical to the content in the first email. The researcher excluded the data from any participants who did not complete the four demographic items and 15 Likert-scale items. The researcher kept all survey responses in a secure filing cabinet, to which only the researcher had a key. Upon completing this study, the researcher will keep all data for a 3-year period.

Data analysis. The researcher used independent *t* tests to compare teacher' perceptions of teacher performance pay. Kim (2015) identified two type of *t* tests: a

paired *t* test and an independent *t* test. Researchers use a paired *t* test when one group of participants receives the same intervention, and they want to compare the group's mean score before and after it receives the intervention (Creswell, 2013; Kim, 2015). According to Kim (2015), an independent *t* test is appropriate when researchers seek to compare the means of two independent groups who receive the same condition. For this study, the researcher used an independent *t* test. This was an appropriate test because the researcher had two independent groups and compared the two groups' mean scores.

Prior to conducting an independent *t* test, the researcher had to adjust the significance level and to calculate the Cohen's d statistic. The significance level is the probability of obtaining a result by chance alone, and it is reported as a *p*-value (Armstrong, 2014; Hall & Richardson, 2016). The Cohen's d is a measure of effect size, and researchers calculate this measure by determining the mean difference between two groups (Borenstein, Rothstein, & Cohen, 2001). The researcher used the Bonferroni correction to adjust the significance level. In Chapter 4 of the document, the researcher provided a thorough rationale for adjusting the significance level and for calculating the Cohen's d statistic to answer research questions 2 through 4.

The researcher had four research questions. For the first research question, the researcher reported the central measures of tendency (i.e., mean, median, and mode). For the second research question, the researcher used an independent *t* test to compare the mean scores of Pre-K, Kindergarten, and Grade 1 teachers and those of Grade 2, Grade 3, Grade 4, and Grade 5 teachers. For the third research question, the researcher compared the mean scores of teachers with tenure status and the mean scores of teachers with non-tenure status. For the fourth research question, the researcher compared the mean scores

of teachers who had more than 10 years of teaching experience and the mean scores of teachers with fewer than 10 years of teaching experience. Finally, the researcher conducted a primary analysis and a secondary analysis for both Research Question 2 and Research Question 4.

Chapter 4: Results

Introduction

The purpose of this quantitative study was to investigate teachers' perceptions of a teacher performance-pay initiative that was initiated in 2015 at the target school (Francilus, 2015). Research investigating the perceptions of teachers' in school districts across the United States regarding performance pay initiatives is scant (Marsh, 2014; Stephens, 2015; Viscardi, 2014). The target school district had not investigated teachers' perceptions of its teacher performance-pay initiative.

The researcher compared teachers' perceptions of a teacher performance-pay initiative across multiple demographic factors could enable educational stakeholders at both the district- and school-level to develop a clear understanding of the extent to which one elementary school's teachers perceived that performance pay impacted their job performance across the factors of grade level, years of experience, and tenure and non-tenure status. The Title 1 elementary school is part of a large urban school district in the southeastern part of the United States. The sample was composed of elementary school teachers who worked at a Title 1 elementary school during the 2018-19 school year. The sample for this study included 54 teachers who worked at a Title 1 elementary school in the southeastern United States; however, the researcher included 52 teachers in the data analysis. Upon reviewing participants' responses to the survey items, the researcher found 50 participants completed the entire survey and four participants did not complete the entire survey. Two teachers did not complete survey items 8-12 and two teachers did not complete the demographic items and background questions. For the data analysis, the researcher excluded the two participants who did not complete survey items 8-12.

Demographic Characteristics

Of the 52 respondents included in the analysis, 49 (94.2%) were females and three (5.8%) were males. The mean age of respondents was 48.69 (*SD* = 11.88), and the mean number of years of teaching experience among them was 19.31 (*SD* = 10.77). More than half the teachers (29 teachers; 55.8%) had tenure, while the remaining subjects (23 teachers; 44.2%) did not have tenure.

There were five teachers (9.6%) who were in the DROP program and there were 47 (90.4%) of the participants who were not in the DROP program. The researcher included from seven different grade levels ranging from pre-kindergarten, kindergarten, and first through fifth grade. Table 1 shows the number of teachers teaching at each grade level.

Table 1

Frequencies of Teachers at Each Grade Level

Grade Level	*n*	*%*
Pre-K	1	1.9
Kindergarten	11	21.2
Grade 1	11	21.2
Grade 2	15	28.8
Grade 3	20	38.5
Grade 4	18	34.6
Grade 5	17	32.7

Note. The numbers add up to more than 52, and to more than 100%, because some teachers taught at more than one level.

Data Analysis

Within this section, the researcher presents the results for each of the research questions. Prior to determining the findings for each of the research questions, the

researcher had methodological considerations to consider. The methodological considerations were statistical adjustments and effect size calculations. The researcher provides a thorough description of and empirical evidence to support the appropriateness for each methodological consideration.

Methodological considerations. The first consideration was the need to adjust the significance level for Research Question 2, Research Question 3, and Research Question 4 because the researcher included multiple research questions and conducted multiple analyses (Pérez, & Pericchi, 2014). By conducting this adjustment, researchers change the critical value below which they consider their findings significant (Chen, Feng, & Yi, 2017). The researcher used the Bonferroni correction to achieve this adjustment. The second consideration was the researcher's use of Cohen's *d* as the measure of effect size. The researcher has discussed each of these two considerations and provided evidence to support their appropriateness in the subsequent sections of this applied dissertation.

Bonferroni correction. Researchers increase the likelihood of having at least one randomly significant result when they conduct multiple analyses involving several statistical tests and thereby increase their chances of not making an incorrect conclusion (Armstrong, 2014; Hall & Richardson, 2016). To decrease the likelihood of making an incorrect conclusion regarding significance between groups, researchers conduct a Bonferroni correction to determine an adjusted *p* value (Armstrong, 2014). A Bonferroni correction is an appropriate adjustment method for offsetting any issues researchers encounter when conducting multiple analyses and for decreasing the likelihood of them finding a significant difference between groups when no difference is present

(Armstrong, 2014). The Bonferroni correction calls for dividing the total significance level, which is .05 in most studies, by the number of analyses. The researcher had four research questions.

The researcher did not conduct significance testing for Research Question 1 but did conduct significance testing and primary analysis for each of the other three research questions. For the Bonferroni correction, the researcher divided the *p* value of .05 by 3, which yielded a *p* value of .0167. Thus, .0167 was the significance required of each analysis to have an overall significance level of .05 for the study. Additional analyses performed were considered secondary and therefore not included in the Bonferroni correction.

Cohen's d. Effect size is reported in order to indicate the importance of a study's results (American Psychological Association, 2010). Although effect sizes tend to be larger when *p* is significant, they measure something different. Effect size measures the standardized difference between the two means in a *t* test, or, in other words, the number of standard deviation units of the effect. Cohen's *d* is the effect size measure the American Psychological Association recommends for reporting the results of *t* tests (Nicol & Pexman, 2010). Cohen's *d* is one measure of effect size: the standardized difference between two means in a *t* test. Cohen's conventions for the social sciences are that $d = 0.2$ is considered a small effect size, $d = 0.5$ is a medium effect size, and $d = 0.8$ is a large effect size (Borenstein et al., 2001). Educational survey data are subjective as participants are expressing their opinions. Therefore, these conventions are appropriate for the current research.

The Teachers' Perceptions of Performance Pay (TPPP) Scale

The researcher used the results from the TPPP scale and a mean of the responses to the following five survey items to address all four research questions (see Appendix):

- A salary schedule without merit pay is an adequate way to pay teachers
- Merit pay is an appropriate way to increase teacher wages
- Merit pay is an appropriate way to reward teacher performance
- I approve of the teacher merit pay system used by M-DCPS in the past
- I approve of the merit pay system used by the state of Florida

All five items had three response options: *Yes* (coded as 1), *Somewhat* (coded as 2), and *No* (coded as 3). In order to make support for merit pay the higher scores, as is conventional, the researcher reversed the codes for the responses on the last four questions. Thus, participants' responses on the scale ranged from 1 (against merit pay on all five questions) to 3 (in favor of merit pay on all five questions).

Research Questions

Research Question 1. Research Question 1 was as follows: What are urban elementary teachers' perceptions of pay-for-performance initiatives?

TPPP analysis. In order to approximate the meaning of the responses to the original items using the TPPP scale scores, the researcher created a grouped variable, TPPPG. Note that the groups in TPPPG are named differently than the scale points in the original items. This is because TPPPG measures averages over all the items, whereas the scale points were used for individual items.

TPPPG had three groups: (a) 1−1.6 (similar to the scale point *No* in the original items); (b) 1.7−2.3 (similar to the scale point *Somewhat* in the original items); and, (c)

2.4−3 (similar to the scale point *Yes* in the original items). As can be seen in Table 2, nearly half (46.2%) of the respondents were mostly opposed to performance pay, approximately one-third (34.6%) had mixed feelings, and only about one in five (19.2%) were mostly supportive of performance pay.

Table 2

Teachers' Perceptions of Performance Pay, Grouped

Items 8-12	*n*	*%*
Mostly no	24	46.2
Mixed feelings	18	34.6
Mostly yes	10	19.2

The lack of support of performance pay was also reflected in the summary statistics for the (ungrouped) TPPP scale: The mean was 1.63 (*SD* = 0.56), the median was 1.6, and the mode was 1. The researcher found 11 of the respondents (21.2%) answered *No* to all five of the items in the scale.

Additional descriptive analysis. In addition to investigating whether teachers were in favor of using performance pay or not, the researcher wanted to investigate teachers' perceptions of various aspects of the performance pay initiative while the initiative was fully in place across the target school district and participants had a clear understanding of the standards used by school district leaders to determine performance pay for their teachers. Table 3, Table 4, and Table 5 include the summary statistics for Items 13, 14, and 15, respectively, and the researcher placed the mean rating for each category in descending order. Note that in these items, unlike in the main analysis

variable TPPP, lower ratings indicate support of performance pay and higher ratings indicate lack of support of performance pay.

The researcher found that teachers did not support the use of any of the methods of merit pay listed on the survey. Further analysis of the results revealed the lowest mean score (which was the highest level of support) was 4.50, which was halfway between partly and slightly supportive, and the highest mean score (lowest level of support) was 6.04, which was minimally supportive and only 0.96 from not being supportive at all.

The most popular (although still not supported) methods were students' individual growth and teachers' years of service while the least popular were the use of parent evaluations and students' district test scores. Table 3 lists the mean score for each method of merit pay.

Table 3

Descriptive Statistics for Item 13, Sorted by Mean Score

Item 13	*M*	*SD*	*Mdn*
Students' individual growth	4.50	1.91	4
Years of service	4.66	1.92	4
Students' year-to-year growth	4.94	1.75	4.5
Students' targeted growth	5.08	1.81	5
Students' classroom test scores	5.30	1.82	6
Administrator observation	5.38	1.50	6
Students' state test scores	5.62	1.44	6
Student portfolios	5.70	1.42	6
Other teachers' performance	5.82	1.71	7
Students report card grades from Common Core	5.82	1.55	7
Students' district test scores	5.86	1.43	7
Parent evaluations	6.04	1.43	7

Note. Scale points for the item: 1 *exclusively*, 2 *almost exclusively*, 3 *mainly*, 4 *partly*, 5 *slightly*, 6 *minimally*, 7 *not at all*.

Next, the researcher found that teachers were more likely to want themselves and teacher professional associations involved in not only developing a merit pay plan but also monitoring and evaluating a merit pay plan; conversely, they were less likely to want the Florida State DOE involved in developing a merit pay plan and least likely to want parents involved in developing a merit pay plan. Table 4 provides the descriptive statistics for Item 14 while Table 5 provides the descriptive statistics for Item 15.

Table 4

Descriptive Statistics for Item 14, Sorted by Mean Score

Item 14	*M*	*SD*	*Mdn*
Individual teachers	3.46	1.97	3
Teacher professional associations	3.92	1.97	4
Florida State DOE	5.24	1.70	5.5
Parents	6.30	1.22	7

Note. Scale points for the item: 1 *exclusively*, 2 *almost exclusively*, 3 *mainly*, 4 *partly*, 5 *slightly*, 6 *minimally*, 7 *not at all*.

Table 5

Descriptive Statistics for Item 15, Sorted by Mean Score

Item 15	*M*	*SD*	*Mdn*
Individual teachers	3.80	2.07	4
Teacher professional associations	4.10	1.94	4
Florida State DOE	5.06	1.77	4.5
Parents	6.38	1.12	7

Note. Scale points for the item: 1 *exclusively*, 2 *almost exclusively*, 3 *mainly*, 4 *partly*, 5 *slightly*, 6 *minimally*, 7 *not at all*.

Research Question 2. Research Question 2 was as follows: How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to grade level?

The researcher included teachers from seven different grade levels (Pre-K, kindergarten, and Grades 1-5); however, in order to perform the planned analysis, which was an independent *t* test, the researcher had to configure the groups because some teachers taught more than one grade level and because an independent *t* test does not allow a respondent to appear in both groups. To address issues pertaining to grouping, the researcher performed a correlational analysis to determine which grouping of grades

provided the cleanest break among teachers. To address these issues, a correlational analysis was performed. The researcher found a significant correlation between teaching kindergarten and teaching Grade 1 and significant correlations among teaching Grade 2 to Grade 5. Nonetheless, there was only one teacher respondent who taught pre-kindergarten. Table 6 includes the results of the correlational analysis across grade levels, with the shading in the table showing the two groupings.

Table 6

Correlations Among Levels Taught

Grade Levels	1	2	3	4	5	6	7
1. Pre-K (*n* =1)	1						
2. Kindergarten (*n* =11)	-.07	1					
3. Grade 1 (*n* = 11)	-.07	.31*	1				
4. Grade 2 (*n* = 15)	-.09	.09	-.02	1			
5. Grade 3 (*n* = 20)	-.11	-.02	-.12	.28*	1		
6. Grade 4 (*n* = 18)	-.10	.12	.02	.34*	.34*	1	
7. Grade 5 (*n* = 17)	-.10	.04	-.06	.46*	.29*	.36*	1

* Correlation is significant at $p < .05$ (2-tailed) or lower.

Based on the results of the correlational analysis, the researcher placed teachers who taught Pre-K, kindergarten, or Grade 1 in Group 1 and those who taught Grade 2, Grade 3, Grade 4, or Grade 5 in Group 2. Five teachers who taught students across all grade levels could have been classified in both groups (these were teachers who taught a special subject in all or most of the grade levels); however, and because the categories for a *t* test must not overlap, the researcher had to assign those five teachers to one of the two groups. The researcher considered both options and assigned all five teachers to each group and conducted two analyses: a primary analysis and a secondary analysis. To

answer the research question and to decrease the likelihood of making an incorrect conclusion by conducting multiple (as noted by discussion of the Bonferroni correction above) analyses, the researcher used the primary analysis to test for significance; however, the researcher did not test for significance in the secondary analysis.

For the primary analysis (used to test Research Question 2), the first group included the teachers who taught pre-K, kindergarten, or Grade 1 (n = 18; 34.6%), and the second group included all the other teachers (n = 34; 65.4%). By configuring the groups in that manner, the researcher reduced the discrepancy between the number of teachers in each group. In the secondary analysis, the first group included the teachers who taught in Grade 2, Grade 3, Grade 4, or Grade 5 (n = 13, 25.0%), and the second group included all the other teachers (n = 39, 75.0%).

The researcher concluded that teachers of Pre-K, kindergarten, and Grade 1 were slightly more supportive of performance pay (mean TPPP = 1.74) than teachers who taught only Grades 2−5 (mean TPPP = 1.57); however, the difference was not significant (t[50] = 1.07, p = .291), and the effect size was small, with Cohen's d = 0.30. Table 7 includes the results for each group.

Table 7

TPPP Results: Pre-K-Grade 1 and Grade 2-Grade 5

Grade Level Taught	*n*	*M*	*SD*	*t(50)*	*p*	Cohen's *d*
Pre-K, Kindergarten, or Grade 1	18	1.74	0.56	1.07	.291	0.30
Grades 2−5	34	1.57	0.56			

Note. Levene's test for equality of variances was not significant (F = 0.04, p = .834), so equal variances were assumed.

For the secondary analysis, the researcher found teachers of Pre-K, kindergarten, and Grade 1 only were slightly more supportive of performance pay (mean TPPP = 1.68) than

teachers of Grades 2−5 (mean TPPP = 1.62); however, the difference was even smaller (0.06) than the difference for the primary comparison (0.17). As noted above, the secondary analysis did not include significance testing.

Research Question 3. Research Question 3 was as follows: How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to tenure and non-tenure status?

The researcher divided teachers into two groups: those who were tenured (29 teachers, 55.8%) and those who were not tenured and worked on an annual contract (23 teachers, 44.2%). The researcher found tenured teachers were slightly less supportive of performance pay (mean TPPP = 1.59) than non-tenured teachers were (mean TPPP = 1.69); however, the difference was small (-0.10) and not significant (t[50] = -0.641, p = .524). Further analysis of the findings revealed that Cohen's d was only 0.16, which was even smaller than a small effect. Table 8 includes the results for Research Question 3.

Table 8

TPPP Results: Tenured and Nontenured Teachers

Tenure Status	*n*	*M*	*SD*	*t(50)*	*p*	Cohen's *d*
Tenured teachers	29	1.59	0.56	-0.64	.524	0.16
Nontenured (contract) teachers	23	1.69	0.57			

Note. Levene's test for equality of variances was not significant (F = 0.29, p = .594), so equal variances were assumed.

Research Question 4. Research Question 4 was as follows: How do urban elementary teachers' perceptions of pay-for-performance initiatives vary according to their years of experience?

The proposed data analysis for Research Question 4 was to compare teachers with 10 or more years of experience to teachers with fewer than 10 years of experience. This

was the primary analysis; however, the researcher conducted additional analyses because the data suggested that these may be an appropriate methodology for analyzing data.

Primary analysis. The researcher divided teachers into one of the following two groups: those with fewer than 10 years of experience (eight teachers, 15.4%) and those with 10 or more years of experience (44 teachers, 84.6%). The researcher found that teachers with fewer than 10 years of experience were more supportive of performance pay (mean TPPP = 2.05) than teachers with 10 or more years of experience (mean TPPP = 1.56).

The researcher also concluded that the difference (0.49) was almost significant ($t[50] = 2.41$, $p = .020$). This is close to the critical significance level of .0167 when recalling that the Bonferroni correction was used. Cohen's *d* was 0.88, which is considered a large effect size. The data illustrated in Table 9 provides the results of this analysis.

Table 9

TPPP Differences: Primary Analysis

Teaching experience	*n*	*M*	*SD*	*t(50)*	*p*	Cohen's *d*
Fewer than 10 years	8	2.05	0.58	2.41	.020	0.88
10 or more years	44	1.56	0.53			

Note. Levene's test for equality of variances was not significant ($F = .28$, $p = .598$), so equal variances were assumed.

Cutpoint of 20 years of teaching experience. A cross tabulation of TPPPG by 5-year experience groups was the next analysis the researcher conducted. A cross tabulation is appropriate when determining the relationship between two or more variables (Umer & Razi, 2018). The researcher determined that teachers with 20 or fewer years of experience (n = 34; 65.4%; mean TPPP = 1.78) were much more likely to support performance pay than teachers with more than 20 years of experience (n = 18; 34.6%; mean TPPP = 1.37). In this case, the difference was larger (0.41) and had a p value that would be considered significant had it been the primary analysis, even if the researcher used the Bonferroni correction (t[49.62] = 3.07, p =.003). In addition, Cohen's d was 0.84S, which is considered a large effect size.

Table 10

TPPP Differences: Secondary Analysis

Teaching experience	*n*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
20 or fewer years	29	1.78	0.60	49.62	3.07	.003	0.84
20 or more years	23	1.37	0.35				

Note. Levene's test for equality of variances was significant (F = 11.32, p = .001), so equal variances were not assumed.

Correlational analysis. When considering the large (> .8) effect sizes and low p values (< .05) in the analyses for the cutpoints of 10 and 20 years, combined with the mean TPPP levels for each of the experience groups, the researcher believed there may be a negative linear correlation between years of teaching experience and perception of performance pay. In other words, as the number of years of teaching experience increased, teachers' perceptions of performance pay decreased. The correlation for this analysis was negative (r = -.32) and very small (p = .019).

Summary

Nearly half (46.2%) of the respondents were mostly opposed to performance pay, approximately one-third (34.6%) had mixed feelings, and only about one in five (19.2%) were mostly supportive of performance pay (The mean was 1.6 on a scale from 1 to 3.). The researcher found no significant difference in perceptions of performance pay among teachers who taught at different grade levels or between tenured and nontenured teachers. However, the researcher did find a relationship between years of teaching experience and teachers' perceptions of performance pay, with the number of years of teaching increasing and teachers' perceptions of performance pay decreasing.

Chapter 5: Discussion

Introduction

The purpose of this quantitative study was to investigate teachers' perceptions of a teacher performance-pay initiative. The data collection tool was a modified version of Pemberton-Albright's 2011 survey. Pemberton-Albright (2011) created the survey and used it in a study titled *The Merit of Merit Pay*. This researcher modified Pemberton-Albright original 25-item survey and included 15 Likert-scale responses and four demographic items addressing the following areas: (a) gender; (b) grade-level or specialty area; (c) overall number of years in teaching; and, (d) tenure/non-tenure status.

Summary of Findings

The researcher found that more than 80% of the respondents were either mostly opposed to or had mixed feelings about performance pay for teachers and less than 20% of respondents were mostly supportive of performance pay for teachers. Further analysis of the findings revealed that teachers who were mostly opposed to performance pay had the largest representation, with close to half (46.2%) of all respondents choosing this option. When comparing teachers' perceptions of performance pay across grade levels, the researcher found that although Pre-K, kindergarten, and Grade 1 teachers were slightly more supportive than Grade 2, Grade 3, Grade 4, and Grade 5 teachers (TPPP difference of 0.17 between the two groups), the difference between the two groups was not significant and the effect size was small.

When comparing teachers' perceptions of performance pay on the basis of tenure and non-tenure status, the researcher concluded that non-tenure teachers were slightly more supportive than tenured teachers (TPPP difference of 0.10 between the two groups);

however, and similar to the result for research question 2, the difference was not significant.

Interpretation of Findings

Although there were no statistically significant differences in perceptions of performance pay among teachers who taught at different grade levels or between tenured and nontenured teachers, the researcher was able to gain a clear understanding of teachers' perceptions of performance-pay initiatives. The researcher found that teachers who were mostly opposed to performance pay had the largest representation across the three categories and teacher who had mixed feelings had the second largest representation. Based on these findings, the researcher concludes that close to half of the teachers at the Title 1 elementary school had negative perceptions of performance-pay initiatives.

The researcher also concludes that the majority of those opposed to performance-pay initiatives have 20 or more years' teaching experience. The researcher uses the results of the correlation analysis as the basis for this conclusion as there was a negative correlation between the years of teaching experience and teachers' perceptions of performance-pay initiatives. Specifically, as the number of years of teaching increased, teachers' perceptions of performance pay decreased.

Although the teachers at the target school represented a small percentage of the overall percentage of teachers across the school district, the researcher believes the findings of this study have implications at the local and state level. Stakeholders from the target school district and across the target state who are directly involved in establishing, passing, and implementing legislation should consider these findings as they determine

the appropriateness and the feasibility of performance-pay initiatives going forward and the environment they may foster within schools across both the target school district and the state. Researchers underscored the transformative impact of collaboration on teachers' instructional effectiveness (Carter, 2015; Darling-Hammond, 2015; Fulbeck, 2014). Collaboration is especially effective when teachers share instructional practices and resources (Russ, 2015). Teachers who engage in consistent and substantive collaboration demonstrated higher levels of instructional effectiveness than teachers who did not. Education policymakers who implement performance pay initiatives may promote unhealthy competition among teachers, decrease the extent of collegiality among teachers, and undermine the overarching goal of performance-pay initiatives, which is to compel teachers to increase student achievement through the use of financial incentives (Carter, 2015; Routh, 2014; Russ, 2015). Carter (2015) found that close to half of teachers from one school district in the Southeast United States believed a performance pay-initiative led to contention among teachers and resulted in teachers being reticent to share pedagogical strategies and resources.

By implementing a performance-pay initiative, stakeholders from the target school district and across the target state may undermine the culture school leaders and teachers are seeking to establish to promote both teacher effectiveness and student achievement (Carter, 2015). The culture they are seeking to establish is one in which teachers work together in professional learning communities and review student data, analyze and refine their instructional practices, and align curriculum and lesson plans to state standards (Russ, 2015). Teachers and administrators who establish that type of a culture in schools create the conditions teachers need to maximize instructional

effectiveness and promote student learning (Routh, 2014). Establishing the conditions teachers need to maximize instructional effectiveness and promote student learning may not be feasible in the context of a performance-pay initiative.

School board members should also consider that a performance-pay initiative will not be successful unless teachers completely support the initiative and the corresponding methods they will use to determine if teachers are eligible for performance-pay bonuses (Viscardi, 2014). Of the 12 methods for determining performance pay listed on the survey, some of which the school board uses to determine the awarding of performance pay, participants were partly to slightly supportive of only three methods, slightly to minimally supportive of eight methods, and minimally to not at all supportive of one method. With this in mind, school board members should seek input regarding the criteria for determining the awarding of performance pay. Stephens (2015) believed seeking input from teachers regarding the criteria for determining performance pay was an effective approach for garnering teachers' support of a performance pay initiative.

In future contract negotiations, school board members should work in conjunction with the bargaining team for the teachers' union. The bargaining team for the teachers' union can disseminate a survey to all teachers across the target school district and ask them to state which items they believe the school board should use to determine the awarding of performance pay. The bargaining team for the teachers' union can then meet with the school board's bargaining team and present the findings from the survey. The two teams can compose a collective bargaining agreement and establish the criteria for determining performance pay. Finally, the bargaining team for the teachers' union can present the agreement to teachers across the school district, and they can either accept or

reject the terms of the agreement. Under this type of negotiation process, the school board is empowering teachers by giving them input, and ultimately, a degree of autonomy over the standards by which the school district would determine their performance.

At the state level, stakeholders should consider the findings of this study and the negative environment they may create in schools across the state by endorsing performance-pay initiatives. Recently, the state's educational policymakers revised a measure and passed reforms aimed at recruiting, retaining, and awarding successful teachers. These reforms, which are part of the target state's Best and Brightest program, provide bonuses for teachers who receive a rating of either effective or highly based on their yearly evaluations and students' tests scores. Prior to these recent reforms, the program awarded bonuses to only teachers who receive a rating highly effective rating. Under these reforms, more teachers will receive performance-pay bonuses. From a union perspective, however, these reforms may cause resentment among teachers who give their best each day; however, and because of circumstances beyond their control, they may not meet the qualifying criteria for the performance-pay bonuses. School-based professionals who are not classroom teachers, including school psychologists, school counselors, and media specialists, are at a distinct disadvantage as they either teach subjects or have roles within their buildings in which evaluating their effectiveness with respect to the performance is not possible because their students do not take standardized tests. Prior to implementing other performance-pay initiatives, educational policymakers should seek input from teachers who work in various parts of the target and seek their input regarding the criteria for determining performance pay for determining performance pay for not only teachers with student test scores but those without them.

Context of Findings

Prior researchers found that teachers had negative perceptions of performance pay **(**Edenfield, 2014; Routh, 2014; Russ, 2015). The researcher's finding that more than 80% of the target school's teacher were either mostly opposed to or had mixed beliefs about performance-pay initiatives on the basis of teacher performance aligned with that of Edenfield (2014), Russ (2015), and Routh (2014). Edenfield (2014) found that approximately 80% of one southeastern state's teachers opposed merit pay based on teacher performance in the area of student achievement. Russ (2015) investigated the perceptions of 250 teachers from two school districts in the midwestern part of the United States regarding merit-pay initiatives and concluded that teachers had negative perceptions of merit-pay initiatives in which the criteria for determining monetary rewards were teacher evaluations, students' standardized test scores, and both school- and district-level performance on standardized tests.

Although much broader in scope by including a random sample of teachers from one midwestern state when compared to the scope of both this study and Russ's 2015 study, Routh (2014) also found teachers had negative perceptions of merit-pay initiatives based on teacher evaluation. Therefore, Routh's finding aligned with that of both the researchers and Russ (2015). The researcher's finding that teachers with fewer than 10 years' experience were more supportive of performance pay than teachers with more than 10 years' experience, although the difference was not quite significant, and teachers with fewer than 20 years' experience were more supportive than teachers with more than 20 years' experience, did not align with that of Stephens (2015) who found the perceptions

of teachers from one large school district in the southeastern part of the United States did not differ based on years of teaching experience.

Implications of the Findings

The researcher found that teachers had negative perceptions of teacher performance pay and were mostly opposed to it. These findings may have implications in developing alternative pay scales for teachers. The researcher also found a relationship between the number of years of experience and perceptions. For each of the experience groups, the researcher believed there may be a negative linear correlation between years of teaching experience and perception of performance pay. In other words, as the number of years of teaching experience increased, teachers' perceptions of performance pay decreased. Because of the link between years of experience and negative perceptions of performance pay, an implication of the findings could allow administrators and professional organizations to take a closer look at alternative ways to pay teachers perhaps through a pay scale built in to the contract for teachers with less than 10 years and more than 10 years.

Limitations of the Study

Every study has limitations, and researchers should not hesitate to describe these limitations **(**Olufowote, 2017). The most common limitations of a study are sample size, sampling procedures, data collection instruments and procedures, research design, accessibility of participants (Olufowote, 2017; Theofanidis & Fountouki, 2019). By identifying their limitations, researchers convey a comprehensive understanding of their topic and the need to increase the body of knowledge within a discipline. The following sections provide each of this study's limitations.

Convenience sampling. The use of convenience sampling was the first limitation of this study. Convenience sampling is appropriate when researchers are unable to ensure that every participant in a population will have an equal chance of being included in their studies and when they must include the most accessible and convenient population (Asiamah, Mensah, & Oteng-Abayie, 2017; Etikan, Musa, & Alkassim, 2016). Reduced costs and simple and efficient execution are the most noteworthy advantages of convenience sampling (Etikan et al., 2016; Jager, Putnick, & Bornstein, 2017). Although researchers who use convenience sampling can glean insight regarding the characteristics of a segment of an accessible population, they cannot generalize their findings to the broader representative population, thereby impacting the external validity of their study (Etikan et al., 2016).

By using convenience sampling, which is a non-probability sampling procedure, the researcher was unable to determine if the participants in this study were an accurate representation of the population. Because the participants in this study were not an accurate representation of the population, the researcher could not generalize this study's findings to the broader population. Finally, the researcher was unable to control for discrepancies in the number of participants in each group. For example, when comparing TPPP differences based on experience, one group had eight participants and the other group had 44 participants.

Data analysis procedures. The second limitation was the data analysis procedures. The researcher conducted multiple analyses and used several statistical tests to answer each research question. Researchers increase the likelihood of making an incorrect conclusion when they conduct multiple analyses involving several statistical

tests (Armstrong, 2014; Hall & Richardson, 2016). To decrease the likelihood of making an incorrect conclusion, researchers conduct a Bonferroni correction to determine an adjusted *p* value (Armstrong, 2014). Although the Bonferroni correction is an appropriate for decreasing the likelihood of making incorrect conclusions, it cannot guarantee that researchers' conclusions are precise (Hall & Richardson, 2016). The researcher used a Bonferroni correction to offset any issues pertaining to multiple analyses and to increase the likelihood of making precise conclusions; however, the researcher cannot guarantee there were no significant differences in perceptions of performance pay among teachers who taught at different grade levels or between tenured and nontenured teachers nor can the researcher guarantee that a statistically significant relationship existed between years of teaching experience and teachers' perceptions of performance pay.

Data collection procedures. The third limitation was the researcher could not guarantee that participants provided truthful responses to the survey items. Participants in this study completed their survey on the Survey Monkey site. Survey Monkey is an online survey tool that researchers use to create, send, and analyze their surveys. Through Survey Monkey's data collection menu, researchers can determine who responds to survey, send an initial email to participants reminding them to complete a survey, send out reminder emails to participants who have not completed their survey, and collect the results of surveys (Forris, 2015). One concern regarding the use of the Survey Monkey platform is the ability of researchers to ensure participant confidentiality if a data breach occurs (Regmi, Waithaka, Paudyal, Simkhada, & van Teijlingen, 2016). Researchers using this form of data collection procedure must make every effort to ensure participant confidentiality (Regmi et al., 2016). To allay participants' concerns regarding

confidentiality, researchers can use the first page of the survey to inform participants that any identifying information will be removed and participants that they have the right to opt out of the survey at any time and not be part of the study. Although this measure may alleviate participants' concerns regarding confidentiality, researchers cannot guarantee that participants will answer all survey items truthfully.

With respect to this study, the researcher addressed any concerns participants may have had regarding confidentiality by stating that participation in this study was voluntary and that no identifying information would be recorded. The researcher included those statements at the beginning of the survey. By completing the survey, participants established their informed consent to participate. Despite assuring participants of the steps taken to ensure confidentiality, the researcher could not guarantee that participants provided accurate responses to the survey items.

Future Research Directions

The researcher found that despite implementing teacher performance pay in 2015, the target school district school board had not investigated teachers' perceptions of the teacher performance pay initiative across any of its schools (Francilus, 2015). The researcher sought to address this gap in knowledge by investigating teachers' perceptions of performance pay at one elementary school; however, and as noted above, the target school district is composed of 392 schools (171 = elementary, K-8 = 53, 48 = middle, 54 = high schools, 8 = combination of K-12 schools, 12 = alternative and exceptional student education centers, 45 migrant pre-k outreach centers and teacher advancement program).

Increase scope of study. When considering the scope of this study and the

overall number of schools under the jurisdiction of the target school district, the researcher investigated the perceptions of teachers from approximately 0.25% of the target school district's schools; therefore, the first recommendation for future research is to conduct a similar study at other schools across the school district. Researchers conducting this type of a study should include multiple elementary schools.

Researchers should also expand this scope of this study by investigating the perceptions of teachers who work at K-8, middle, and high schools and alternative learning centers across the school district. The researcher notes that the target school district is composed of three geographic regions, and each region is under the jurisdiction of a region superintendent and a central office staff. Each region is subdivided into feeder patterns, and within each region, the school district has established feeder patterns based on geography. Students living within a feeder pattern attend an assigned elementary school, middle school, and high school. Each feeder pattern has varying number of elementary, middle, and high schools. Within each of the target school district's regions, future research should investigate teachers' perceptions of performance pay across all elementary, middle, and high schools within a feeder pattern and then conduct comparisons across school levels (elementary schools vs. middle schools vs. high schools (Routh, 2014).

Replicate study in other settings. A review of the related literature also revealed that target school districts across the southeastern state had not investigated teachers' perceptions of the teacher performance pay initiative. With this in mind, the researcher recommends replicating this study in other school districts across the southeastern state. Similar to the recommendations for further research within the target school districts,

researchers should conduct this study at the middle, elementary, and high school levels. To further expand the body of knowledge, future research should compare teachers' perceptions of the teacher performance pay initiative based on setting (rural vs. suburban vs. urban); hence, researchers could compare the perceptions of teachers who work in rural settings and those who work in urban settings and the perceptions of teachers who work in suburban settings and those who work in urban settings (Marsh, 2014).

Use random sampling technique. As noted in the section on limitations, the researcher used convenience sampling procedures, thereby generalizing findings to other settings was not possible. Hence, the third recommendation is for future research to replicate this study across other schools within the target school district and in school districts across the state and use random sampling procedures. Researchers who use a random sampling technique decrease the chance of selection bias and increase the chance of selecting a representative sample from the target population (Sharma, 2017).

Implement a qualitative study model. Another recommendation is for future research to use a qualitative research approach to investigate teachers' perceptions of performance pay. A qualitative approach is appropriate when researchers seek to provide an in-depth description of a phenomenon from the perspective of those with first-hand knowledge of the phenomenon (Hammarberg, Kirkman, & de Lacey, 2016). This in-depth description is not possible with a quantitative approach (Creswell, 2013). Researchers who conduct a qualitative study could use a case study design and collect multiple forms of data. The forms of data they could use include interviews, observations, and artifacts. Although researchers who use a qualitative approach and a case study design may include only 10 to 20 participants, they may be able to provide rich

descriptions of teachers' perceptions regarding performance pay initiatives and gain a clear understanding of the factors informing teachers' perceptions and what if performance pay models they prefer.

References

Alger, V. E. (2014). *Teacher incentive pay that works: A global survey of programs that improve student achievement*. British Columbia, Canada: Barbara Mitchell Centre for Improvement in Education, Fraser Institute.

Ali, Z., & Bhaskar, S. B. (2016). Basic statistical tools in research and data analysis. *Indian Journal of Anesthesia*, *60*(9), 662–669. doi:10.4103/0019-5049.19062

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.

Amrein-Beardsley, A., Pivovarova, M., & Geiger, T. J. (2016). Value-added models: What the experts say. *Phi Delta Kappan*, *98*(2), 35-40. doi:10.1177/0031721716671904

Anderson, A. D., Hunt, A. N., Powell, R. E., & Dollar, C. B. (2013). Student perceptions of teaching transparency. *The Journal of Effective Teaching*, *13*(2), 38-47.

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502-508. doi:10.1111/opo.12131

Asiamah, N., Mensah, H. K., & Oteng-Abayie, E. (2017). General, target, and accessible population: Demystifying the concepts for effective sampling. *The Qualitative Report*, *22*(6), 1607-1621.

Assessment, Research, and Data Analysis Division. (2014). Explaining the value-added model. Retrieved from http://www.oada.dadeschools.net

Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher

performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher, 44*(2), 77-86. doi:10.3102/0013189x15574904

Bauries, S. R., Sutherland, B. J., & Legare, C. B. (2014) *Brief of education law and educational measurement professors as amici curiae in support of plaintiffs-appellants, Cook v. Steward* (Policy Brief). Retrieved from http://uknowledgeuky.edu/law_facpub_advocacy/3

Blazer, C. (2011). *Status of teacher performance pay programs across the United States.* (Information Capsule).

Borenstein, M., Rothstein, H., & Cohen, J. (2001). *Power and precision*. Englewood, NJ: Biostat, Inc.

Britton, J., & Propper, C. (2016). Teacher pay and school productivity: Exploiting wage regulation. *Journal of Public Economics*, *133*, 75-89. doi:10.1016/j.jpubeco.2015.12.004

Buck, S., & Greene, J. P. (2011). Blocked, diluted, and co-opted. *Education Next, 11*(2), 26-31.

Buddhi, R. (2007). Merit pay for Florida teachers: Design and implementation issues. RAND Education.

Carlon, L. G. (2015). *Merit pay for educators: An investigation of components significantly impacting student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3680827)

Carruthers, C., Figlio, D., & Sass, T. (2018). Did tenure reform in Florida affect student test scores? *Evidence Speaks Reports 2*(52), 1-14. Retrieved from https://pdfs.semanticscholar.org/0bbe/70264955dbb4d2befb1f58e6a19ee19ac2ab.

pdf?_ga=2.99473305.689421002.1565303599-368544635.1560637627

Chen, S. Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, *9*(6), 1725–1729. doi:10.21037/jtd.2017.05.34

Cocke, J. W. (2014). *Florida teachers' knowledge and perceptions about student learning growth, assessment, and the Florida value-added model* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3637958)

Collins, C., & Amrein-Beardsley, A. (2014). Putting growth and value-added models on the map: A national overview. *Teachers College Record*, *116*(1*)*, 1-18.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation. *Educational Researcher*, *44*(2), 132-137. doi:10.3102/0013189X15575346

Davis, J. B. (2014). *The relationship between the growth score and the overall TEAM observation rating for teachers in Tennessee* (Unpublished doctoral dissertation). East Tennessee State University, Johnson City, TN.

Dean, J. D. (2015). *Comparing schools: From value added to sound policy* (Unpublished doctoral dissertation). University of Arkansas, Fayetteville, AR.

Deckers, L. (2014). *Motivation: Biological, psychological, and environmental* (4th ed.). Boston, MA: Allyn & Bacon.

Doran, H. C. (2014) Methods for incorporating measurement error in value-added models and teacher classifications. *Statistics and Public Policy*, *1*(1), 114-119. doi: 10.1080/2330443X.2014.955228

Edenfield, J. (2014). *Teachers' perceptions of merit pay in Georgia* (Unpublished doctoral dissertation). Georgia Southern University, Statesboro, GA.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4. doi:10.11648/j.ajtas.20160501.11

Farley, A. N. (2017). *Review of for good measure? Teacher evaluation policy in the ESSA era*. Boulder, CO: University of Colorado, School of Education.

Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? lessons from Chicago's low-performing schools. *American Educational Research Journal, 44*(3), 594;629;-629. doi:10.3102/0002831207306767

Florida Department of Education. (2008). *Florida School Recognition Program: 2008 general information*. Retrieved from http://www.fldoe.org/evaluation/geninfo.asp

Florida Department of Education. (2012). *No Child Left Behind*. Retrieved from http://www.fldoe.org/nclb/

Florida Department of Education. (2013). *Governor Scott signs Student Success Act* [Press release]. Retrieved from http://www.flgov.com/2011/03/24/governor-scott-signs-studentsuccess-act/

Florida Department of Education. (2014). *Florida School Recognition Program*. Retrieved from http://www.fldoe.org/how-do-i/evaluation-reporting.stml

Florida Department of Education. (2016). *2015-16 guide to calculating school and district grades*. Retrieved from http://www.fldoe.org/core/fileparse.php/18534/urlt/SchoolGradesCalcGuide16.pd

f.

Florida Educational Accountability Act. (1971). *Fla. Stat. ch. 76-223*.

Forand, S. (2012). *Teachers' attitudes and perceptions about pay-for-performance* (Unpublished doctoral dissertation). Northeastern University, Boston, MA.

Forris, S. E. (2015). *The quest for work and family balance using flexible work arrangements* (Unpublished doctoral dissertation). Walden University, Minneapolis, MN.

Francilus, J. (2015, September). Miami-Dade teachers ratify new performance-pay system. *Politico*. Retrieved from https://www.politico.com/states/florida/story/2015/09/miami-dade-teachers-ratify-new-performance-pay-system-025380

Fulbeck, E. S. (2014). Teacher mobility and financial incentives: A descriptive analysis of Denver's ProComp. *Educational Evaluation and Policy Analysis, 36*(1), 67-82. doi:10.3102/0162373713503185

Gamson, D., & Hodge, E. (2016). Education research and the shifting landscape of the American school district, 1816 to 2016. *Review of Research in Education, 40*(1), 216-249. doi:10.3102/0091732X16670323

Gemeda, F. T. (2015). Exploring teachers' motivation for teaching and professional development in Ethiopia: Voices from the field. *Journal of Studies in Education, 5*(2), 169-186. doi:10.5296/jse.v512.7459

Gius, M. (2013). Using a difference-in-differences approach to estimate the effects of teacher merit pay on student performance. *Eastern Economic Journal, 39*(1), 111–120. doi:10.1057/eej.2012.22

Guerere, C. (2013). *Value-added and observational measures used in the teacher evaluation process: A validation study* (Unpublished doctoral dissertation). University of South Florida, Tampa, FL.

Gurnsey, R. (2017). *Statistics for research in psychology: A modern approach using estimation*. Thousand Oaks, CA: Sage Publications.

Haertel, E. (2013). *Reliability and validity of inferences about teachers based on student test scores*. Princeton, NJ: Educational Testing Service.

Hall, M., & Richardson, T. (2016). Basic statistics for comparing categorical data from 2 or more groups. *Hospital Pediatrics*, *6*(6), 383-385. doi:10.1542/hpeds.2015-0273

Hammarberg, K., Kirkman, M., & de Lacey, S. (2016). Qualitative research methods: When to use them and how to judge them. *Human Reproduction*, *31*(3), 498–501. doi:10.1093/humrep/dev334

Han, J., & Yin, H. (2016). Teacher motivation: Definition, research development and implications for teachers. *Cogent Education*, *3*(1), 1-50. doi:10.1080/2331186X.2016.1217819

Harrison, C., & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narratives surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education*, *87*(5), 517-534. doi:10.1017/CBO9780511490934

Jager, J., Putnick, D. L., & Bornstein, M. H. (2017). II. More than just convenient: The scientific merits of homogeneous convenience samples. *Monographs of the Society for Research in Child Development*, *82*(2), 13–30. doi:10.1111/mono.12296

Jensen, N. C. (2011). The value of value-added measures. *Office for Education Policy, 8*(2), 1-3.

Jones, K. (2014). *Teachers' perspectives on the third grade retention policy at the primary grade level* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3632216)

Jones, M. D. (2013). Teacher behavior under performance pay incentives. *Economics of Education Review*, *37*, 148-164. doi:10.1016/j.econedurev.2013.09.005

Kane, T., & Staiger, D. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains.* Seattle, WA: Bill and Melinda Gates Foundation.

Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, *68*(6), 540–546. http://doi.org/10.4097/kjae.2015.68.6.540

Laliberte, M. D. (2015). *Florida's A++ plan: An expansion and expression of neoliberal and neoconservative tenets in state educational policy* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3706265)

Lauen, D. L., & Kozlowski, K. P. (2014). *Performance incentives in North Carolina*. Retrieved from the Consortium for Educational Research and Evaluation-North Carolina website https://publicpolicy.unc.edu/files/2015/07/Race-to-the-Top-Performance-Incentives-in-North-Carolina-August-2014.pdf

Lavania, M., Cohen-Vogel, L., & Lang, L. B. (2015). The Common Core State Standards initiative: An event history analysis of state adoption. *American Journal of Education*, *121*(2), 145-182. doi:10.1086/679389

Liang, G., & Akiba, M. (2015). Characteristics of teacher incentive pay programs: A

statewide district survey. *Journal of Educational Administration, 53*(6), 702-717. doi:10.1108/JEA-09-2013-0106

Loeb, S. (2013). How can value-added measures be used for teacher improvement. *The Carnegie Knowledge Network.* Retrieved from https://cepa.stanford.edu/sites/default/files/CKN-Loeb_Teacher-Improvement.pdf

Lowe, J. C. (2013). *Assessment of teachers' reactions to a knowledge and skills-based pay structure at an international school* (Unpublished doctoral dissertation). Lehigh University, Bethlehem, PA.

Maranto, R. (2014). Tough teacher evaluation and high morale? *Educational Leadership, 71*(5).

Marsh, J. S. (2014). *Performance-based pay: Perceptions of elementary teachers of the Pandora School District* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3680827)

McCullough, M., English, B., Angus, M. H., & Gill, B. (2015). *Alternative student growth measures for teacher evaluation: Implementation experiences of early-adopting districts*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from http://ies.ed.gov/ncee/edlabs.

Meier, K. J., & Rutherford, A. (2016). *The politics of African-American education : Representation, partisanship, and educational equity*. New York, NY: Cambridge University Press.

Messerli, J. (1972). *Horace Mann, a biography.* New York: Random House.

Moss, P. A., & Haertel, E. H. (2016). Engaging methodological pluralism. In D. Gitomer and C. Bell (Eds), *Handbook of research on teaching* (5th Ed), (pp. 127-247). Washington, DC: AERA.

Munroe, A. (2017). Measuring student growth within a merit-pay evaluation system: Perceived effects on music teacher motivation career commitment. *Contributions to Music Education*, *42*, 89-105.

Najera, T. L. (2017). *Confirming predictors of rural teacher expectancy* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.

Nicol, A. A. M., & Pexman, P. M. (2010). *Presenting your findings: A practical guide for creating tables*. Washington, DC: American Psychological Association.

Nieberg, A. J., Pieper, J. R., & Trevor, C. O. (2016). Pay-for-performance's effect on future employee performance: Integrating psychological and economic principles toward a contingency perspective. *Journal of Management*, *42*(7), 1753-1783. doi:10.1177/0149206313515520

O'Connor, J. (2013). *What the Florida teacher evaluation system lawsuit could mean for other states.* Retrieved from State Impact Florida website:http://stateimpact.npr.org/florida/2013/05/15/what-the-florida-teacher-evaluation-lawsuitcould-mean-for-other-states/

Olcum, D., & Titrek, O. (2015). The effect of school administrators' decision-making styles on teacher job satisfaction. *Procedia - Social and Behavioral Sciences*, *197*, 1936-1946. doi:10.1016/j.sbspro.2015.07.575

Olufowote, J. O. (2017). Limitations of research. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (pp. 863-864). Thousand Oaks,

CA: Sage.

Owens, J. (2013). *Confessions of a bad teacher*. Naperville, IL: Sourcebooks.

Ozoemena, P. O. (2013). Improved incentive system in teaching profession as strategy for quality delivery and enhanced productivity in contemporary Nigeria. *International Journal of Education and Research, 1*(7), 1-12.

Parijat, P., & Bagga, S. (2014). Victor vroom's expectancy theory of motivation–an evaluation. *International Research Journal of Business and Management, 7*(9), 1-8.

Pemberton-Albright, M. M. (2011). *The merit of merit pay* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3453880)

Pérez, M. E., & Pericchi, L. R. (2014). Changing statistical significance with the amount of information: The adaptive α significance level. *Statistics & Probability Letters*, *85*, 20–24. doi:10.1016/j.spl.2013.10.018

Perry, T. (2016), English value-added measures: Examining the limitations of school performance measurement. *British Educational Research Journal*, *42*(6), 1056-1080. doi:10.1002/berj.3247

Pivovarova, M., Amrein-Beardsley, A., & Broatch, J. (2016). Value-added models (VAMs): Caveat emptor. *Statistics and Public Policy*, *3*(1), 1-10. doi:10.1080/2330443X.2016.1164641

Postal, L. (2017, May 8). Massive education bill alters testing, recess, teacher evaluations, bonuses, charters. *Orlando Sentinel*. Retrieved from https://www.orlandosentinel.com/features/education/school-zone/os-legislature-education-recess-testing-bonuses-20170508-story.html

Pressley, T. M. F. (2015). *Teachers view of the role VAM plays in their work in the school and school community* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3724350)

Regmi, P. R., Waithaka, E., Paudyal, A., Simkhada, P., & van Teijlingen, E. (2016). Guide to the design and application of online questionnaire surveys. *Nepal Journal of Epidemiology*, *6*(4), 640–644. doi:10.3126/nje.v6i4.17258

Research Services, Dade County Public Schools, (2017). Retrieved from http://drs.dadeschools.net.

Rice, J. K., Malen, B., Jackson, C., & Hoyer, K. M. (2015). Time to pay up: Analyzing the motivational of financial awards in a TIF program. *Educational Evaluation and Policy Analysis*, *37*(1), 29-49. doi:10.3102/0162373714524622

Ritter, G. W., & Jensen, N. C. (2010). The delicate task of developing an attractive merit pay plan for teachers. *Phi Delta Kappan, 91*(8), 32-37. doi:10.1177/003172171009100807

Robertson-Kraft, C. (2014). *Teachers' motivational responses to new teacher performance management systems: An evaluation of the pilot of Aldine ISD's inVEST system* (Unpublished doctoral Dissertation). University of Pennsylvania, Philadelphia, PA.

Robles, Y. (2015, July 27). Evolution of teacher evaluations is leading performance pay reforms. *The Denver Post*. Retrieved from https://www.denverpost.com/2015/07/27/evolution-of-teacher-evaluations-is-leading-performance-pay-reforms/

Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. N. (2013). Feeling the Florida

heat? How low-performing schools respond to voucher and accountability pressure. *American Economic Journal: Economic Policy, 5*(2), 251-281. doi:10.1257/pol.5.2.251

Routh, J. B. (2014). *Missouri teachers, administrators, superintendents' perceptions of teacher performance pay* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3713984)

Russ, R. D. (2015). *Teacher attitudes regarding performance-based pay* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3707142)

Sharma, G. (2017). Pros and cons of different sampling techniques. *International Journal of Applied Research, 3*(7), 749-752.

Smith, C. (2018, March 19). Richardson ISD looks to Dallas for inspiration for its new turnaround program. *Dallas News*. Retrieved from https://www.dallasnews.com/news/education/2018/03/19/richardson-isd-looks-dallas-inspiration-new-turnaround-program

Springer, M., & Winters, M. (2009). *The NYC teacher pay-for-performance program: Early evidence from a randomized trial.* Center for Civic Innovation at the Manhattan Institute. Retrieved from http://www.manhattan-institute.org/pdf/cr_56.pdf

Stephens, D. M. (2015). *Mississippi teachers' perception of merit pay* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3733079)

Strauss, V. (2017, February 17). Trump backs merit pay for teachers-but one Florida

school system now says it doesn't work. *The Washington Post,* 3-5.

Stutz, T. (2013, October 13). Texas merit pay plan for teachers quietly disappears. *Dallas Morning News*. Retrieved from http://www.dallasnews.com/news/localnews/20131013-texas-merit-pay-plan-for-teachers-quietly-disappears.ece

Suen, L. J., Huang, H. M., & Lee, H. H. (2014). A comparison of convenience sampling and purposive sampling. *Hu Li Za Zhi, 61*(3), 105-111. doi:10.622/JN.61.3.105

Theofanidis, D., & Fountouki, A. (2019). Limitations and delimitations in the research process. *Perioperative Nursing, 7*(3), 155–162. doi:10.5281/zenodo.2552022

Toch, T. (2016). Grading the graders: A report on teacher evaluation reform in public education. Retrieved from https://educationpost.org/wp-content/uploads/2016/05/Thomas-Toch-Grading-the-Graders.pdf

Tucson Unified School District. (2017). 301 pay for performance plan. Retrieved from https://www.nctq.org/dmsView/Tucson_301Plan17-18

Umer, M., & Razi, S. (2018). Analyzing research methodologies and publication trends in service marketing literature. *Cogent Business & Management, 5*(1), 1-16. doi:10.1080/23311975.2018.1446265

U.S. Department of Education. (2012). *Race to the Top: Florida report.* Washington, DC: U.S. Department of Education.

U.S. Department of Education. (2016). *Race to the top fund*. Washington, DC: U.S. Department of Education. Retrieved from https://www2.ed.gov/programs/racetothetop/index.html.

Vacca, R. S. (2016). Teacher classroom effectiveness and student learning outcomes:

Policy implications. *The CEPI Education Law Newsletter*, *15*(2), 1-7.

Viscardi, D. (2014). *The teacher pay for performance phenomenon* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3667762)

Vroom, V. H. (1964). *Work and motivation.* New York, NY: Wiley.

Warne, R. T. (2017). *Statistics for the social sciences: A general linear model approach*. New York, NY: Cambridge University Press.

Wells, P., Combs, J. P., & Bustamante, R. M. (2013). Team performance pay and motivation theory: A mixed methods study. *Journal of Research in Education*, *23*(2), 116-125.

Wesson, L., Potts, K., & Hill, K. (2015). *Use of value-added in teacher evaluations: Key concepts and state profile.* Nashville, TN: Tennessee Comptroller of the Treasury, Office of Research and Education Accountability. Retrieved from http://www.comptroller.tn.gov/Repository/RE/ValueAdded2015.pdf

Will, M. (2018). An instructional coach can improve school performance. Here's proof. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/teacherbeat/2018/02/in_tennessee_will_teacher_lead.html

Yuan, K. V., Le, V., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2013). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation and Policy Analysis*, *35*(1), 3–22. doi:10.3102/0162373712462625

Appendix

Teacher Perceptions of Performance Pay

Teacher Perceptions of Performance Pay

Welcome to Teacher Perceptions of Performance Pay
by Wendy Firtell

Thank you for participating in my survey. Your feedback is important.

Teacher Perceptions of Performance Pay

Demographics

Please answer the first 7 questions which are regarding demographics:

* 1. Gender:

* 2. What is your age?

* 3. Are you currently in the DROP Deferred Retirement Option Plan?

Teacher Perceptions of Performance Pay

* 4. Years of Experience as a teacher?

* 5. What is your present teaching status?

- ○ Tenured Teacher (Grandfathered in on a professional contract)
- ○ Annual Contract Teacher
- ○ Teacher and Administrator (certified teacher but acting as a school administrator)

Other (please specify)

* 6. What grade level(s) do you currently teacher? (May choose more than one)

- ☐ Pre-Kindergarten
- ☐ Kindergarten
- ☐ Grade 1
- ☐ Grade 2
- ☐ Grade 3
- ☐ Grade 4
- ☐ Grade 5

Other (please specify)

* 7. What area(s) of instruction do you teach?

- ☐ Pre-K Regular Ed.
- ☐ K-2 Gifted/Extended Foreign language (EFL)
- ☐ K-2 English as a Second Language (ESOL)
- ☐ K-2 Exceptional Student Education (ESE)
- ☐ K-2 Regular Ed.
- ☐ 3-5 Regular Ed.
- ☐ 3-5 Gifted/Extended Foreign Language (EFL)
- ☐ 3-5 Exceptional Student Education (ESE)
- ☐ 3-5 English as a Second Language (ESOL)
- ☐ Fine Arts (Music, Art)
- ☐ Bilingual Specialist
- ☐ Physical Education
- ☐ Math/Reading Coaches

Other (please specify)

Teacher Perceptions of Performance Pay

Teacher Opinions

Please state how much you agree with the following:

* 8. A salary schedule without merit pay is an appropriate way to pay teachers.

- [ ] Yes
- [ ] Somewhat
- [ ] No

* 9. Merit pay is an appropriate way to increase teacher wages.

- [ ] Yes
- [ ] Somewhat
- [ ] No

* 10. Merit pay is an appropriate way to reward teacher performance.

- [ ] Yes
- [ ] Somewhat
- [ ] No

* 11. I approve of the teacher merit pay system used by M-DCPS in the past.

- [ ] Yes
- [ ] Somewhat
- [ ] No

* 12. I approve of the merit pay system used by the state of Florida

- [ ] Yes
- [ ] Somewhat
- [ ] No

Teacher Perceptions of Performance Pay

Teacher Opinions

* 13. To which of the following should merit pay be tied?

	Exclusively	Almost exclusively	Mainly	Partly	Slightly	Minimally	Not at all
Administrator Observations	○	○	○	○	○	○	○
Students' **state** test scores	○	○	○	○	○	○	○
Students' **district** test scores	○	○	○	○	○	○	○
Years of Service	○	○	○	○	○	○	○
Other Teachers' Performance	○	○	○	○	○	○	○
Parent Evaluations	○	○	○	○	○	○	○
Student Portfolios	○	○	○	○	○	○	○
Students' Individual Growth	○	○	○	○	○	○	○
Students' Targeted Growth	○	○	○	○	○	○	○
Students' Year-to-Year Growth	○	○	○	○	○	○	○
Students' Classroom Test Scores	○	○	○	○	○	○	○
Students' Report Card Grades from Common Core Curriculum	○	○	○	○	○	○	○

Teacher Perceptions of Performance Pay

Teacher Opinions

* 14. Who should be involved in **developing** a merit pay plan?

	Exclusively	Almost Exclusively	Mainly	Partly	Slightly	Minimally	Not at all
Individual Teachers	○	○	○	○	○	○	○
Teacher Professional Organizations	○	○	○	○	○	○	○
Parents	○	○	○	○	○	○	○
Florida State Department of Education (DOE)	○	○	○	○	○	○	○

* 15. Who should be involved in **monitoring** and **evaluating** a merit pay plan?

	Exclusively	Almost Exclusively	Mainly	Partly	Slightly	Minimally	Not at all
Individual Teachers	◯	◯	◯	◯	◯	◯	◯
Teacher Professional Organizations	◯	◯	◯	◯	◯	◯	◯
Parents	◯	◯	◯	◯	◯	◯	◯
Florida State Department of Education (DOE)	◯	◯	◯	◯	◯	◯	◯

16. Any additional comments (optional)

Teacher Perceptions of Performance Pay

Thank you for taking my survey.