PAY FOR MY PERFORMANCE:
TEACHERS’ EXPERIENCES WITH REFORM-BASED EVALUATION TOOLS IN A LARGE SUBURBAN SCHOOL DISTRICT

A thesis presented
by
Paul Satty

to
The School of Education

In partial fulfillment of the requirements for the degree of
Doctor of Education

in the field of
Education

College of Professional Studies
Northeastern University
Boston, Massachusetts
December 2016
Abstract

Teacher performance evaluations continue to be one of the most debated practices within the formal education system. While the purpose of teacher evaluation is to improve classroom instruction, provide data for staff development, and assure that the district's vision is developed through teacher instruction, many teachers express and experience fear and anxiety about the process, although others welcome the standard process of evaluation and the ability to receive feedback and input on their professional work.

The purpose of this study was to investigate, analyze, and describe how teachers perceive the new Broward County, Florida, school district evaluation tool in regards to its effectiveness, fairness, and equity, as well as its relationship to performance pay. The goal was to make Broward County Public Schools policymakers and leaders aware of current trends in teacher attitudes and to provide an analysis of the successes and failures of the evaluation tool.

This qualitative study is based on eight in-depth interviews with teachers from one Broward County high school. The data revealed the purposes, perceptions, and attitudes of these education professionals toward the current evaluation methods.

The following conclusions were drawn from the responses of the teachers interviewed in this study: a large percentage of teachers (a) did not believe that the evaluators were fair and objective enough to conduct the evaluations, (b) did not feel that the evaluations were accurate as they did not necessarily measure real student academic gains and achievement levels of the students, (c) did feel undue stress and anxiety, (d) did feel negative impacts to the evaluation when students are uncooperative and unfocused, (e) did not accept the evaluation standards being tied into performance pay, and (f) did not understand the evaluation system very well. As a result, these teachers had negative perceptions about the evaluation and did not support
continuing it in its present form. The results of this study provide a basis for mutual communication and discussion among the teachers, the teachers’ union, and the school district to revise and improve the teacher evaluation process.

*Keywords:* evaluation, teacher, performance pay
Acknowledgements and Dedication

While there are many people to thank on this journey, I want to especially express my deepest appreciation and gratitude to my lead adviser, Dr. Billye Sankofa Waters. Not only is she professional and objective, but she is also a giving and kind individual who has kept me on track and sane throughout this process. I could not have made it to this point without her guidance and support, and for that, I am forever grateful.

I would like to thank Dr. Karen Reiss Medwed, my second reader, who has consistently offered her guidance and knowledge. She provided her time and attention to my quest and I will not forget her professionalism. In addition, I want to give a huge shout-out to Dr. Valerie Taylor, who came in as my interim adviser at the end of this hard-fought trek. Not only did she “hit the ground running,” but she provided another voice to get me to the end of this voyage.

I am grateful to the teachers of the Broward County Public School who work tirelessly, day in and day out, and provide the best that they have to support student growth and achievement, both in and out of the classroom, to the students of the sixth largest school district in the country.

I have the greatest family any man could ask for (Dad, Jill, Dana, Melisa, Shai, Jake, Sydney, and Marilyn). Your sheer support, faith, and love helped to transform my dream of becoming a doctor into reality. If not for your shoulders to lean and cry on, I would not be here today. Your unconditional love and guidance got me through the rough spots and I am so fortunate to have all of you in my life. To my friends (you know who you are), thank you for your patience and understanding when I needed them the most. You made this adventure worthwhile.
Finally, I want to dedicate all who I am along with the completion of this long journey to my beautiful mother, Rhoda Satty, who would have been behind me every step of the way throughout this process. She is missed every day, yet she is always in my heart and thoughts. After God made you, Mom, He broke the mold!!!!!!
Table of Contents

**Chapter 1 - Statement of the Problem**
- Topic .............................................................................................................. 8
- Research Problem .......................................................................................... 8
  - Justification for the Research Problem ....................................................... 10
  - Deficiencies in the Evidence ...................................................................... 11
  - Relating the Discussion to Audiences ......................................................... 11
- Significance of the Research Problem ............................................................ 12
- Positionality Statement .................................................................................. 12
- Research Question .......................................................................................... 16
- Theoretical Framework .................................................................................... 17
  - Agency in Education and the Role of Teachers ........................................ 18
- Summary and Conclusion ............................................................................ 20

**Chapter 2 – Literature Review**
- Introduction .................................................................................................... 20
- Race to the Top Initiative ............................................................................. 22
- Performance Evaluations ............................................................................. 27
- Factors of Evaluation ..................................................................................... 33
  - Self-efficacy .................................................................................................. 33
  - Correlation between teacher self-efficacy and teacher evaluations .......... 35
  - Evaluator objectivity in evaluating faculty .................................................. 39
  - Teacher stress and the impact of evaluations and performance pay ......... 42
  - Linking teacher evaluation to professional development (Innovation) .... 45
  - Relationship of classroom management, behavior, and time-on-task with evaluations ...................................................................... 49
- Pay for Performance – History and Background .......................................... 51
  - Performance Pay: For and Against .............................................................. 53
  - Experimentation of Pay for Performance .................................................... 56
- Summary ........................................................................................................ 58

**Chapter 3 - Qualitative Design**
- Methodology .................................................................................................. 60
- Research Design ............................................................................................. 61
- Research Tradition .......................................................................................... 62
- Participants and Context .................................................................................. 65
- Recruitment and Access .................................................................................. 67
- Data Collection ................................................................................................. 67
- Data Storage .................................................................................................... 68
- Data Analysis ..................................................................................................... 69
- Validity, Trustworthiness, and Reliability ...................................................... 72
- Coding ............................................................................................................... 75
- Reciprocity and Protection of Human Subjects and Ethical Considerations ...................................................................................... 75
- Limitations ....................................................................................................... 78

**Chapter 4 - Analysis of Data and Research Results**
Introduction and Approach ........................................................................................................... 79
Participants’ Profile ..................................................................................................................... 79
Themes ........................................................................................................................................ 81
  Theme 1 ..................................................................................................................................... 81
  Theme 2 ..................................................................................................................................... 93
  Theme 3 ..................................................................................................................................... 100
  Theme 4 ..................................................................................................................................... 107
Summary ...................................................................................................................................... 117
Conclusion and Personal Findings ............................................................................................... 118

Chapter 5 - Discussion, Conclusion and Recommendations
Introduction .................................................................................................................................. 120
Discussion of Findings ................................................................................................................ 121
  Evaluation as a valid, equitable, and useful tool in determining performance pay ............. 121
  The evaluation tool affected by individual and independent teacher self-efficacy .......... 123
  Stress from teacher performance evaluations, and performance-related pay ................. 125
Other factors that teachers perceive to influence performance evaluations and
performance-related pay ............................................................................................................ 128
  Test scores matter .................................................................................................................... 128
  Student behaviors matter .......................................................................................................... 128
  Teacher collaboration matters .................................................................................................. 129
  Outside the classroom considerations ..................................................................................... 130
Conclusion .................................................................................................................................. 130
Implications of the Results for Practice ..................................................................................... 131
Implications for Policymakers and School Districts ................................................................. 133
Recommendations for Future Research ..................................................................................... 134
Personal Reflections and the Authentic Voice ........................................................................... 135

References .................................................................................................................................... 139

Appendices
  Appendix A: In-Depth Interview Protocol ............................................................................. 166
  Appendix B: Informed Letter of Consent ................................................................................ 168
  Appendix C: Invitation Letter to Interview Participants ......................................................... 170
  Appendix D: Principal Request Letter ..................................................................................... 171
  Appendix E: Study Participants ............................................................................................... 172
  Appendix F: Theme 1 Subthemes ............................................................................................. 173
  Appendix G: Frequency of Subthemes for Theme 1 ............................................................... 174
  Appendix H: Theme 2 Subthemes ............................................................................................. 175
  Appendix I: Frequency of Subthemes for Theme 2 ............................................................... 176
  Appendix J: Theme 3 Subthemes and Topics ......................................................................... 177
  Appendix K: Frequency of Subthemes and Topics for Theme 3 ........................................... 178
  Appendix L: Theme 4 Subthemes ............................................................................................. 179
  Appendix M: Frequency of Subthemes for Theme 4 ............................................................. 180
  Appendix N: Marzano’s iObservation Learning Map .............................................................. 181
Chapter 1: Statement of the Problem

Topic

With the advent of No Child Left Behind (NCLB), signed into law in 2001, the federal government made student achievement and teacher quality a national priority (U.S. Department of Education, 2003). As a result, teaching became a major focus of reforms. State and federal policies have targeted teacher evaluation as a strategy for improving the quality of teaching. Such evaluation events typically are not associated with teacher learning, even though these processes are in fact designed to develop a teacher’s practice in addition to assessing performance for employment decisions (Danielson, 2008; Danielson & McGreal, 2000; Stronge, Richard, & Catano, 2008). Teacher evaluation systems best serve the goals of school improvement, student learning, and teacher growth when they balance accountability requirements with ongoing teacher learning (Ovando & Ramirez, 2007). For all of these purposes to be realized; however, the evaluator must value the outcomes as well as the bureaucratic requirements (Erickson, 2014). The teacher evaluator, most often a school administrator or assistant principal, must view the evaluation events as opportunities for improving the quality of teaching and learning (D. R. Davis, Ellett, & Annunziata, 2002). Therefore, the principal’s attitude and actions regarding teacher evaluation, along with the capacity of the teacher’s classroom management, the teacher’s knowledge of the curricula, and the teacher’s monitoring of student achievement and growth, link directly to how the teacher evaluation process is realized at a school.

Research Problem

This study examined the teacher evaluation process in the Broward County (Florida) Public Schools (BCPS). That process is based on Robert Marzano’s teacher evaluation model
which can be considered an aggregation of the research on those elements that have traditionally been shown to correlate with student academic achievement. The model includes four domains: classroom strategies and behaviors, planning and preparing, reflecting on teaching, and collegiality and professionalism (Marzano, 2012). The model was developed by drawing on substantial research on the elements it addresses. The model was initially based on thousands of studies that span multiple decades, chronicled and catalogued in books that have been widely disseminated in the United States (Marzano, 2012).

Broward County, Florida, implemented the iObservation tool, part of the Marzano evaluation model, during the 2012-2013 school year. iObservation is an instructional and leadership improvement system. It collects, manages, and reports longitudinal data from classroom walkthroughs, teacher evaluations and teacher observations. Teacher growth and leadership practices inform professional development, differentiated to individual learning needs for every teacher and leader to increase his or her classroom effectiveness each year. Marzano (2012) stated, “iObservation supports districts as they establish a common understanding of effective classroom instruction; create a singular focus on enhancing teaching to increase student learning; develop trust with teachers through a focused and fair walkthrough, evaluation and observation process; engage in reflection and collaboration within virtual professional learning communities; save time and money by implementing effective professional development programs; connect teacher learning to student learning; help principals and administrators monitor and support effective teaching; and build capacity and implement for sustainability” (p. 17).

While the fundamental purpose of teacher evaluation is to provide feedback to improve the personal and pedagogical skills of professionals, a secondary purpose is to make summary
judgments about the overall quality of professional performance (Graf, 1998). Graf (1998) reports, “Teacher evaluation produces information about quality of instruction, contributes information about a teacher’s professional growth needs, and furnishes information regarding the teacher’s success at improving student outcomes” (p. 15).

**Justification for the research problem.** In March 2011, Florida lawmakers passed the “Student Success Act,” which introduced some of the most sweeping educational reforms in the state’s history. These reforms included the introduction of teacher evaluation systems based on value-added modeling, mandatory “performance pay” for teachers, and the elimination of long-term professional service contracts (Harrison & Cohen-Vogel, 2012). Harrison & Cohen-Vogel (2012) mentioned, “Under the new law, teachers who had been traditionally compensated through pay scales that rewarded advanced education or experience faced a new system in which performance-based evaluations drive salary increases. These new systems had to base at least 50 percent of a teacher’s evaluation on their value added to student achievement. In addition, Senate Bill (SB) 736 eliminated the state’s previous policy, which required that districts award tenure to teachers after 3 to 5 years of service by offering them long-term professional service contracts. Under the new law, teachers hired after July 1, 2011, were offered annual contracts that expired at the end of each school year, regardless of length of service” (p. 517-518).

Performance-based pay systems are a form of compensation reform, and, like any other institutional reform initiative, can fail for many reasons. The purpose of this research was to determine the relationship between professional teacher performance evaluations and performance pay based on the results, and if and how the evaluation process may lead to teacher stress. Teacher experiences with the new Florida evaluation system are identified so we can place a value on the significance of the effects that come from pay-for-performance.
**Deficiencies in the evidence.** School districts around the country are placing a greater emphasis on teacher performance evaluations, value-added measurements, and how the results relate to pay-for-performance (Liang & Akiba, 2015; Harrison & Cohen-Vogel, 2012). While there has been quite a bit of research conducted on teacher performance evaluations, value-added measurements, and performance pay independently, there is little research on how teacher evaluations and value-added measurements affect pay-for-performance outcomes. Therefore, there is still a need to research and review the factors that contribute to teacher perceptions about the merging of performance evaluation results with pay-for-performance (M.D. Jones, 2013; Woessman, 2012). This qualitative study provided a holistic approach to teachers’ impressions of their own performance evaluations and how the pressures of performance-based pay may affect teacher attitudes.

**Relating the discussion to audiences.** Examination into how teachers support the integration of performance pay with teacher performance evaluations yielded knowledge that is beneficial in furthering the understanding of teachers’ perceptions, and how district personnel, school-based administration, and teachers themselves can make the process more transparent and equitable. This study informed district and school-based stakeholders about the factors that affect teacher morale and teacher impressions regarding how the results of their performance evaluations affect their teacher pay. This research provided understanding and information about what needs to be accomplished to build teachers’ comfort with the combination of teacher evaluations and pay-for-performance. This study, through qualitative interviews, also added to the negligible body of literature about the merger of teacher performance evaluations with teacher performance pay. Finally, this study demonstrated how Bandura’s self-efficacy theory could be used to investigate how teachers’ belief in their own ability to complete tasks and reach
goals relates to their understanding and support of performance evaluations that are connected to performance pay.

**Significance of the Research Problem**

With Florida shifting from a tenure-based evaluation tool to a performance-based system, in which teachers are classified according to performance categories, it was important to understand how the expanded teacher evaluation ratings impacted teacher efficacy and agency, recognition, effectiveness, growth, remediation, and dismissal (Harrison & Cohen-Vogel, 2012). As Florida continued to tweak and expand its teacher performance rating categories, it was important to further analyze the intended and unintended impacts of this expansion, as well as educators’ perceptions about performance-based pay. This analysis increased understanding of teacher attitudes and perceptions about relating their appraisals to pay as the state of Florida continues to monitor its new evaluation system (Bullis, 2014).

Much has been written about the pros and cons of performance-based pay systems, including that many teachers’ raises and bonuses in some Florida counties are directly linked to how their students perform on standardized (high stakes) tests (Whoriskey, 2006). There are, however, few qualitative studies that use rich descriptions and qualitative tools to root out a deeper understanding of the true nature of teacher concerns about performance-based pay being tied to performance evaluations (Marsh, 2014).

This research helped teachers understand the beneficial qualities of the new evaluation tool as it relates to their professional commitments. This study provided insights into the teachers’ rich thoughts, asserted implications for practice, and helped provide alignment and avoid misalignment of the evaluation tool.

**Positionality Statement**
I have been working at Coral Glades High School for the past 10 years and in BCPS for 17 years, thus I am intimately familiar with the progression of teacher pay-for-performance and the implementation of the new evaluation tool to measure teacher effectiveness. Coral Glades High School is in a middle class neighborhood. It is in a diverse community, with the population being 43% white, 25% black, 25% Hispanic, 5% Asian, and 2% other; 47% of the student population is on free and reduced lunch and the school is designated Title I eligible. There are over 120 faculty members and most of them, regardless of their personal reactions to the Florida Accountability System, are concerned about the welfare of their students. As a school assistant principal, I am fully mindful that increased pressure to raise test scores is one of the more salient issues reported by my teachers, due to the effect that these tests and scores have on their own performance. As an example of one of the factors determining pay for performance, one teacher reported, “The MCAS (Massachusetts Comprehensive Assessment System) is one of the reasons I’m leaving, because of the pressure it puts on the kids and everyone else. And, we don’t even have the right materials to teach for it” (Dawson, 2012).

BCPS, the sixth largest school system in the nation, was chosen as the site for this study. BCPS has a culturally diverse student population of over 250,000 students. In addition, the school district employs over 14,000 teachers in grades K-12, adult centers, and vocational locations. The current student to teacher ratio of 18:1 is higher than the state average of 15:1. Presently, the district includes 33 high schools, 42 middle schools, 138 elementary schools, 16 adult/vocational schools, and 16 centers.

For the past 5 years, I have been through countywide workshops that trained and continue to train administrators on how to properly implement, observe, and judge teaching practices through the iObservation tool. During this time, I have seen its transformation from a systematic
evaluation tool to a new more objective, accountable tool that looks at measurable benchmarks and domains to observe teacher instructional practices and student learning and input. While the new tool is meant to identify effective teachers and reward them accordingly, more information is still needed to effectively link performance pay to the iObservation tool.

As a former teacher of middle school English, I am aware of how teacher evaluations influence the classroom teacher. In today’s schools and classrooms, teachers are “under the gun” to ensure student growth through learning objectives, increase student achievement, and teach to the test. All of these factors affect the teacher’s psyche and cause stress and burnout. Today, teachers not only reference the issues of everyday classroom teaching, such as discipline, attendance, and curriculum enforcement, but they also express that students’ performance was felt to be a direct reflection of teaching practices. Along with the stress of student performance, teachers experience additional stress when being evaluated in the classroom, realizing that the evaluation results may determine their salary (Jacob & Springer, 2008). According to Boyer (2005), “it is normal for teachers to experience stress caused by the pressures of high-stakes factors. Teachers in secondary schools who teach high-stakes tested courses experience stress associated with accountability and assessment in education (McDaniel, 2012, p.28). Even though teachers need feedback and objective assessment data in order to plan meaningful developmental activities, the stress and anxiety created by these evaluations is tenable. Stress is a psycho-physiological process that results from the interaction of the individual with the environment (Akhaq, Amjad, Memood, ul-Hassan, & Malik, 2010) and results in disturbances to the physiological, psychological and social systems, depending upon individual characteristics and psychological processes.
As the researcher and practitioner for this study, I was very mindful that there were inherent biases. Despite occasional reservations about the new evaluation tool not living up to its promise, my own bias when I began training, observing, and giving teacher feedback leaned toward inquisitive enthusiasm about Marzano and the iObservation evaluation tool. I traced this bias to my own experiences as a classroom teacher tasked with teaching all students using creativity, compassion, and empathy. I remembered allowing myself to stray from my lesson plans in order to revisit many lessons that were not initially understood by my students. Today, the curricula are so specific and the instructional focus calendar so inflexible that teaching can no longer allow for creativity and sidebars. In addition, the era of teaching to the test does not allow teachers to deviate from the district-mandated learning calendar.

Teachers remain the major force in educating our children and ensuring that their future endeavors are safe and successful. With this in mind, I supervise my 25 teachers with respect, trust, professionalism, and compassion. There are so few opportunities to generate positive faculty and staff morale that it is incumbent upon me to alleviate some of the stress and burnout that occurs from teacher evaluations that could dictate teacher pay.

McCloy, a principal staff scientist for the Human Resources Research Organization in Louisville, Kentucky, noted that because feeling a lack of control is a major stress factor, it is predictable that teachers would be experiencing high stress levels (Overman, 2013). McCloy, who has testing experience with several large-scale assessment programs, including the National Assessment of Educational Progress, said he recognizes why teachers find it unfair “for my salary and livelihood to depend on how my students do, when I only have so much control” over their performance (Overman, 2013). In addition, McCloy said “educational reform efforts have
left teachers feeling “stove piped to a particular type of behavior in the classroom” that is “foreign to the way they’ve been trained” (Overman, 2013, p.1).

Frances Banales, president of the Tucson Education Association, agreed. She reported that much teacher stress comes from “the idea that we are not doing what’s best for kids,” adding, “the anxiety can lead you to question the profession you are dedicated to” (Overman, 2013). “Many schools are ‘in improvement’ or near it. Principals are under pressure to improve and there’s a lot of fear,” Banales says (Overman, 2013).

McDaniel (2012) stated, “The stress associated with tying teacher evaluations to pay-for-performance potentially poses a great problem for education. Teachers abandoning the profession because of the stress of high-stakes factors being at forefront of education and accountability could be detrimental for educators and education. Addressing stress as it is associated with teacher evaluations and performance pay could benefit educators, administrators, and school leaders by providing them with resources, support systems, and positive initiatives for managing stress while teaching in a high-stakes environment” (p. 32-33).

As the administrator over several grade levels, I can relate to the faculty perspective that disruptive classroom behaviors have an effect on teacher performance evaluations. I monitor all student referrals from my teachers and I keep a record of the problematic students. Not only do disruptive students cause a distraction to the lesson and learning plans, but they may also hinder and distract the evaluator from objectively evaluating the effective elements and data points that are to be measured during the observation. The knowledge gained through this study was valuable to faculty and administrative members at the interviewees’ school who are continuously striving to advance the iObservation evaluation practices throughout the district.

Research Question
This qualitative study examined the perceptions and opinions of district educators who are being evaluated by the new evaluation model in Broward County, Florida. This information may inform the future success of this tool as the monitoring and refining process evolves through a fifth year. The specific research question is: What are teacher understandings and perceptions about the new Broward County, Florida, evaluation tool?

**Theoretical Framework**

The theoretical framework, also known as conceptual framework, describes the key assumptions, theories, constructs, or variables that define the operable issues of the study (Miles & Huberman, 1994). Understanding these topics provided grounding and meaning to the findings (Patton, 2002). In his book Bandura (1997) stated, “The interrelated nature of teachers’ gateway status as front-line instructors and their beliefs, training, and practices, all of which clearly implicate teachers in the school-change process as a matter of agency, or their sociocultural-mediated capacity to affect positive change in schools, speaks to the theoretical framework of teacher agency, related to self-efficacy, that guided this study” (p. 6).

**Agency.** While agency per se has been extensively theorized, Fuchs (2001) suggested that there has been a tendency in social research to either focus on an over-socialized, macro view of agency – thus ignoring the local and specific – or to concentrate on overly individualized notions of agency (Priestly, Edwards, Priestly, & Miller, 2012). Bandura (1997) described agency as something individuals possess based on their ability to process information from their environment and also engage in actions that shape this environment (Cofield, 2013). Others defined agency as the capacity of actors to “critically shape their response to problematic situations” (Biesta & Tedder, 2006, p. 11), or the “capacity for autonomous action…[independent] of the determining constraints of social culture” (Calhoun, cited in Biesta
According to Archer (2000), “agency has been seen as autonomy and causal efficacy. Such statements may be taken to suggest an overly individualistic view of agency, rooted in psychological views of human capacity, and indeed many writers have taken such a view” (Priestly, Edwards, Priestly, & Miller, 2012, p. 193).

Biesta and Tedder (2007) tended to focus on developing an ecological view of individual human agency. However, collective agency and the agency of human and non-human assemblages, in which action is not linked to conscious intention alone, have also been posited as important in social and educational theory (Fenwick & Edwards, 2010). Nevertheless, viewing agency in such terms helps us understand how humans are able to be reflective and creative, acting counter to societal constraints, and also how individuals are enabled and constrained by their social and material environments (Priestly, Edwards, Priestly, & Miller, 2012). Archer (2000) confers, “Thus, human agents are reflexive and creative and can act counter to societal constraints as well as with societal possibilities. As reflexive people, agents are influenced by, but not determined by, society” (p. 175).

**Agency in education and the role of teachers.** Theorists such as Bowles and Gintis (1976) and Bordieu (1977) applied understanding of human agency to the educational setting in part as a means to explain the persistence of social inequality (Cofield, 2013). Bowles and Gintis (1976) argued that schools operate to reproduce the necessary conditions for capitalism to thrive (Cofield, 2013). Teacher agency is typically viewed as a quality within educators, a matter of personal capacity to act (Priestly, Biesta, & Robinson, 2012), usually in response to stimuli within their pedagogical environment. It describes an educator who has both the ability and opportunity to act upon a set of circumstances that presents itself within that individual’s leadership, curricular or instructional roles. The educator described would then draw from
acquired knowledge and experience to intercede appropriately and effectively. Agency is increasingly rare in the educational world of prescriptive improvement, and the term is too “often utilized as a slogan to support school-based reform” (Priestley, Biesta & Robinson, 2012).

Most educators would probably agree that out of all of the professions, they feel that their voices have the least amount of power; theirs are the ones least heard of any profession when voicing desires, needs, and innovative ideas (Gerstein, 2013). Samuel A. Culbert, a professor in the Anderson School of Management at the University of California, Los Angeles, noted in the New York Times opinion piece How to Raise the Status of Teachers: Allow More Autonomy,

“The way to make stars out of teachers is to let teachers be stars, to let them be as innovative as they can be, to let them find the path that works best for them and their students. If they are allowed to search for the best answers, they’ll find them. Instead, we’re doing the opposite: we’re telling them that if they want to keep their jobs, they have to do what people who know so much less than they do about education tell them to do. They have to dance to some constantly changing, politically created tune that is guaranteed to leave them demoralized and their students floundering” (Gerstein, p. 1).

Biesta and Tedder (2006) posited, “The bottom line is that teachers need to reclaim their perceived and real teacher agency, voice, and empowerment. They need to develop a growth mindset that they can and do have agency in their profession. With all that is happening in the education profession today, it is important to remember that teachers have power to change the system. This power for change can be called “agency,” which is defined as the capacity of teachers to shape critically their responses to educational processes and practices” (p. 134). With all the external push from various sectors, ultimately teachers are the ones who can cut through
all of the cross-purposed mandates and transform their own process and practices to ensure the best educational experiences for their students.

Within the context of school reform and high stakes accountability, teachers may be exposed to a wide variety of strategies and initiatives that are deemed to improve instructional practice. Sweeping legislation and school reform initiatives rarely solicit teacher input regarding school, district, and/or state-based changes (Ruggles, 2009). Since teachers are often provided changes that are expected to be implemented by state-wide, district-wide, and school-wide administration as a result of school reform initiatives, an exploration of teacher agency, or the power to influence prescribed changes within the school reform context, is vital to understanding how teachers respond to change (Ruggles, 2009). By exploring teacher agency, we will be better informed about school reform and why many initiatives, including but not limited to pay for performance and teacher evaluation and accountability, meet a great deal of teacher resistance while some are embraced. If teachers are to become agents of change, as is strongly supported by contemporary curricular policy, then school managers must carefully consider the relational conditions through which teachers achieve agency, bearing in mind that a collaborative culture to strengthen agency is to a large extent dependent upon the nature and scope of relationships within the school (Priestly, Biesta, & Robinson, 2012).

**Summary and Conclusion**

This study is presented in five chapters. This first chapter provided a brief overview of the study, including my personal connection to the topic, as well as the research questions that guided the research. Chapter 2 reviewed the literature related to teacher evaluation and pay-for-performance including a historical perspective, the relationship of teacher evaluations to pay-for-performance, and the factors affected by legitimate concerns over the influences of performance
evaluations on pay-for-performance. Chapter 3 discussed the research design of the study, the site selection, the recruiting of participants, the data collection, the data analysis, and all ethical considerations. In chapter 4, we heard the voices of the teacher participants as part of a discussion of the research conducted. Chapter 5 provided the results and the analysis of data, with conclusions, recommendations for future research, and its significance for educational practice and preparation.

Chapter 2: Literature Review

Introduction

The purpose of this literature review was to identify and focus on teachers’ views, thoughts, ideas, concerns, questions, and perceptions of the impact of the teacher evaluation process on pay-for-performance, also known as performance pay. The literature review was designed to report on research that addressed the research question in order to dissect if and how there is a relationship between teacher performance evaluations and performance pay, and how that relationship affects those teachers.

In 2006, the United States government created a $600 million federal grant program called the Teacher Incentive Fund (TIF) to support projects that reform teacher compensation (Razo, 2014). Gratz (2009) posited, “The Teacher Incentive Fund was developed as a companion to No Child Left Behind and focuses on rewarding teachers and schools for closing the achievement gap, raising student achievement, and producing real results for all children” (p. 3428).

The five goals of the TIF grant were similar to the elements found in Denver’s ProComp program. These goals were to improve student achievement by improving teacher and principal effectiveness; tie teacher and principal compensation to increases in student achievement;
increase the number of effective teachers in hard-to-staff subjects and schools; create sustainable performance pay systems; and examine multiple approaches to providing teacher incentives (Humphrey, Gallagher, & Yee, 2012). TIF grants were awarded to 33 state educational agencies, 22 local education agencies, and 8 non-profit organizations beginning in 2006 (Humphrey, Gallagher, & Yee, 2012). Denver was awarded a TIF grant for its ProComp program. The TIF grant allowed awardees to expand an existing program and/or increase payouts for teachers (Razo, 2014). TIF grant award winners had to shift their teacher evaluation from personnel-driven reporting, with subjective ratings, to targeted instructional support and improvement systems, with numerous observations from trained administrators (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; McGuinn, 2012). These challenges resulted in mixed student achievement results from the first round of the TIF grant winners and encouraged change to the TIF structure. These changes included support for the implementation of data systems and minimum requirements for annual teacher observations (Razo, 2014). These modifications to TIF helped shape the largest federal school reform grant, Race to the Top. This grant was the cornerstone of the Obama administration’s education reform policy (Razo, 2014).

**Race to the Top Initiative**

In February 2009, President Barack Obama signed the American Recovery and Reinvestment Act of 2009 (ARRA); historic legislation designed to stimulate the economy, support job creation, and invest in critical sectors, including education (U.S. Department of Education, 2013). ARRA provided $4.35 billion for the Race to the Top fund, of which approximately $4 billion was used to fund comprehensive statewide reform grants under the Race to the Top program. In 2010, the U.S. Department of Education (2012) awarded Race to the Top Phase 1 and Phase 2 grants to 11 states and the District of Columbia. The Race to the
Top program is a competitive four-year grant program designed to encourage and reward states that are creating the conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, and improving high school graduation rates; and ensuring students are prepared for success in college and careers. In order to apply for Race to the Top funding, states had to meet four core educational reform areas. These are adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy; building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction; recruiting, developing, rewarding, and retaining effective teachers and principals; and turning around the lowest-achieving schools (U.S. Department of Education, 2009). As part of the Department’s commitment to supporting states as they implement ambitious reform agendas, the Department established the Implementation and Support Unit (ISU) in the Office of the Deputy Secretary to administer, among others, the Race to the Top program (U.S. Department of Education, 2013). The goal of the ISU was to provide assistance to states as they implement unprecedented and comprehensive reforms to improve student outcomes. Consistent with this goal, the Department has developed a Race to the Top program review process that not only addresses the Department’s responsibilities for fiscal and programmatic oversight, but is also designed to identify areas in which Race to the Top grantees need assistance and support to meet their goals. Specifically, the ISU works with Race to the Top grantees to differentiate support based on individual state needs, and helps states work with each other and with experts to achieve and sustain educational reforms that improve student outcomes (U.S. Department of Education, 2013). In partnership with the ISU, the Reform Support Network (RSN) offers collective and individualized technical
assistance and resources to Race to the Top grantees. The RSN’s purpose is to support Race to the Top grantees as they implement reforms in education policy and practice, learn from each other, and build their capacity to sustain these reforms (U.S. Department of Education, 2013).

As part of its education reform agenda, Florida set ambitious goals for students and educators in its Race to the Top application, including doubling the percentage of incoming high school freshmen who ultimately graduate from high school, go on to college, and achieve at least a year’s worth of college credit; cutting the achievement gap in half by 2015; and increasing the percentage of students scoring at or above proficient on the National Assessment of Educational Progress (NAEP) by 2015 to or beyond the performance levels of the highest-performing states (U.S. Department of Education, 2012a). Florida was supported in these efforts not only by the projects funded through its $700,000,000 Race to the Top grant, but also through its existing strategic plan. In December 2010, the Florida State Board of Education approved the Next Generation Pre-K-20 strategic plan to advance the state’s education reform efforts. The six strategic areas of the plan included strengthening foundational skills, improving college and career readiness, expanding opportunities for postsecondary degrees and certificates, improving the quality of teaching in the education system, improving K-12 educational choice options, and aligning resources to strategic goals (U.S. Department of Education, 2012a). In developing its Race to the Top plan, Florida carefully considered the best approach for aligning the six strategic areas and the Race to the Top core education reform areas in order to build upon its existing education agenda (U.S. Department of Education, 2012a).

Florida’s education reform agenda also included the passage of the Student Success Act (the Act) in March 2011, which mirrored many of the goals in the state’s strategic plan and Race to the Top application. The Act made the following changes: (a) established a comprehensive
evaluation system for teachers and principals based on multiple measures of effectiveness, which include primary emphases on student growth and observations of educator practice; (b) tied compensation to evaluation results beginning in school year (SY) 2014-2015; and (c) eliminated tenure except for those instructional personnel who already had a professional or continuing service contract (U.S. Department of Education, 2012a). The Act put into law many of the elements of the teacher and principal evaluations proposed in the state’s Race to the Top application.

Florida is using its strategic plan, its Race to the Top plan, and the Act to further its education reform agenda. The state continues to believe that the ambitious goals set for students and educators within these reform efforts will increase the academic achievement of its students.

Florida received a Race to the Top award in September 2010 as part of Phase 2 of the Race to the Top competition (U.S. Department of Education, 2013). In Year 1, the state made progress in implementing some of the projects outlined in its Race to the Top plan. These projects included assisting Local Educational Agencies (LEAs) in redesigning teacher and principal evaluation systems to incorporate multiple measures, including instructional practices and student growth; helping LEAs begin the transition to new Common Core State Standards (CCSS); launching the Local Systems Exchange that allows LEAs to share information on their Local Instructional Improvement Systems (LIIS); and engaging stakeholders through the creation and engagement of eight implementation committees (Department of Education, 2013). Despite progress in the areas discussed above, Florida had difficulty implementing other aspects of its Race to the Top grant, including executing the large number and scope of contracts associated with its plan. Contributing to significant delays in Year 1 were leadership changes, legal challenges, disparate vendor quality in some initial responses, and difficulties in hiring qualified

In Year 2, the state made progress in executing contracts and implementing activities, and almost all projects were on track with the state’s amended timelines (U.S. Department of Education, 2013). In Year 2, Florida implemented the CCSS, now the Florida Standards Assessment (FSA), in kindergarten. Training also began for teachers across all grade levels, with approximately 7,500 educators receiving training on implementation of the Florida Standards Assessment (U.S. Department of Education, 2013). The Florida Department of Education (FDOE) began work on updating the Teacher Standards Instructional Tool and the Student Tutorial to include FSA materials (U.S. Department of Education, 2013). The state is also working to develop formative assessments in mathematics and English language arts (ELA) that align to the FSA.

In Year 2, all LEAs in Florida received approval for their teacher and principal evaluation systems and began using these systems for evaluations in SY 2011-2012. Florida Department of Education supported LEA and institutions of higher education partners in launching job-embedded teacher and principal preparation programs, UTeach replication, and a recruitment program for minority teachers (U.S. Department of Education, 2013).

Florida’s plan was that implementation of the teacher and principal evaluation systems will continue in Year 3 and beyond. LEAs will have the option to revise their evaluation systems based on lessons learned in the first year of implementation (U.S. Department of Education, 2013). Florida will support efforts to turn around the lowest-achieving schools through the continued support of the STEM (science, technology, engineering, math) and reading coordinators and its ongoing focus on growing STEM focused career and technology education programs in low-achieving high schools (U.S. Department of Education, 2013).
Teachers in Broward County, Florida, like those in other public school districts across the country, historically were unenthusiastic about evaluation policies for a plethora of reasons. Administrators mainly developed such policies with no teacher input (Harrison & Vogel, 2012). Additionally, principals who at times may have allowed personal opinions to affect their judgment administered these evaluations (Harrison & Vogel, 2012). The evaluations reduced the complicated task of teaching to a simplistic numeric value and they were used to sort teachers into specific categories, which many teachers felt divided rather than supported them in their pursuit of the goal of educational excellence (Sullivan, 2012).

In addition, BCPS requires teachers to prepare annual growth plans. Within the portfolio, teachers must submit their annual growth plans for the upcoming year. Annual plans must be approved by the supervisor at the beginning of the year and address what the teacher will do to advance school and district learning goals. An end of the year evaluation of actual accomplishments is also required as part of the plan.

**Performance Evaluations**

With the national movement to improve educational results focusing especially on test scores, some suggested that another way to improve public education would be changing the way that teachers are evaluated. Many argue that the most influential school-based factor affecting student achievement is teacher quality (Rice, 2003). Some see teacher evaluations as an important part of measuring teacher quality and are unimpressed with the current teacher evaluation systems, especially since evaluations often neglect measures of instruction quality and student learning in a direct and serious way (Toch, 2008). An Economic Policy Institute study argued, “in practice, American public schools generally do a poor job of systematically developing and evaluating teachers” (Lynn, 2013). Evaluations often consist of checklists based
on short classroom visits, with some items on the checklist not directly addressing the quality of instruction. In addition to checklist items often being unsatisfactory, evaluators often give the same ratings to most of the teachers they evaluate and fail to discuss the evaluation results with them afterwards (Toch, 2008). While current teacher evaluations have many problems, education experts offer a variety of potential solutions. Choices include more detailed standards that focus on indicators of instruction and planning ability, evaluations based on teacher portfolios as well as observations, and multiple evaluations by multiple evaluators (Toch, 2008). Some of these systems focus more on improving instruction than on eliminating ineffective teachers (Toch, 2008).

One method of teacher evaluation uses student test scores to measure teacher performance. Three test-based accountability program approaches are most commonly used by school districts in the United States: status models, cohort-to-cohort change models, and value-added models (Koretz, 2008). No Child Left Behind (NCLB) combines two of these approaches to test-based accountability programs (Koretz, 2008). NCLB measures student scores of one group at one time against the state’s annual measurable objective (a type of status model), while also comparing the change in student scores over time and increasing the state’s annual measurable objective each year (a type of cohort-to-cohort change model) (Koretz, 2008). Koretz (2008) went on to say that some argue that tests are a “direct and simple” measurement tool. Proponents of using student test scores as part of measuring teacher performance assert that they provide an effective measure of how well students are learning, especially when value-added data is used to help account for factors teachers cannot control (Lynn, 2013).

Using student test scores to evaluate teachers has recently taken the form of a value-added model, which attempts to “incorporate information on the value-added by individual
teachers to the achievement of their students” (Lynn, 2013). Value-added models generally compare a student’s end-of-year test scores with test scores from the beginning of the school year, and adjust the difference in scores using other factors that may have affected a student's score (Lynn, 2013). These factors are sometimes based on student background or school-wide factors outside the teacher’s control (Lynn, 2013). Value-added models are still being developed, and different versions of these models are in use to measure the value a teacher adds each year (Koretz, 2008). Teachers and principals can gain valuable information about teachers’ strengths and weaknesses from value-added measures of student test scores and can compare teacher performance with other teachers outside their school (Lynn, 2013). Value-added models are generally believed to be a better way to measure performance than status models or cohort-to-cohort change models, but some argue that the information used by the model has limitations, making the model insufficient on its own as a measure of teacher performance (Koretz, 2008).

Some have suggested that student test scores should play a supporting role in evaluations, rather than the leading role many states have given them (Toch, 2008). These policy changes were partially triggered by Race to the Top. It has been proposed that reliance on student test scores when evaluating teachers be reduced, although student test scores—especially value-added models—could still be used as a tool to improve teacher performance by highlighting areas that may need more work (Lynn, 2013). Several U.S. states that have recently incorporated teacher evaluations of this kind; often, these policies resulted in conflicts (Lynn, 2013).

A Florida law passed in 2011 that ties teacher evaluations to student standardized test scores generated controversy and a lawsuit from teacher unions and individual teachers. Teachers narrowed in on the legal problems with Florida State Board of Education rulemaking procedures and substance.
For a number of reasons, Florida’s recently-passed law is an appropriate one to review in order to understand how this method of teacher evaluation may affect education systems and create ongoing litigation and conflict. The law imposed a value-added model of teacher evaluations based on student test scores and used its new teacher evaluation model as a way to apply for Race to the Top funds (Postal, 2011). This legislation also invited controversy and led to a litigation battle (Postal, 2011). The following section will detail how the litigation created conflict between teachers and the education reform movement, invalidated a rule, and caused uncertainty about future law and policy.

In 2011, the Florida legislature passed Senate Bill 736, codified in § 1012.34 of the Florida Statutes (Jordan, 2012). Entitled the “Student Success Act,” the law provoked a legal challenge from the statewide teachers’ union, the Florida Education Association (Jordan, 2012). The statute changed teacher evaluation requirements. Also, the bill reinforced Race to the Top, which requires that half of the evaluation for classroom teachers and other instructional personnel be based on the performance of students assigned to them over a three-year period (Lynn, 2013). If less than three years of data is available, this part of the evaluation still must account for not less than 40 percent of the evaluation (Lynn, 2013). The law states that the “Commissioner of Education would establish a learning growth model ... to measure the effectiveness of a classroom teacher ... based on what a student learns” (Fla. Stat. § 1012.34 (3)(a)2, 2012). This model would be used for the state assessments. The portion of the teacher evaluation measures that are not based on test scores are based on instructional practice (Fla. Stat. § 1012.34 (2)(e), 2012). Four evaluation levels are required under the law: “highly effective; effective; needs improvement or, for instructional personnel in the first [three] years of employment who need improvement, developing; and unsatisfactory” (Fla. Stat. § 1012.34(2)(e),
The Commissioner of Education is required to consult with experts, instructional personnel, school administrators, and education stakeholders in developing criteria for these levels (Fla. Stat. § 1012.34(2)(e), 2012).

Teachers evaluated as highly effective or effective will receive a pay raise under the system (Fla. Stat. § 1012.22(1)(c)5b, 2012). A school district will not renew a teacher’s annual contract if the teacher receives any of the following: two unsatisfactory annual ratings in a row, two unsatisfactory ratings in a three-year period, three needs improvement ratings in a row, or a combination of unsatisfactory and needs improvement evaluations in a three-year period (Fla. Stat. § 1012.33(3)(b), 2012). In addition, these performance evaluation ratings can subject a teacher to dismissal for cause. The law also requires that superintendent recommendations for promotions or transfers be primarily based on the employee’s effectiveness according to the standards of § 1012.34 (Fla. Stat. § 1012.22(1)(e), 2012).

Many Florida teachers reacted negatively to the new law and its requirements that they would have to follow, believing that some measures would not help further educational goals; some who agreed with its basic tenets thought it was implemented too quickly (Postal, 2012). The research organization that worked to create the policy also revealed concerns that some districts implemented the plan too quickly (Postal, 2012). Teachers unions challenged the most controversial part of the law in court—that which uses student test scores as part of the measure of teacher performance (Postal, 2012).

Another part of the Florida law that generated controversy is the nature of the formula that helps determine teacher pay (Isensee & Gonzalez, 2012). The formula predicts what a student should score on the state standardized test, the Florida Comprehensive Assessment Test (FCAT), or, starting in the 2014-15 school year, the Florida Standards Assessment (FSA), and
compares the actual score with the predicted score to evaluate teacher performance (Isensee & Gonzalez, 2012). While factors such as class size, attendance, and the students’ previous FCAT scores are included, socioeconomic status is not included in the formula, despite a known link between poverty and success in school (Isensee & Gonzalez, 2012). Other objections to the formula are its complexity and its reliance on variables that may not produce accurate results (Gonzalez, 2012). As one teacher commented, it is “only as good as the variables that you’re actually looking out for as well as the test that you’re using to measure” (Gonzalez, 2012). This raises concerns about the reliability of test scores as a general measurement of student success (Lynn, 2013).

In May 2014, a federal judge in Florida ruled that the state’s new teacher-evaluation system is legal, even though he was persuaded that it had been poorly implemented and was unfair (Sawchuk, 2014). The pressure of the decision was worrisome enough that lawmakers made some legislative fixes to the system. In an 18-page ruling issued on May 6, federal District Court Judge Mark Walker noted that the evaluation system stands to affect teacher’s pay, promotion, retention, and even reputation, given Florida newspapers’ penchant for publishing the value-added ratings (Sawchuk, 2014).

The suit, brought by seven teachers, charged that the law violated teachers’ constitutional equal-protection and due-process rights (Sawchuk, 2014). However, Judge Walker said that the classification of employees using test scores of students who were not in their classes was not done to discriminate, but was rather a “practical consequence” of the lack of uniform measures for assessing student-achievement growth across all subjects and grades (Sawchuk, 2014). The Florida Education Association (FEA) promises to continue the fight. FEA President Andy Ford said, “We will continue to point out this unfairness and we will continue to work to find an
evaluation system that is fair, open, and provides a sensible way to properly evaluate our public school teachers” (Sawchuk, 2014). Rulings regarding other value-added evaluation systems are pending in Tennessee and Texas.

**Factors of Evaluation**

**Self-efficacy.** For teachers, high levels of work stress may result in low morale, reduced effectiveness, higher levels of absenteeism, and reduced commitment to the profession (Hakanen, Bakker, & Schaufeli, 2006), potentially resulting in the decision to leave the profession (Jepson and Forrest, 2006). For beginning teachers, high levels of teaching stress may dissuade entry into the profession, lower occupational commitment, or encourage abandoning the profession soon after entry (Conley & You, 2009; Rots, Aelterman, Vlerick, & Vermeulen, 2007). Early success in managing teaching stress may be key to building confidence for teachers: “A strong sense of self-efficacy is developed through repeated successes, (whereas) occasional failures are unlikely to have much effect on judgments of one’s capabilities” (Bandura, 1997, p. 399). The relationship between teachers’ work stress and occupational commitment is important: Initial level of teacher commitment has been linked to decisions to leave the profession in the first five years of teaching (Rots et al., 2007) and to decisions to decline initial entry into the profession (Chaplain, 2008).

Self-efficacious teachers would perceive the objective demands of daily teaching as being less threatening than would those teachers who harbor self-doubts about their professional performance. Successful adaptation to stressful demands, in turn, would prevent the emergence of job burnout (Caprara, Barbaranelli, Borgogni, & Steca, 2003; Skaalvik & Skaalvik, 2007).

Schwarzer & Hallum (2008) explored the cross-sectional associations between teacher burnout and putative precursors, that is, self-efficacy and job stress. At the same time, job stress
was considered a result of low personal coping resources. In addition to the mere associations among the variables, they examined whether there is a mediation effect (self-efficacy $\rightarrow$ job stress $\rightarrow$ burnout) to gain more insight into possible mechanisms of the development of burnout and of protective factors. The study was guided by the assumption that self-efficacy constitutes a resource factor and that job stress may operate as mediator between self-efficacy and burnout (Schwarzer & Hallum, 2008). The unexpected absent relationship between self-efficacy in the classroom and causes of stress needed closer scrutiny (Helms-Lorenz, Slof, Vermue, & Canrinus, 2012).

Perceived self-efficacy is critically important to those who teach. Perceived self-efficacy is defined as the confidence in one’s coping abilities and the beliefs that people are capable of meeting the expectations of their jobs at designated levels of performance, which, in turn, influences their lives (Christian, 2011). Bandura (1994) asserted that beliefs about self-efficacy determine how individuals act, feel, think, and motivate themselves. Self-efficacy aims at the broad and stable sense of confidence that helps people deal with a variety of stressful things (Christian, 2011). Teachers who have a strong perception of self-efficacy will perform their teaching responsibilities more efficiently, believing in their own competence and effectiveness. In other research by Bandura (1994), in the general field of occupational stress, and, more specifically, in the study of teacher stress, he suggested that self-efficacy can be a mediating factor and an enhancement for human accomplishment and personal well being. In the workplace, confidence can help to keep stress levels low and intrinsic virtue in one’s job high (Christian, 2011). Those who doubt their own capabilities may view difficult tasks as personal threats while those with high levels of self-efficacy embrace them (Bandura, 1994).
Application of Bandura’s social-cognitive theory explains how important a teacher’s self-efficacy is in the classroom. If students perceive tentativeness or lack of belief in a teacher, it will undermine the students’ motivation to learn and belief in their teacher’s ability to teach (J. L. Smith, 2011). Teachers with high perceptions of self-efficacy are able to engage students and give them a concrete plan of attack for an assignment, which helps them believe in their efforts (J. L. Smith, 2011). Current research indicates that teachers with perceptions of high self-efficacy can strengthen struggling learners’ beliefs in their academic abilities and increase their willingness to engage in academic tasks (Margolis & McCabe, 2006). Research done on the importance of high perceived self-efficacy and teacher effectiveness; therefore, confirms a need to determine the relationship between perceived self-efficacy and personal stress among teachers who are mandated to administer the high-stakes test (J. L. Smith, 2011).

In an effort to further the study of teacher efficacy, Gibson and Dembo (1984) developed the Teacher Efficacy Scale (TES). Henson, Kogan, & Vacha-Haase (2001) wrote that the TES was the first major attempt to empirically develop a data collection instrument to tap into this potentially powerful variable in teachers. The outcome of Gibson and Dembo’s study was a 16-item instrument (reduced from 30 items) in 6-point Likert scales consisting of two essentially uncorrelated subscales. The TES subsequently became the predominate instrument in the study of teacher efficacy, leading Ross (1994, p. 382) to label it a “standard” instrument in the field. Largely using the TES, researchers have linked teacher efficacy to multiple positive variables in teaching effectiveness as well as to positive student outcomes, including achievement variables.

**Correlation between teacher self-efficacy and teacher evaluations.** According to Bandura’s (1977) social cognitive theory, self-efficacy, defined as “people's beliefs about their capacities to produce designated levels of performance and exercise influence over events that
affect their lives,” represents one of the most important predictors of human motivation (Bandura, 1994, p. 71). Researchers have found a correlation between teachers’ self-efficacy and the teacher evaluation instrument (Pisciotta, 2014). Pisciotta said that Brock (2005) found that “tenured teachers believed classroom changes occur by reflecting on their own teaching practices and not on the evaluation process” (p.13). Pendergast, Garvis, and Keogh (2011) suggested that an initial decline in self-efficacy was evident for novice teachers due to their over-estimated anticipation of teaching readiness. Uzor (2005) suggested that the California teaching evaluation process did not improve teacher performance; however, other factors including collegiality, collaboration, and teacher self-efficacy had improved practices as well as student achievement scores. Tan (2012) found that teachers’ self-efficacy is influenced by the academic performance of their students. Shaughnessy (2004) said, “Perfect curricula or perfect teaching strategies would never happen. On the other hand, teachers who had high self-efficacy were open-minded, set high goals, and were willing to try new teaching strategies were the ones who were likely to have successful students” (p. 162). Keogh, Garvis, Pendergast, and Diamond (2012) believed that teachers’ effectiveness with students depended on their personal teaching beliefs and attitudes about their own effectiveness. Teachers who showed high self-efficacy saw themselves as being more efficient at teaching and educating, as well as making a greater contribution to the overall purpose (Calik, Sezgin, Kavgaci, & Cagatay Kilinc, 2012). With these attitudes and beliefs within themselves, teachers may see their actions as being capable of achieving the effectiveness needed to be successful (Pisciotta, 2014).

The role of self-efficacy in instruction has been explored with respect to teachers’ beliefs in their ability to promote learning and engagement in their students (Tschannen-Moran & McMaster, 2009; Tschannen-Moran & Woolfolk Hoy, 2001; Woolfolk Hoy, Hoy, & Davis,
Influenced by Bandura’s socio-cognitive model and self-efficacy research with teachers, Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) defined teachers’ self-efficacy as beliefs regarding one’s ability to teach, to regulate classroom behavior, and to motivate students to learn (see also Tschannen-Moran & Woolfolk Hoy, 2001). As outlined in their comprehensive model, the effects of teacher self-efficacy, motivation, and mastery experiences are a cyclical process with the core assertion being that higher self-efficacy should lead to better instruction due to self-efficacious teachers being more willing to invest effort in their teaching, thereby creating mastery experiences that further bolster their self-efficacy (Tschannen-Moran & Woolfolk Hoy, 2001).

As such, Tschannen-Moran et al.’s (1998) model contributes two critical ideas underlying the present research. First, it asserts that teacher self-efficacy is best evaluated with respect to three underlying components: instructional strategies, classroom management, and student engagement (Wang, Hall, and Rahimi, 2015). Self-efficacy regarding instructional strategies concerns teachers’ beliefs about their ability to effectively use various teaching strategies, whereas self-efficacy concerning classroom management refers to beliefs concerning one’s ability to regulate students’ behavior during class (Wang et al., 2015). In contrast, self-efficacy regarding student engagement reflects teachers’ beliefs in their ability to motivate students to value and actively participate in the learning process (Wang et al., 2015). Secondly, this model is consistent with social learning theory in suggesting that higher self-efficacy should lead to greater effort and higher expectations for success, once again implicating teachers’ beliefs in personally controllable contributors to teaching effectiveness (e.g., effort) as a proximal consequence of high self-efficacy that, in turn, contributes to teacher development (Wang et al., 2015).
With respect to interactions with students, teachers with higher self-efficacy tend to be more patient, make better use of class time, criticize students less, encourage student autonomy and responsibility, and persist longer when dealing with challenging students (Gibson & Dembo, 1984). For example, a study by Woolfolk, Rosoff, and Hoy (1990) with language teachers in the United States found that higher levels of self-efficacy predict more autonomy-supportive and less controlling behavior with students; less self-efficacious teachers were more authoritarian in their classrooms (Wang et al., 2015). Additionally, teacher self-efficacy has been shown to correspond with classroom practices, as evidenced by studies showing that teachers with high self-efficacy use effective teaching strategies more frequently (Chacon, 2005), are more willing to implement innovative teaching methods (Guskey, 1988), and focus on student collaboration and interaction as opposed to drill and practice methods (Woolfolk et al., 1990). Finally, students have also been found to have better academic performance with self-efficacious teachers (Abernathy-Dyer, Ortlieb, & Cheek, 2013).

With respect to the effects of teacher self-efficacy on psychological adjustment, recent findings further indicate that one’s objective teaching ability does not predict job satisfaction directly, but rather that perceptions of teaching-related self-efficacy lead to greater positive affect and job satisfaction (Klassen & Chiu, 2010). These findings are consistent with North American teachers in showing that self-efficacious instructors take greater responsibility for their teaching and are more willing to commit to their teaching career (Coladarci, 1992). When a teacher performs with high self-efficacy, the resulting product may result in successful learning for students. With this attitude and belief within themselves, teachers may see their actions as being capable of achieving the effectiveness needed to be successful with their performance.
Evaluator objectivity in evaluating faculty. Without a doubt, a teacher’s final evaluation rating is related to a school’s accountability. This is a critical area of decision-making for administrators because of the direct influences that teachers have on the outcomes of student achievement. According to the New Teacher Project (2011), “evaluations should give schools information to build the strongest possible instructional teams, and help districts hold school leaders accountable for supporting teaching development” (p. 47). Through purposeful evaluations, administrators can help those teachers in need and recognize those who are excelling. “It is when administrators make the decision to perform perfunctory, unrealistic, generally satisfactory evaluations for most teachers that accountability falls victim” (New Teacher Project, 2011 p. 47).

Administrators must make good decisions that enhance their teacher evaluation policies in order to extract the most potential from each teacher, who then, in turn, will cultivate student achievement (Kerrn, 2015). With teachers affecting student achievement, and student achievement being crucial to reaching a school’s accountability standards, administrators need to understand that their decisions about teacher evaluations and accountability for their schools are linked (Kerrn, 2015). W. Lewis and Young (2013) stated, “student achievement is the primary gauge of how effective teachers are within schools” (p. 191). This assertion means that if administrative decisions are made to “water-down” evaluations, or provide meaningless feedback within the evaluations to their teachers, they are ultimately affecting the accountability of their school (Kerrn, 2015).

Administrators must use teacher evaluation systems correctly in order to be effective and hold teachers accountable within the system. Administrators with insufficient training or who cannot provide effective feedback to teachers may be ultimately responsible for inconsistent
instruction to students (Weisberg, Sexton, Mulhern, & Keeling 2009). Investments in administrator capacity to provide specific and impactful feedback are necessary. These investments include extensive training in maintaining objectivity and providing targeted feedback, efforts to keep the evaluation process manageable, and professional development in effective assessment of instructional quality (Crystal, 2014). Teachers are impervious to feedback from a coach or administrator when they have conflicting definitions of quality. The single most essential thing that a school leader can do is to reach consensus with the staff about what constitutes teaching quality (Fisher & Frey, 2010). Administrators also must be able to practice the art of conducting difficult conversations with teachers (Donaldson & Peske, 2010).

Another likely obstacle to implementation of the new evaluation plan is teacher and administrator buy-in. Teachers must own change and understand that reform is an incremental process (Brown & Anfara, 2003). Reeves (2009) posited that the gap between what leaders claim to value, and what their actions show they truly value, is the greatest obstacle to cultural change. Leaders who do not fully believe in the intent of the value-added plan or who allow biases to impede the process may not attain the intended result of instructional improvement (Crystal, 2014).

There is an important aspect of evaluation and respect for objectivity on the part of the administrator. Research suggests that teacher feedback (McLaughlin & Pfeifer, 1988; Stiggins & Duke, 1988) and fairness (Keeping & Levy, 2000; Milanowski & Henneman, 2001) are associated with valuation acceptances and success (Kimball, 2002). Feedback is an important aspect of evaluation that has been shown to relate to teacher perceptions of evaluation quality (Kimball, 2002). Among a number of variables studied by Stiggins and Duke (1988), certain attributes of feedback were identified as having the highest correlation with teachers’ perceptions
of evaluation quality (Kimball, 2002). These attributes included perceived evaluator credibility, quality of ideas, depth of information, persuasiveness of rationale for suggested changes, usefulness of suggestions, trustworthiness of the evaluator, and capacity to demonstrate needed changes (Stiggens & Duke, 1988). Similarly, McLaughlin & Pfeifer (1988) identified several important dimensions of feedback, including timeliness, specificity, credibility, and intent (Kimball, 2002).

Teachers generate their own understandings of their performance through self-reflection, which can be a powerful source of professional growth (McLaughlin & Pfeifer, 1988; P. L. Peterson & Comeaux, 1990). Specific feedback from a knowledgeable and trusted observer can help facilitate growth (Kimball, 2002). When feedback is tied to evidence and agreed-upon criteria, the opportunity for teacher reflection and improvement may be enhanced (Danielson & McGreal, 2000). Teachers and evaluators must have access to resources, such as training, and have the time to conduct evaluation activities, reflect and act on feedback, and document progress (Kimball, 2002).

Perceptions of fairness also influence employee satisfaction with performance appraisal (evaluation), attitudes of evaluation acceptance, motivation to improve, trust in supervisors, organizational commitment, and legal defensibility in evaluation decisions (Gilliland & Landon, 1998). Three central aspects of fairness are procedural fairness, interpersonal fairness, and outcome fairness (Gilliland & Landon, 1998). Procedural fairness can be fostered by employee opportunity to participate in decisions, consistency in all evaluation processes, ensuring that appraisals and feedback are job related and unbiased, and provision of a formal channel for employees to rebut evaluation decisions (Gilliland & Landon, 1998). Interpersonal fairness is established by keeping employees informed, offering timely and informative feedback, and
treating employees with courtesy, respect, and trust (Gilliland & Landon, 1998). Outcome fairness is generated by providing outcomes that are expected, and maintaining and communicating a formally structured incentive system (Gilliland & Landon, 1998).

Another tension that exists in teacher evaluation systems is between the increased burden generated and the goal of systematically obtaining comprehensive evidence to improve the likelihood of reliable and valid decisions (Kimball, 2002). Construct validity considers whether evaluation evidence is adequate to make performance inferences using the standards (Kimball, 2002). Reliability refers to consistency in evidence gathered across teachers and whether evaluators are consistent in judgments of performance (Kimball, 2002).

In Kimball’s (2002) study, teachers’ perceptions of the evaluation tool, particularly the nature of feedback and fairness, were enabling conditions that influenced the use of feedback. The evaluation systems used added more work for teachers and administrators and may have inhibited evaluation feedback, but despite the increased workload, teachers largely saw the system standards, procedures, and outcomes as fair (Kimball, 2002). As Kimball’s (2002) case study suggested, simply adopting a standards-based evaluation system is not enough to transform teacher evaluation (Kimball, 2002).

**Teacher stress and the impact of evaluations and performance pay.** Teachers’ work stress reflects the experience of unpleasant emotions as a result of teaching (Kyriacou, 2001). This is highly relevant not only to teachers but also to school administrators and policy makers, given that the profession of teaching has been labeled as highly stressful by many researchers (al-Fudail & Mellar, 2008; Collie, Shapka, & Perry, 2012; Kyriacou, 2001). In fact, various international studies have shown that up to one-third of teachers are stressed or extremely stressed (Borg & Riding, 1991; Geving, 2007; Thomas, Clark, & Lavery, 2003). In these
studies, many different causes of teacher stress have been cited; however, two types of stress that have consistently been mentioned in the literature are stress related to students’ behavior and discipline and stress related to workload (Borg & Riding, 1991; Chaplain, 2008; Klassen & Chiu, 2010). Indeed, research has shown that these two types of stress are associated with several negative outcomes for teachers, including increased burnout (McCarthy, Lambert, O’Donnell, & Melendres, 2009) and reduced sense of teaching efficacy (Klassen & Chiu, 2010), job satisfaction (Klassen & Chiu, 2010), and commitment (Klassen & Chiu, 2011).

Researchers on work stress appear to agree that job stress is a serious problem in schools (Cooper & Cartwright, 1994). Concern over the mounting evidence linking work stress to negative individual and organizational outcomes has prompted an interest in identifying and developing strategies to prevent, control, and manage stress and its consequences (Murphy, 1995).

One work practice that may lead to stress is the annual performance review. This may result in some form of stress because both the appraisee and the appraiser are confronted with opportunities, constraints, and demands (Gbadamosi & Ross, 2012). For the appraisee, a good performance review may lead to a promotion, greater responsibilities, and higher salaries through a merit-pay program, while a poor review may prevent the teacher from being promoted and, if extremely poor, may result in job loss or job reprisal (Gbadamosi & Ross, 2012). Many studies have found that teachers report an increase in their job stress after the implementation of incentive pay programs. For instance, 72% of Charlotte-Mecklenburg Schools teachers and 87% of Kentucky teachers reported that the award program put more job pressure and stress on them (Kelley, 1999). Nearly half of the responding teachers in the Adkins (2004) study reported they experienced increased stress as a result of the pay-for-performance program and performance
evaluations (Yuan et al., 2013). Half of the teachers who were required to participate in Denver’s ProComp reported increased stress (Wiley, Spindler, & Subert, 2010). However, given the lack of control group teachers’ responses, it is unknown whether the increased job stress was due solely to the incentive pay programs (Yuan et al., 2013).

Gbadamosi and Ross (2012) investigated to what extent these variables were predictive of perceived stress. A key finding of their study was the emergence of gender, core self-evaluation, and performance appraisal discomfort as important predictor variables of perceived stress. Gender and core self-evaluation emerged as stronger predictors of perceived stress than performance appraisal discomfort (Gbadamosi & Ross, 2012). The results implied that demographic and psychological traits (individual characteristics) were more significant predictors of perceived stress than discomfort with performance appraisal (Gbadamosi & Ross, 2012). This finding has important implications for education administrators seeking to develop stress reduction strategies. While some studies have recommended prioritizing the removal of organizational (school site) sources of stress over individually-focused interventions, this study implied that interventions that develop individuals’ core self evaluations may have significant impacts upon their perceptions of stress (Gbadamosi & Ross, 2012).

In the absence of studies that have investigated these relationships (perceived stress, performance appraisal discomfort, core self-evaluations, and gender) simultaneously in the literature, the preliminary findings provided potentially useful insights into some of the causes of teacher stress and anxiety (Gbadamosi & Ross, 2012). The results indicated that more attention may need to be placed and focused on the individual teacher’s psychological and demographic characteristics if his or her perceived stress and responses to performance appraisal are to be
predicted, and their possible negative impacts on work performance avoided (Gbadamosi & Ross, 2012).

Khan, Shah, Khan, and Gul (2012) posited and summarized, “teachers under stress from performance evaluations and performance pay cannot perform well. Job satisfaction and motivation levels are decreased and they show unwanted behaviors like absenteeism, mistakes during work, and violence at work. Furthermore, they have more health related physical and psychological complaints. Their students’ satisfaction level is also decreased because they cannot impart quality instruction to the students. As a resultant, complaints come from both parents and employers, thus the overall image of the educational institution gets damaged. During the development of stress among the teachers, certain factors act as moderators, which buffer the negative effects of stress on the teachers’ performance. These factors can be termed as teachers’ resources. They include both personal as well as job related resources. The teachers’ resources minimize the negative effects of stress by giving strength and support to the teachers in fighting stress. Thus the teachers’ resources have vital importance in the management of teachers’ stress and performance” (p. 25).

Educational institutions should create general awareness among teachers regarding the existence of job stress and its related negative consequences. Furthermore, they should provide suitable job resources to teachers in such a way that it could help in fighting stress and increasing performance. Apart from job resources, the teachers should also try to utilize their personal resources for managing their job related stress and performance. This should be done in such a way that there is a balance between the teachers’ resources utilized and the stress dealt with.

**Linking teacher evaluation to professional development (innovation).** The alignment of teacher evaluation results with professional growth opportunities can be considered in terms
of the collection and use of evidence. This begins with identifying sources of evidence that will be used for evaluating teachers. Once state and district leaders have determined those sources, they must identify instruments and create processes that will ensure that evidence is gathered with high standards of validity and reliability. Evaluating the impact of professional development in terms of schools having an awareness of the effectiveness of professional development undertaken by teachers has been reported as the weakest link in the professional development chain (Office for Standards in Education, 2006). This is despite it being described as “a learning tool that improves the quality of both the continuing professional development and the outcomes achieved” (Early & Porritt, 2010).

There are three key standards for high-quality and effective professional development, as outlined in the National Staff Development Council revised standards of 2001. These include professional development that is data-driven, evaluative, and research-based (Pritchett, 2010). Dufour, Eaker, and Dufour (2005) called this a focus on results. In professional learning communities, “Every teacher team participates in an ongoing process of identifying the current level of student achievement, establishing a goal to improve the current level, working together to achieve that goal, and providing periodic evidence of progress” (p. 40-41). They also emphasized that “Educators must begin to embrace data as a useful indicator of progress” (p. 41).

Data should be collected at various pre-determined intervals in order to evaluate student achievement and program effectiveness (Pritchett, 2010). Husby (2005) cautioned, “Although common assessment of student achievement relies on annually derived standardized scores, for the purposes of evaluating project effectiveness, learners should consider not only broad scale data, but also the spectrum of data sources” (p. 39).
With this being said, according to the article, *Schools as Learning Communities* (2004), data should be used to design the professional development systems that will be used to improve school achievement (Dufour, 2004). According to Diaz-Maggioli (2004), “traditional professional development plans were made by top administrators and consultants and then presented to teachers. Today, specific goals and priorities for student learning should drive the content of your school’s adult learning” (Tallerico, 2005, p. 20).

In evaluating and determining professional development opportunities, it is important that the method chosen be research-based (Harwell, 2003). A study by the American Research Association (2005) indicated that research-based professional development that is tied to curriculum materials that teachers use leads to better instruction and improved student learning, and then to potentially stronger teacher performance evaluations. Haller and Kleine (2001) pointed out that research relies on experience, evidence, and observation. As educators implement new practices, it is imperative that these things be considered.

Guskey (1992) noted that when it comes to maximizing the time, money, and energy spent on staff development opportunities, educators must familiarize themselves with a wide range of educational resources that show sufficient evidence that they can improve student learning and focus (Pritchett, 2010). For this reason, he pointed out that one must take a critical look and attitude toward research-based initiatives (Pritchett, 2010).

Pritchett (2010) concluded, “Principals should analyze multiple data forms to determine effective teaching practices by individual teachers, and document that data within the teacher evaluation tool. As current research suggests, evaluation should no longer be about what the teacher is doing, but it should focus on quantitative evidence about what the students are learning, including state achievement data and internal assessments of learning” (p. 97). Based
on this scientific data, evaluations should accurately reflect instructional effectiveness to
determine which probationary teachers earn tenure and which tenured teachers need professional
development to meet the quality teaching indicators (Marzano, 2007) and ultimately score
effective performance evaluations (Pritchett, 2010).

In Florida, classroom observations are a valuable part of a performance evaluation, as
well as one of the best sources of information to guide teacher professional growth (Goe,
Biggers, & Croft, 2012). With information from observations, teachers and observers can focus
on evidence of both positive and negative interactions with students and develop strategies to
expand the positive interactions and reduce the negative ones (Goe, et al., 2012). Engaging in
evidence-based conversations with observers encourages teachers and has long been used in the
“lesson study” model of teacher professional development (Rock & Wilson, 2005).

A benefit from an aligned professional development system and program, especially one
that is purposely transparent and includes teachers in every stage of development, is that it will
receive greater buy-in from teachers (Goe et al., 2012). Goe et al. (2012) state, “if teachers
believe a system will be used strictly for accountability, then it will be received as punitive and
less likely to be accepted. On the other hand, if teachers understand that the key role of the
evaluation system is to improve teaching and learning, they can take an active role in their own
development. If they understand what is expected of them, are able to spend time in professional
learning activities, and can see their own practice improve, they will experience the payoff of
their efforts and know that it was time well spent. As instruction improves as a whole, teaching
will elevate as a profession and teachers and schools will receive the trust, resources, and support
they need to sustain excellence in U.S. education. If teachers can see these possibilities, they
will carry the system forward and achieve the ultimate goal of improvements in teaching and learning” (p. 16-18).

**Relationship of classroom management, behavior, and time-on-task with evaluations.** The classroom environment is relevant to the instructional evaluation process in that the creation of a positive learning climate is directly under the teacher’s control (Nor, Pihie, & Ali, 2008; Phillips, Balan, & Manko, 2014). Classrooms are unique environments that offer a myriad of possibilities to foster improved practice and facilitate student learning (Meyer & Turner, 2006). A trusting environment promotes risk taking, deep learning, and engagement where process and not product is emphasized (Robinson & Kakela, 2006). Trust is an important component in relationship building; when teachers build relationships with their students they have fewer discipline problems (Gregory & Ripski, 2008). To promote trust, teachers show respect for all students, value diversity, and encourage imagination and creativity (Robinson & Kakela, 2006). Students’ emotions play a role in their learning, and the environment influences emotion (Phillips et al., 2014).

Lazarus’ relational-motivation-cognitive theory supports the importance of creating a safe, purposeful, and equitable learning environment. Lazarus defined emotions as “emerging through person-environment interactions that change over time and situations” (Meyer & Turner, 2006, p. 378). Contemporary theorists concur with the notion that emotions, which influence learning and behavior, are contextualized (Meyer & Turner, 2006). Teachers’ attitudes and behaviors definitely affect students’ perceptions, classroom discipline, and performance (Meyer & Turner, 2006). Thus, the classroom environment category and the classroom management piece are appropriate aspects for consideration when conducting teacher evaluation.

The management of problem behaviors or the “lack of discipline” has been identified by
the public as the most persistent and possibly the most troublesome issue facing schools (Elam, Rose, & Gallup, 1992; V. Jones, 1993). Increasingly, public school personnel are facing problem behaviors that occur more frequently and significantly affect staff and student safety. In addition to various classroom disruptions including student disrespect of the teacher, student disruption, lack of attention, and unruly play and behavior, these behaviors include physical assault, weapons possession, gang violence, substance abuse, and physical and sexual abuse (Cotton, 1990). Public schools are far from being a safe environment where teachers and students can focus their attention solely on the learning process.

Students’ problem behaviors in the classroom interfere with teaching and learning. The behaviors not only interrupt the students’ learning, but also cause distress and interrupt learning for the other classmates (Burton, 1993). Creating a classroom that is conducive to learning involves maintaining order. This is said to be a process of interactions between the teacher and the student (Mikesell, 1998). According to Witteman (1995), it is said, “within the classroom, a teacher may have up to 1,000 interactions each day. If one understands that a school day lasts approximately 6 hours and the school year is 180 days, by middle school, a student has completed 7,000 hours, and by high school graduation, over 13,000 hours or 2,250 days. Therefore, it is possible that students will have witnessed over 2,000,000 interactions by the time they graduate from high school” (p. 87).

Classroom teachers are under enormous pressure to meet the academic and behavioral needs of their students, fulfill state achievement requirements, and satisfy the demands of the No Child Left Behind (NCLB) Act (Vallaire-Thomas, Hicks, & Growe, 2011). High quality instructional time is essential to the success of all students in any grade. In order to maximize this instructional time, educators must minimize interruptions and distractions from the daily
routine (Vallaire-Thomas et al., 2011). When a student’s disruptive behavior repeatedly halts instruction, all of the students in that classroom suffer the consequences. In order to avoid a lack of student interest and negative classroom behaviors, the time-on-task approach has been implemented to ensure student compliance and achievement. Numerous studies show that the amount of time a pupil is engaged in a learning task directly affects achievement, and that teachers can be trained in classroom management practices that increase students’ engagement rate (Bossert, Dwyer, Rowan, & Lee, 1982). However, time-on-task has a curvilinear effect on learning and may be less important than “success rate” in producing student motivation and achievement. Time-on-task has relatively direct implications for both learning and behavior.

Time-on-task has long been recognized as an important contributor to academic success because learning is partly a function of the time spent engaged in a task. Individual differences in time-on-task contribute to individual differences in academic skills (Bloom, 1974). Time-on-task is also, by definition, a measure of the absence of behavior problems, whereas time-off-task is a risk factor for disruptiveness and conflict (Gest & Gest, 2005). Increases of time-on-task are expected to improve academic learning and to reduce a possible cause of conflicts with teachers and peers (Gest & Gest, 2005).

Because time-on-task and classroom discipline have major effects on teacher practices, the rethinking of teacher compensation is a first step to school and classroom reform. But restructuring the teaching job is the way to truly have an impact on the bottom line: student performance, strong teacher retention, and strength in the classroom.

**Pay for Performance – History and Background**

Attempts at restructuring teacher compensation can be traced back to the mid 1800s in England through experimentation with a system known as “payment by results” (Dillon, 2008).
The model relied on annual student exams, and the results determined how much money the government would give to the school (Dillon, 2008). Incorporated into its education system in 1862, England eventually abandoned the system more than three decades later due to charges it narrowed the curriculum and hindered teacher creativity (Dillon, 2008; Gleason, 2000).

Performance-based pay, also called merit pay, appeared in 1908 in Newton, Massachusetts, but did not last (Ryan, 2008). In the United States, the idea of a “performance contract” debuted on a larger scale in 1969, but the efforts were eventually abandoned due to concerns over fairness, objectivity, funding support, and poor results (Gleason, 2000). Despite these attempts to reform teacher pay, currently 96% of public schools, accounting for nearly 100% of teachers, use a single salary schedule (Podgursky, 2008).

New federal policies, such as the Race to the Top program put into place by the Obama administration, have signaled the process of abandoning more traditional compensation systems in education (Marsh, 2014). The rationale for this change is not new, but is steeped in the wisdom of national test scores that fall short of meeting our national expectations for excellence on the world stage and fail to adequately prepare our students for a college education (U.S. Department of Education, 2012a). The main premise of Race to the Top is to encourage a more stringent teaching of the subjects of mathematics and language arts (U.S. Department of Education, 2012a). These policies are not new and have been repeated for decades (Odden & Kelly, 2002). The push for performance-based pay incentives is founded in business as a strategy for improvement by attracting, hiring, and retaining the best of the best talent (Sturman, Trevor, Boudreau, & Gerhart, 2003).

Marsh (2014) posited that the perceptions of the teacher while moving from one compensation system to another are largely unresearched in a true qualitative design. Preparing
for a similar study, he investigated a series of databases only to find that the qualitative research on teacher perception of performance-based pay is weak at best (Marsh, 2014). The results concerning performance-based pay yielded just over 100 studies, and when the field was narrowed further, by filtering for teacher perceptions, the results dropped to zero (Marsh, 2014).

The absence in literature, combined with the current state and national direction being taken towards compensating teachers based on the outcomes of their evaluation performances, provided the rationale for this study. The research does indicate that the attitudes of teachers about performance-based pay should be taken into account before any move within a district is attempted (Troman, Jeffrey, & Raggle, 2007).

**Performance pay: For and against.** Recent studies have yielded important findings on teacher attitudes toward performance pay. For example, younger teachers are more receptive toward performance pay than older ones (Goldhaber, 2008). Teachers of disadvantaged and low-achieving students are also more inclined to support pay for performance (Goldhaber, 2008), as are Black and Hispanic teachers (Jacob & Springer, 2007). About 50 percent of public school teachers surveyed in 2003 supported a move away from the single salary schedule (Goldhaber, 2008). Breaking those numbers down into subgroups revealed that elementary teachers are less supportive of pay for performance than secondary teachers (Goldhaber, 2008), and veteran and female teachers are also less supportive of pay reform in education (Jacob & Springer, 2007).

Teachers with positive views of their principals and administrators and negative impressions of other teachers in their school are more likely to support performance pay programs for highly effective teachers (Jacob & Springer, 2007). In a 2007 survey, when asked what pay structures were favored more than others, 72% of teachers favored “combat pay” (added pay for hard-to-staff schools), 47% favored incentives for National Board certification,
41% favored added pay for hard-to-staff subject and/or content areas, and only 17% favored merit pay (Goldhaber, 2008). Many other studies found that teachers favored additional pay for responsibilities, for teaching in low-performing schools, and for receiving outstanding evaluations. Teachers gave less favorable reactions to extra pay for teaching in hard-to-fill subjects or an incentive system based on student test scores (J. J. Lewis & Springer, 2008).

A study released in January 2007, using data from several districts in the United States, showed evidence that performance pay plans can help attract and retain quality teachers, which could lead to enhanced student performance over time (Gonring, Teske, & Jupp, 2007). Studies of comprehensive school reform show that reform is more likely to be successful when the plan has the support of the teachers, the principal and administration offers strong leadership, the district gives its full support, the school receives ongoing financial assistance, and the plan is implemented in small schools (Scherer, 2003). These findings indicate that certain factions of teachers might accept teacher compensation, in the form of performance pay, as a viable compensation model.

Performance pay programs often provide a financial incentive in the form of a cash bonus or salary increase whenever student test scores achieve a particular level or if performance appraisals are effective. These programs affect teacher behavior, which in turn affects student academic performance (M. D. Jones, 2013). In order for performance pay programs to affect student outcomes, there must be an upstream relationship between incentives and teacher behavior (M. D. Jones, 2013).

Proponents of performance pay argue that an effective incentive structure rewards teachers for their hard work and impact on student achievement (M. D. Jones, 2013). A few empirical studies have shown that teachers respond to these incentives by altering their work
effort under performance pay (M. D. Jones, 2013). Ahn (2013) found that teachers have fewer absences from work when the bonus outcome is in doubt, but no change in absences when there is either a low or high probability of receiving the bonus (Ahn, 2013). Another study reported that there was no increase in teacher absences or homework assignments (Glewwe, 2010). Glewwe (2010) found that the pedagogy did not change, although there was evidence that teachers conducted more test preparation sessions (Glewwe, 2010). In a performance pay program in Israel, Lavy (2009) found that teachers increased after school teaching and changed their teaching methods (M. D. Jones, 2013). Fryer (2011) found no effect of performance pay incentives on teacher absences or retention in New York City (M. D. Jones, 2013).

Opponents of performance pay argue that it discourages cooperation and encourages teachers to “teach to the test” (M. D. Jones, 2013). A leader of the United Teachers Los Angeles union wrote “teacher unions have historically resisted performance pay proposals because they undermine one of the core principles of teaching and learning: collaboration…as teachers we understand teaching is about working together to help students, not competition for better pay” (L.A. Teachers Sue, 2007, p. 6). Although few, if any, studies specifically examine teacher cooperation under performance pay, Jacob and Levitt (2003) found that teachers responded to high-stakes testing in Chicago public schools by altering student test scores. Neal and Schanzenbach (2010) found strong improvements in student test scores at the middle of the achievement distribution, but little changes at the ends of the distribution after the introduction of No Child Left Behind standards in Chicago (M. D. Jones, 2013). These findings suggest that teachers respond strategically to incentives, warranting a closer examination of cooperative behavior under performance pay (M. D. Jones, 2013).
**Experimentation of pay for performance.** While interest in performance-related pay programs is growing across the United States, there are several current programs that started at the district and state levels and have progressed through to national and federal-level initiatives. Today these programs are found in a number of major school systems, among them Denver and Houston, and states such as Idaho, Minnesota, Mississippi, Texas, and Florida. Public school experimentation with merit pay appears to be less controversial when whole schools are evaluated, and a bonus is awarded to entire groups of teachers (Goldhaber, 2010). South Carolina and Tennessee, for instance, have experimented with awarding schools additional funding based on the performance of their students (Clotfelter & Ladd, 1996). Although they differ from one another in some respects, both state systems compare school-wide gains in student scores over time to what statistical models predict the gains to be, after controlling for students’ demographic characteristics (Goldhaber, 2010). South Carolina’s “school-gain index,” for instance, was based on how well students in a particular school perform on both a state and a national standardized test relative to that same school’s performance in the previous year (Goldhaber, 2010). This index accounted for the socioeconomic status of students, which was based on the percentage of students receiving free or subsidized lunches (Goldhaber, 2010). Awards were shared among schools each year so there was no fixed sum of money received by outstanding schools; typical awards per school ranged from $15,000 to $20,000 (Goldhaber, 2010).

In contrast to South Carolina’s focus on measures of the gain in student performance, Kentucky set future proficiency standards and rewarded schools based on progress towards those goals over time, with more points awarded for higher-performing students. The awards in
Kentucky amounted to about $2,000 per teacher in eligible schools, or a total appropriation of $26 million (Odden & Kelley, 2002).

Experimentation with individually based merit pay plans, such as the widely publicized experiments in Denver and Houston, has occurred in a number of school systems over the last couple of decades (Hatry, Greiner, & Ashford, 1994). Although the Denver teacher pay plan has been touted as “merit pay,” it is really a hybrid system that awards teachers additional compensation based on four categories: knowledge and skills, demonstrated by holding an advanced degree; professional evaluation from administrators; market incentives for teaching in hard to staff positions or hard to serve schools; and various measures of student performance (Goldhaber, 2010). Excluding the incentives for an advanced degree, an individual teacher stands to gain upwards of $5,000 per year by meeting all the criteria (Wiley et al., 2010).

By contrast, the merit pay program adopted in January 2006 by the Houston Independent School District is designed to be a true merit plan, where teachers are to be evaluated based on student performance on the Texas state assessment (Goldhaber, 2010). As constructed, the Houston school system’s plan will pay annual bonuses worth as much as $3,000, with the intention to increase this amount to $10,000 within the next 5 years, to teachers whose students demonstrate strong gains on standardized tests (Cook, 2006).

Another well-known, long-standing performance pay plan is the one implemented in Denver, Colorado. In 1999, the Denver Classroom Teachers Association and the Denver Public Schools reached agreement on an alternative teacher plan that linked pay to student achievement and professional evaluations (Podgursky & Springer, 2007). ProComp included components of existing incentive pay including knowledge-and-skills along with hard-to-staff subjects and schools, but the most innovative portion of the plan was linking student achievement to the
teacher and incorporating teacher evaluations as outputs for performance pay (Goldhaber & Walsh, 2012). The pioneering ProComp model used student growth data in an attempt to accurately measure the value quality teachers add to a student’s education (Razo, 2014). The model was created to capture the inputs into student success and then measure and reward those outputs (Mihaly, McCaffrey, Staiger, & Lockwood, 2013). “ProComp is the first attempt to quantify the impact of a teacher during one year of instruction based upon a student’s previous academic standing and is defined as ‘value added’” (Razo, 2014, p. 23).

Denver’s was one of the first successful pay-for-performance programs able to link teacher evaluation scores and student achievement. The inclusion of hard-to-staff subjects and schools and knowledge-skills-based pay created the first comprehensive performance pay program (Razo, 2014).

Summary

This chapter summarized research that supports the relationship between teacher evaluations and teacher performance. The theories discussed provide an explanation for how teacher evaluations are affected by specific factors leading to performance pay rankings, and whether teachers perceive these factors to be fair and equitable. In the United States, the focus of teacher evaluation has shifted from personal characteristics in the early 20th century to observable teaching behaviors from the 1950s to the 1980s and to accountability, professional development, and school improvement from the 1980s into the 21st century. Teacher evaluation continues to be a major focus of school reforms for accountability, promotion, and staff development (Ellett & Teddlie, 2003; Teddlie, Stringfield, & Burdett, 2003).

The most common method of data collection for teacher evaluation is classroom observations conducted by principals or administrators with pre and post observation meetings
However, reliance on one or two formal principal/administrator observations presents serious problems because their teaching effectiveness, influenced by their lesson preparation, professional learning activities, and classroom management, may not be demonstrated during the administrators’ short classroom visits (Liang & Akiba, 2015).

Jensen, Yamashiro, & Tibbetts (2010) mention, “pay-for-performance is appearing as a strategy increasingly implemented at the district, state, and even national levels to address issues of teacher quality, recruitment, and retention in hopes that it will improve academic outcomes. Challenges related to teacher quality, attrition, and recruitment lay at the foundation of proposed reforms to traditional teacher compensation systems. While the potential for pay-for-performance systems to improve teacher quality is unclear, what is clear is that there is a perceived need to recognize and reward high-quality instruction” (p. 2-4).

As suggested in the research, increasing teacher satisfaction, recruitment, and retention is likely to require a multi-pronged approach involving the improvement of school-site conditions and professional development in conjunction with a financial reward system (Jensen et al., 2010). While the inclusion of different professional development and recruitment components may vary significantly, the one component that remains constant is pay. How best to change the way teachers are financially compensated is the most controversial piece of pay-for-performance models, as it inherently changes the long-standing single salary schedule that honors experience (Springer, 2009).
Chapter 3: Qualitative Design

Methodology

The purpose of this qualitative study was to determine teachers’ attitudes toward consequences of performance pay incentives based on their own performance evaluations. This study also sought to identify teacher characteristics and contributing factors that affect attitudes toward the relationship of performance evaluations to performance-related pay. Additionally, this study was created to help policymakers determine whether appropriate performance pay incentives should be related to teacher performance evaluations. This study described, interpreted, and analyzed the following question: What are teacher understandings and perceptions about the new Broward County, Florida, evaluation tool?

Gerhart, Trevor, and Graham (1996) reported, “Every pay program has its advantages and disadvantages. Programs differ in their sorting and incentive effects, their incentive intensity and risk, their use of behaviors versus results, and their emphasis on individual versus group measures of performance. Because of the limitations of any single pay program, organizations often elect to use a portfolio of programs, which may provide a means of reducing the risks of particular pay strategies while garnering most of their benefits” (p. 156).

In regard to performance pay, Hassel (2002) reported, “states, districts, and private funders should invest heavily in documenting alternate approaches to compensation and assessing their effects on teacher recruitment, retention, practices, and effectiveness in raising student achievement” (pp. 3-4). Teachers’ attitudes towards performance-based incentives in education need more documentation, analysis, and research. This study helped to further analyze teachers’ attitudes toward performance incentives and how their personal performance evaluations affect their pay and salary, thus furthering the literature and helping establish what
they think should be the basis for receiving performance rewards. Finally, this study took into account teacher backgrounds and school context variables that have rarely been included in discussion and research about performance incentives.

Interpretivism is a way to gain insights through discovering meanings, thereby improving the comprehension of the whole. Qualitative research explores the richness, depth, and complexity of phenomena. Qualitative research, broadly defined, means “any kind of research that produces findings not arrived at by means of statistical procedures or other means of quantification” (Strauss & Corbin, 1990). The underlying assumption of interpretivism is that the whole phenomenon needs to be examined in order to understand it. Interpretivism proposes that there are multiple realities, not single realities of phenomena, and that these realities can differ across time and place. This researcher used interpretivism to gain insight into how teachers perceive the world of high-stakes testing and its effects on teacher stress.

Research Design

The nature of this study dictated a qualitative research design. Qualitative research is by definition exploratory, and it is used when we do not know what to expect, how to define the problem, or need to develop an approach to the problem. It is also used to go deeper into issues of interest and explore nuances related to the problem at hand. A common data collection method used in qualitative research is in-depth interviews, the method chosen for this study. In order to delve into teachers’ perspectives on performance pay and evaluations, this qualitative research developed an initial understanding of the problem, looked for a range of ideas and feelings about tying performance pay to performance evaluations, projected the results to a larger population, identified evidence regarding cause-and-effect relationships, described
characteristics of relevant groups of people, tested specific hypotheses and examined specific relationships, and identified and sized market segments.

Research Tradition

An interpretative phenomenological analysis (IPA) design was selected to create an in-depth look at the effects of the new Broward County teacher evaluation on pay-for-performance. The emphasis within IPA research is on using a purposive homogenous sample rather than random or representative sampling to ensure the topic is relevant for the sample and can be explored in depth (Bramley & Eatough, 2005). The aim of IPA is to explore in detail how participants are making sense of their personal and social world, and the main currency for an IPA study is the meanings that particular experiences, events, and states hold for participants (J. A. Smith, 2011). J. A. Smith, Flowers, & Osborn (1997) said, “the approach is phenomenological in that it involves detailed examination of the participant’s life-world; it attempts to explore personal experience and is concerned with an individual’s personal perception or account of an object or event, as opposed to an attempt to produce an objective statement of the object or event itself” (p. 70). At the same time, IPA also emphasizes that the research exercise is a dynamic process with an active role for the researcher in that process (J. A. Smith, 2011). One is trying to get close to the participant’s personal world, to take, in Conrad’s (1987, p. 2) words, an “insider’s perspective,” but one cannot do this directly or completely (J. A. Smith et al., 1997). Smith et al. state, “access depends on, and is complicated by, the researcher’s own conceptions; indeed, these are required in order to make sense of that other personal world through a process of interpretive activity”. Thus a two-stage interpretation process, or double hermeneutic, is involved. The participants are trying to make sense of their
world; the researcher is trying to make sense of the participants trying to make sense of their world” (p. 74-76).

J. A. Smith, Flowers, and Larkin (2013) wrote that IPA is informed by three key positions: phenomenology, hermeneutics, and idiography. They went on to say that phenomenology describes the “what” and “how” of individuals’ experienced phenomena and develops descriptions of the essences of experiences, but does not explain or analyze those descriptions (Creswell, 2013). Hermeneutics is a theory of interpretation concerning textual meaning; as in the techniques used in speaking and writing that divulge the intentions and context of the speaker or writer (J. A. Smith et al., 2013). Finally, idiography relates to details and thorough analysis of small cases, which differs from mainstream psychological studies that are nomothetic in nature (J. A. Smith et al., 2013).

In regards to a phenomenological study, Creswell (2007) posited, “A phenomenological study describes the meaning for several individuals of their lived experiences of a concept or a phenomenon. Phenomenologists focus on describing what all participants have in common as they experience a phenomenon” (Creswell, 2007, pp. 57-58). Phenomenology describes the structures of experience as they present themselves to consciousness, without recourse to theory, deduction, or assumptions from other disciplines (Newberg-Long, 2010). Phenomenology and interpretative phenomenological analysis fit the purposes of this study by contributing insight into teachers’ lived experiences of the current environment, specifically the effects of performance pay based on the results of teacher performance evaluations, along with their beliefs about their role and responsibilities as teachers.

Phenomenology presents conscious experience as experienced from the subjective or first-person point of view and is a research technique that involves the careful description of
aspects of human life as they are lived (Newberg-Long, 2010). To provide insight into the world of teacher experiences and beliefs concerning performance pay and teacher performance evaluations, the phenomenological approach documented the experiences of teachers’ views on how these phenomena impact teaching.

Due to the complexities of teaching in the era of No Child Left Behind, this IPA study shed unique light on teacher perceptions through stories that communicated the experience of teachers pertaining to accountability and responsibility, stress, and their beliefs about teaching. Exploring teachers’ beliefs about their role in educating children and their perspective on the effects of teachers’ performance evaluations on performance pay informed us about both intended and unintended consequences of this practice on secondary school teachers.

Two prominent schools of thought within phenomenology exist. These are perspectives that come from the Utrecht school (Netherlands) and the Duquesne school (operating out of the Psychology Department at Duquesne University in Pittsburgh in the United States). The work of Max van Manen (1982, 1991) is said to have been “influenced by the Utrecht school, while Amadeo Giorgi is a proponent of phenomenological work that has come out of the Duquesne school. In contrast with van Manen’s work, Giorgi’s phenomenological research method has been categorized as “empirical phenomenological research in psychology” (Hein & Austin, 2001, p. 5). There are two schools of thought in the phenomenological realm. In hermeneutic phenomenology, researchers interpret human experience as though it were a text, and the outcomes of these studies are viewed as texts that offer rich and deep accounts of phenomena (Hein & Austin, 2001). The holistic and poetic approach (Meyer-Drawe, 1997) reveals a depth and insight into the human condition that is poetic in its sensitivity and reflection.
The aim of phenomenological psychology, following Giorgi, Fischer, and von Eckartsberg (1971), is to produce accurate descriptions of human experience. For this reason, phenomenologists operating within this tradition mainly utilize descriptions provided by others (obtained through interview or through written texts), although they can use their own experiences (Giorgi, 1985a). This is in contrast to hermeneutic phenomenology, which tends to use a much greater range and variety of data collection material (for instance fiction, poetry, and literature) (Hein & Austin, 2001). Ehrich (2005) posited, “in contrast with hermeneutic phenomenology, which uses less prescriptive methods of data analysis, phenomenological psychologists analyze the data utilizing a systematic and rigorous process. For example, Giorgi (1985b) outlined four key steps; these include reading the entire description to get a sense of the whole statement, discriminating meaning units within a psychological perspective, transforming the subject’s everyday expressions into psychological language, and synthesizing transformed meaning units into a consistent statement of the structure of the phenomenon. Specific statements are written for individual participants and a process of analysis is used whereby common themes across these statements are elicited and then used to form a general structural description that becomes the outcome of the research” (p. 8-9).

**Participants and Context**

The Broward County Public School district, the sixth largest in the United States, was selected for this study. It offered selection of elementary and secondary schools from which to choose seven teachers for this study. With approximately 250,000 students, this urban school district has over 200 schools (elementary, middle, and high). Specific school selections were made based on the willingness of the principals and teachers to participate in the study.
The district has a wide range of demographics, including schools that have mostly White middle to upper class student bodies, to those with nearly all Black students from low-income homes. The district’s strategic plan and reform initiatives guide the focus of district requirements for raising student achievement in its schools. Concerns exist about the district’s goal of increasing teacher accountability, including evaluation of every teacher every year, while using fewer resources and keeping teacher attrition to low levels.

The sampling strategy was purposive. Also known as judgment sampling, in this most common sampling technique, the researcher actively selects the most productive sample for answering the research question. This can involve developing a framework of the variables that might influence an individual’s contribution, based on the researcher’s practical knowledge of the research area, the available literature, and evidence from the study itself. This is a more intellectual strategy than the simple demographic stratification of epidemiological studies, though age, gender, and social class might be important variables. Homogeneous sampling was used in this study. This is the process of selecting a small homogeneous group of subjects or units for examination and analysis. Crabtree & Miller (1992) describe homogeneous sampling as “used when the goal of the research is to understand and describe a particular group in depth. In addition, participants may be able to recommend useful potential candidates for study resulting in a snowball sampling strategy as well. Snowballing is a method of expanding the sample by asking one informant or participant to recommend others for interviewing” (p. 253). Bailey (1996) and Greig and Taylor (1999) called those through whom entry is gained gatekeepers, and those persons who volunteer assistance key actors or key insiders.

This study documented the experiences of seven teachers’ views on the effects of teacher performance evaluations on performance pay. The reason for this number of participants in this
study is that, as within any research area, different participants can have diverse opinions. Qualitative samples must be large enough to assure that most or all of the perceptions that might be important are uncovered, but at the same time if the sample is too large data becomes repetitive and, eventually, superfluous. The researcher remained faithful to the principles of qualitative research where the sample size in the majority of qualitative studies should generally follow the concept of saturation. Creswell (1998) recommended having at least six participants in a phenomenological study.

**Recruitment and Access**

Recruiting the right participants is the foundation of effective research because the research results are only as good as the participants involved. Representative, well-spoken, and thoughtful research participants can provide invaluable feedback. Recruitment of participants consisted of contacting the school district to ask permission to recruit teachers for this study. Once permission was granted by the Superintendent’s office, the principals from the various high schools were contacted requesting permission to select teacher participants. Each principal was asked to provide the names of two or three potential candidates for the interview. Teachers were then contacted by phone and provided information regarding the nature of the study and invited to participate. There were no incentive awards offered to participants, as there were a number of opportunities to fulfill the sample size criteria.

Participants were informed of their anonymity and confidentiality during the entire process. The participants were informed of their ability to withdraw, without penalty, at any time during the interview process. Fairness and equity were maintained for all participants involved in this research.

**Data Collection**
Semi-structured in-depth interviews, no longer than 90 minutes each, were done with seven teachers who were teaching at the high school level. Interviews were scheduled at the convenience of the interviewees with the intent to sustain continuity of ideas and time and to thoughtfully identify areas for probing and in-depth understanding. Teachers with high school teaching experience were selected because of their experiences with the iObservation evaluation tool and the pay-for-performance structure in its current environment. Each participant had at least one year of teaching experience at the high school level in public schools. The years in public education allowed the researcher to receive information about the effects of performance pay through teacher evaluations, based on having experience in public school teaching before and after the institution of these specific protocols.

Individuals were chosen who could unequivocally make themselves available to the researcher to share their experiences and stories of teaching over time. In the world of public education, the administration often makes decisions that affect teaching in the classroom. The researcher used the actual voice of the teachers who are on the front lines teaching children on a daily basis. Open-ended interview questions provided the flexibility to allow the teachers to use stories and feelings to reflect on their experiences with high-stakes testing, their perceptions of their role as educators relative to district expectations, the environment of accountability, and their impact on students. These semi-structured in-depth interviews furnished meaningful, knowable, and explicit information that assisted in answering the research questions.

**Data Storage**

Data was qualified, quantified, and verified using recordings and notes from the interviews. The interviewer depended upon the recordings for details on quotes and much of the
data. Each interview was transcribed to allow for coding that would ultimately support the data analysis and results.

Centralizing the data or keeping them all together in one place was a key element of managing the data in an organized and systematic fashion. All physical data (notes, tape recordings, transcripts, etc.) were kept in a locked filing cabinet or equally secure location. All documents related to a given data collection event were placed in a secure location at each field site, in one large, heavy-duty archival envelope per event with an archival information sheet. The envelope’s contents included typed transcripts, expanded field notes, debriefing notes, handwritten versions of the notes, and possibly cassette tapes. Backup copies of tapes were stored separately. Eventually, both copies of the tapes will be destroyed for reasons of confidentiality. Once the researcher destroyed the tapes, it was necessary to put documentation of this in the archival envelope.

Data Analysis

Analyzing qualitative data is the activity of making sense of, interpreting, or theorizing data (Schwandt, 2001). Analysis begins with the process of organizing, reducing, and describing data; drawing conclusions or interpretations from the data follows. The entire process is intended to summarize the data collection process and give an account for what is actually happening. For this study, the data analysis process answered the questions:

- Do teachers perceive the new evaluation system as a valid, equitable, and useful tool for determining their performance-related pay?

- Does self-efficacy affect teachers’ perceptions of the new performance-based evaluation tool?
• Have teacher performance evaluations and performance-related pay caused stress on teachers?

• What other factors do teachers perceive to influence performance evaluations and performance-related pay?

In an interpretative phenomenological analysis, data must be analyzed using a particular method. Once all the relevant statements were taken, Moustakas’ (1994) data analysis steps were used to analyze the data. The researcher incorporated narrator voice into the conclusions as a means of preserving accuracy. Interpretive analysis was used to make comparisons, allowing for cross-case analysis (Newberg-Long, 2010). “Interpretive display of the narrator’s story requires the discernment of salient information to answer the research question(s) so that an excerpted narrative can be created to disclose the data in an accessible and evocative manner” (Mears, 2009, p. 71). Handwritten notations and audio-recorded interviews were transcribed and every significant statement was extracted and coded, and given equal value, in a process known as “putting that situation at a distance to better view it without assumptions or bias” (Mears, 2009, p. 71). Once all significant statements were extracted, duplicate statements were purged. Next, meanings were formulated from the significant statements. These meanings were formed as a result of reading, re-reading, and reflecting upon the original statements made by the participants. Next, collective meanings were formulated into clusters of themes. These clusters represented themes that materialized from the interviews and were common to all of the participants’ descriptions. NVivo, a coding software program was used to aid in the analysis process. Thesis and dissertation software programs are designed to help researchers by providing a system for storing and organizing files, searching for themes, crossing themes,
diagramming, and creating a template(s). This software program was used to create diagrams of themes in order to aid in the presentation of the data.

From the results of the analysis, a general description (textural description) of what was experienced as well as an exhaustive description of how (structural description) the phenomenon was experienced by the individuals in the study was presented. All findings were presented in a rich-descriptive narrative. Hence, teachers’ perceptions of whether or not high-stakes testing has an effect on teacher stress, as described by the participants, was explained and detailed.

“A well crafted narrative provides a richly tapestried account that enables a sort of vicarious participation so that appreciation for the circumstances becomes possible” (Mears, 2009, p. 123). According to Gubrium and Holstein (2002), “Narrative analysis allows for the systematic study of personal experience and meaning…Personal narratives provide windows into lives that confront the constraints of circumstances” (pp. 706-707).

The process of data analysis involves “making sense out of text and data…and preparing the data for analysis, conducting different analyses, moving deeper and deeper into understanding the data, representing the data, and making an interpretation of the larger meaning of the data” (Creswell, 2009, p. 183). The researcher looked for patterns, themes, and dimensions in the data through analysis of the interviews, coding of the data, and further analysis as themes and patterns emerged. The researcher’s goal was to describe the participants’ subjective experiences and their views.

The first level of identification occurred during the initial review of each interview transcript. Upon receiving the transcripts, the researcher read each transcript, analyzed the data for each interview, and then conducted open coding utilizing NVivo software, which is an analytic tool to facilitate the coding process.
The researcher used open coding, which utilizes a brainstorming technique described by Corbin and Strauss (2008) to “open up the data to all potentials and possibilities contained within them” (p. 160). In open coding, the researcher thoroughly reviews the data contained within the data set before beginning to group and label concepts. The process of coding involves taking the raw data and pulling out concepts, further categorizing data by their properties and dimensions, and then grouping data into themes. The data analysis process included the following steps:

1. Review all interview transcripts notes.
2. Import the data into NVivo.
3. Code the data in NVivo using open coding.
4. Define the properties of the dominant themes.
5. Create categories that represent primary themes and subthemes.

The resulting themes and subthemes are described in the summary of the research findings.

Validity, Trustworthiness, and Reliability

In qualitative research, trustworthiness or credibility corresponds to the concept of validity (Gall, Borg, & Gall, 1996), and design validity is described using a wide range of terms and perspectives (Creswell, 2007). According to Polkinghorne (1989), “trustworthiness in phenomenological research involves asking the following question: “Does the general structural description provide an accurate portrait of the common features and structural connections that are manifest in the examples collected?” (p. 57). As cited in Creswell (2007), Polkinghorne believed that validation of a concept is achieved when it is well-grounded and well-supported (Harris, 2012).

Angen (2000) argued that validation is a judgment about the trustworthiness or goodness of a piece of research. Additionally, Creswell (2007) noted “a researcher’s accuracy emerges by
spending extensive time in the field, transcribing precise and deeply detailed descriptions, and connecting or developing rapport with the participants” (p. 37). Conversely, Lincoln and Guba (1985) used a more naturalistic approach by stating that trustworthiness is a term that establishes credibility, authentic transferability, dependability, and confirmability as being equivalent to internal and external validation, reliability, and objectivity.

To ensure trustworthiness in this study, the researcher used the following strategies: triangulation; rich, thick descriptions; debriefing interviews; and member checking. As collaborative evidence, triangulation involves using multiple methods, sources, investigators, and theories for collecting data validly (Creswell, 2007; Miles & Huberman, 1994). The process allows the researcher to condense data into themes or codes to essentially illuminate the understanding of an experience (Creswell, 2007). Therefore, direct observation and semi-structured individual interviews were means for collecting data.

According to Creswell (2007), “presenting data with rich, thick description is an important way in which to provide credibility to research findings. Moreover, detailed descriptions allow the reader to determine if information is transferable to other settings. Therefore, the researcher presented rich descriptions of data through verbatim transcripts of individual teacher interviews. Further detailed notes pertaining to the participants’ specific interactions, behaviors, and any other events or activities were noted. Finally, descriptive details were included in the data interpretation, reporting, and results” (61-62).

In qualitative research, the researcher is known as the main instrument for data collection, analysis, and interpretation (Creswell, 2007; Onwuegbuzie, Leech, & Collins, 2008). Peer review or debriefing is a means of providing an external evaluation of the research process (Creswell, 2007; Lincoln & Guba, 1985). Being human, researchers have innate biases that must
be acknowledged and identified (Miles & Huberman, 1994). According to Onwuebuzie et al. (2008), bias occurs when a researcher has presumptions or personal prejudices that he or she is unable to bracket (Harris, 2012). Thus, bias can occur at any phase of the research process, including data collection, analysis, and interpretation (Harris, 2012).

Lincoln and Guba (1985) described the role of a peer reviewer as “a person who keeps the researcher honest by posing difficult questions concerning the methods, procedures, and interpretations of the research study. The peer reviewer may be a colleague or any person who has no interest or direct relationship to the research process” (p. 212).

As this research study focused on teacher stress, the debriefing process supported the validation and accuracy of the potential findings. A colleague who was familiar with interviewing skills, qualitative research, and willing to become familiar with the effects of high-stakes testing on teacher stress was selected as a peer reviewer. In addition, this peer reviewer read interview transcripts and participated in several debriefing interviews.

Member checking or informant feedback consists of soliciting participants’ views on the credibility of data interpretations and conclusions (Lincoln & Guba, 1985; Miles & Huberman, 1994). Lincoln and Guba (1985) asserted, “Member checking is the “most critical technique for establishing credibility” (p. 314). In order to eliminate any possibility of misinterpretation or misrepresentation, the researcher convened with the participants by asking them to reflect and judge the accuracy of the preliminary analysis of descriptions and themes drawn from the study (Harris, 2012).

Gay and Airasian (2000) cautioned, “Unbiased and perfectly valid data are nearly impossible to obtain in a qualitative study. The interpretive role of an investigator makes bias and invalid data serious concerns. To preserve the integrity of qualitative research, investigators
must address these concerns. Therefore, the researcher endeavored to preserve the integrity of this research study using the safeguards suggested by qualitative researchers” (Rubin & Rubin, 2005, p. 193).

The researcher did a continual check during the coding process to ensure that coding did not drift from its original intent as the coding process evolved. The researcher used an electronic codebook within NVivo to code the data. As only one researcher was responsible for analyzing the data, there was no need to cross check for intercoder agreement.

Coding

The coding process identified a total of 18 primary themes. The themes were delineated into four categories, with each category focusing on one of the four sub-research questions. The data gathered for this study were intended to answer the overarching research question: What are teacher understandings and perceptions about the new Broward County, Florida, evaluation tool?

This question was addressed via the exploration of four sub-questions: (a) do teachers perceive the new evaluation system as a valid, equitable, and useful tool for determining their performance-related pay; (b) does self-efficacy affect teachers’ perceptions of the new performance-based evaluation tool; (c) have teacher performance evaluations and performance-related pay caused stress on teachers; and (d) what other factors do teachers perceive to influence performance evaluations and performance-related pay? The findings for each sub-question were summarized and exemplars from the interviews were used to illustrate the themes and subthemes.

Reciprocity and Protection of Human Subjects and Ethical Considerations

Ethical considerations must be addressed prior to conducting research, as well as throughout the study. The ethical considerations regarding protection of the participants in the
phases of this study were anonymity and confidentiality that the researcher ensured were employed throughout the study. Informed consent, from persons capable of such consent, was obtained. This required informing participants about the overall purpose of the research and its main features, as well as of the risks and benefits of participation. Consent was given in written format, verbally and audiotaped, or videotaped. If the investigator did not know in advance the questions that a participant might be asked, or what potential risks might be involved in the future, this was made clear to the participant at the outset.

The investigator’s responsibility to the participants included issues such as ensuring confidentiality, avoidance of harm, reciprocity, and feedback. In ensuring confidentiality, an investigator may not report private data that identifies participants. One of the safest ways to ensure anonymity is not to record the names of the participants at all and to provide an information sheet that asks for verbal rather than signed consent. Categories of sensitive information requiring anonymity are the following: sexual attitudes, preferences or practices; use of addictive substances; illegal conduct; information that could damage an individual’s financial standing, employability, or reputation; medical record information that could lead to stigmatization or discrimination; and any information about an individual’s psychological well-being or mental health.

The first ethical consideration consisted of obtaining permission from the board of education. Afterwards, the researcher requested approval from the Northeastern University Institutional Review Board (IRB). Once approval was granted, the researcher contacted principals through email and/or written documentation, soliciting teachers who would voluntarily participate in face-to-face interviews with the researcher. Consent forms informing the participants of the purpose, benefits, risks, confidentiality, and the option to withdraw from the
study were given to the participants prior to collecting data. To ensure anonymity, completed documentation was given a numerical code in the first phase of research.

The risk of harm to participants should have been negligible. The sum of potential benefits to the participant and the importance of the knowledge gained should have outweighed the risk of harm to the participant and thus supported a decision to carry out the research. Qualitative interviews on sensitive topics may provoke powerful emotional responses from a participant. An appropriate referral source for professional help should have been ready, should referral have been necessary. Such referral may have included authorities responsible for responding to illegal conduct. Participants were assigned pseudonyms to ensure anonymity and confidentiality.

Ideally there should be reciprocity in what participants give and what they receive from participation in a research project. The investigator is indebted to participants for sharing their experiences. Reciprocity may entail giving time to help out, providing informal feedback, making coffee, tutoring, or being a good listener. The reciprocity should fit within the constraints of research and personal ethics, and within the framework of maintaining one’s role as investigator. Participants should receive feedback on research results, because this is a form of recognition and gratitude to participants for their participation. Finally, the qualitative data was stored in a locked metal file cabinet in the researcher’s residence. All confidential data will be destroyed three years after the conclusion of the study.

Each interview began with an overview of the study and an explanation of informed consent. Before the interview, an introduction to the use of the tape recorder was made and the assurance of confidentiality was presented through a brief discussion of data collection and the software used. Each session concluded with a summary in order to enhance the experience for
the participants involved and to be sure they received some gratification after contributing their
time and experience to the study. In addition, each participant was provided a copy of the
transcript following the interview session for revisions. This gave the participants the
opportunity to add ideas to the transcript. The consent forms and interview transcripts were kept
on file for future reference.

**Limitations**

A potential limitation to this study is that only secondary (high school) teachers were
studied. Although the teacher participants met the qualifications outlined in the study,
perspectives of site-based leaders, as well as the willingness of principals and teachers to
participate, limited selection of schools and teachers. The study was limited to several schools,
several teachers, and only one school district. Therefore, the result might not be generalizable to
other schools, teachers, students, or districts.
Chapter 4: Analysis of Data and Research Results

Introduction and Approach

This study determined what factors affect the perceptions of Broward County, Florida, teachers concerning the implementation of performance pay based on teacher performance evaluations. The objective was to identify factors relevant to an overarching research question through four subquestions or themes. Each interview was viewed as a single incident. That is, each interview was considered individually in the analysis. Common subthemes were identified across the data with regard to addressing the themes. This chapter will present an analysis of the data and an interpretation of the findings from an analysis of qualitative data about performance pay and teacher performance evaluations.

The study was conducted to ensure the interview questions would provide data to answer the research question. In April and May 2016, interviews were conducted with seven participants at one secondary school in the Broward County Public Schools.

A qualitative method was used to explore teachers’ attitudes towards an evaluation tool and its relationship to pay-for-performance after the implementation of a new evaluation model. The qualitative approach allowed the researcher to ask the participants questions that generated conversation. It also provided a way to ask follow up questions during the interview based on the answers collected from the participants.

The study took place at one high school. The high school had approximately 120 teachers and educated students in Grades 9 through 12. The school utilized a departmental approach for core academic instruction. In addition, each department member was assigned to a Professional Learning Community (PLC) team that may or may not be departmental.

Participants’ Profiles
This study targeted teachers of different academic curricula and differing grade levels. Volunteer participants in the study included the following:

- David, a male marketing/DECA teacher, had been teaching for 2 years. In most of his classes, he had a mix of 9th–12th grade students. David had been teaching for 3 years and is in the Career Technology Department.

- Nina, a female English and reading teacher, had been teaching for 13 years. She taught literacy, reading, and writing to 11th grade students and was a member of the PTSA Executive Board. In addition, Nina served as a sponsor to several of the school’s clubs.

- Anne, a female social studies teacher, had been teaching for 17 years. She taught 10th graders and was a voting member of the school advisory committee. In addition, Anne was working on her educational leadership certification to become a school-based administrator.

- Eunice, a female Spanish teacher, was in her 38th year of teaching. She taught all levels of Spanish and has taught Advanced Placement Spanish Literature in the past. She currently taught 11th and 12th graders. In addition, Eunice served as the Spanish Honor Society sponsor and mentored teachers who were new to the school and district.

- Melissa, a female French teacher and chair of the World Languages Department, was in her 34th year of teaching. In addition to teaching all levels of French, she worked with all grade levels and had created a “students helping students” ESOL mentoring program for students who need help academically and socially.

- Laura, a female 10th and 12th grade English/literacy teacher, had been teaching for 19 years. Not only was she a PLC team leader, but she also served as the school’s in-service program facilitator, monitoring all required faculty trainings, workshops, and in-service
points.

• Sophie, a female math teacher, taught 9th and 10th grade liberal arts math and Algebra I. Sophie had been teaching for 1 year and was a former student at the school. She served as the student government sponsor and volunteered her services to the Special Education Department.

Participant numbers were assigned in the order that individual interviews took place. The interviews consisted of 21 questions, divided among three sections. Questions 1-6 focused on teacher perceptions of the methods and purposes of the evaluation tool. Questions 7-14 centered on teacher attitudes about how the evaluation system related to their own specific teaching practices and work behaviors. Questions 15-21 asked the participants to provide their perceptions about performance-based pay structures. Many of the questions were followed up with probing questions related to each individual interview that helped address the research question. All responses were confidential and anonymous.

From the results of the interviews, four major themes developed that helped answer the research question, “What are teacher understandings and perceptions about the new Broward County, Florida, evaluation tool?”

Themes

**Theme 1.** Theme 1 was generated from the first question posed to the interviewees, asking if teachers perceive the new evaluation system as a valid, equitable, and useful tool for determining their performance-related pay. The seven primary subthemes related to this theme and question are summarized in this section. This section includes tables, displayed in the appendices, summarizing the definitions of the identified subthemes, the frequency of occurrence for the subthemes, as well as the number of participants who mentioned a specific subtheme.
As reflected in Table 2 (see Appendix F), the primary subthemes were: (a) *It is not fair or equitable*, (b) *It is inaccurate*, (c) *Teacher impact on students is what matters*, (d) *The whole teacher is not evaluated*, (e) *Consistency is needed*, (f) *The new evaluation system is fair*, and (g) *The new evaluation system is useful if feedback is given*. Table 3, in Appendix G, shows the frequency with which the subthemes appeared across interviews and across the data.

The most frequent subtheme for Theme 1 was *It is not fair or equitable*, which refers to the perception that the new evaluation system is not fair or equitable as a tool for determining performance-related pay. This subtheme was mentioned 19 times during five of the interviews.

Anne said,

Well, I don’t think it’s very fair. I don’t think it’s very fair for a number of reasons. First, the evaluation was conducted before any formal training or information was accurately provided to all teachers. I think it should have been a gradual roll out and not just done in one shot and then do the lessons after the fact. I think the other thing about the evaluation tool is, it is structured in a way where it is very unlikely for anyone to be able to get innovating across the boards. You can be innovating on a Tuesday in a particular element, but if an administrator comes back on a Thursday and you’re using that same element you may not be innovating two days later. I think that’s really hard for teachers to come with the vision of how versatile or how effective a teacher is.

Eunice felt the ratings were unfair. She stated,

I know that no matter how many marks you put down for me, I am never going to reach this exceedingly effective or whatever because the system is set up to just gloss over me. It’s going to give the money or the recognition to someone else, someone who teaches in what the state considers to be an important area. It always frustrates me when the report
cards come out because the report cards have an asterisk next to my class as being a core subject area, yet our governor doesn’t think so. Our superintendent doesn’t think so. Nobody gives my subject area any bit of credit, except the colleges. The colleges think that what I do is important. They think that what we all do is important, but I think my subject area, a core subject area, is important for college entrance. But nobody else does because they don’t want to be – it has to limit my classes to 25 students. That’s what this amounts to, I think.

Melissa also explained why it was unfair.

First of all, I think because a lot of people don’t really know what we teach, in the sense how we teach to the content of what we’re teaching. It makes it kind of an unfair evaluation to some extent because what might look like a very simplistic lesson to foreign language is not necessarily simple. Basically, I think that’s a part of the problem because we teach a very specific content. It can’t really be compared or evaluated the same way as certain other subject areas. I think that’s part of the problem where people might not think it’s so fair.

Laura shared,

I don’t understand it. I understand we need to have evaluations because everybody has to have evaluations everywhere. So I get that part of the pay for performance piece, although I don’t really quite understand it. This is because again we cannot, we don’t impact external factors that may influence my evaluation. I can’t stop at students’ houses and say make sure you’re reading everything, why aren’t you reading, and why aren’t you doing your homework? I suppose if I could do that I can understand the essence of the external factors within this evaluation tool.
In a final example of this theme, Sophie mentioned,

Again, even though it’s a 50-50 question, I still feel like there are some teachers who are going to get higher ratings for things because they have students who are going to achieve on state tests, regardless, because they are good test takers. Those teachers who have all those “good test taker” students, across the board, whether they wanted that schedule or whether because that’s what they were assigned, they now pretty much only have to rely on learning gains. This is opposed to teachers who have a large population of students in their classes who are going to have a harder time passing. I just feel like teachers with higher level students are generally tied to higher evaluation scores.”

The next subtheme for Theme 1 was it is inaccurate, which refers to the perception that the new evaluation system is inaccurate as a tool for determining performance-related pay. This subtheme was mentioned 15 times during five of the interviews.

With regard to the evaluation, Nina explained,

I don’t believe it’s 100% accurate. I do feel it’s very focused on standardized testing and the focus of my classroom time is dedicated to state exams and away from my curricula. But again, I just feel that a lot of it is based on the standardized testing and I definitely don’t agree with that.

Anne also commented on perceived inaccuracy of the evaluation tool. She stated, to receive an “innovating” score (the highest level I can receive for any data mark) in Domain 1, which is the instructional practice piece, when they [evaluators] come in, looking for a particular element, it is required that you show evidence that every single student has succeeded in that particular element. If it’s something like providing feedback in a 15-minute glimpse into a classroom, an administrator may not be able to
observe you providing feedback to 35 students at that particular given time. The feedback may come from a feedback that’s written on an assignment and given back to a student, and/or feedback that comes in increments. You do a couple of students one day, a couple of students the next day, but to be able to do that for 30 somewhat students in a 15 minute time frame is near impossible.

The researcher continued with this statement by saying, “Right. You’re saying the accuracy is not necessarily there;” Anne replied, “Right.”

Eunice shared,

Yes and I'm sure we will get in to this a little bit later, but in terms of the things that Marzano thinks are important, many of them or perhaps all of them on certain levels, many of them just don’t have anything to do with what I teach. They [the benchmarks] aren’t good. They are not good benchmarks of how foreign language is taught. It is just that way because, for example, when you are teaching vocabulary, you have to teach it differently in a foreign language then you teach it when you teach. When you’re working in English, it simply has to be different in dealing with particular vocabulary or vocabulary words, but you cannot ask them to look for it in context because they didn’t understand the context either. You have to teach in terms of what that word means and there can’t be any guessing about it because then they really get bored. You have to have a way of presenting it that’s interesting to them and be direct enough to get the job done quickly. Employing the means that we used in English classes or with language arts classes is very different in teaching my classes because it’s just too conversant in a foreign language class. When an administrator or an evaluator I should say – when an evaluator watches this they don’t really see or know what they’re saying. It is something
that is a beginning level when it really is not a beginning level. It’s simply the way a world language has to be taught.

The researcher probed on this statement by saying, “Got it, so indicators aren’t necessarily there?” Eunice replied, “The indicators aren’t the same as they are with other subject areas.”

Melissa expressed a similar perception:

That’s an interesting question because in my particular field what my teacher struggles with sometimes with me are that some of those standards on there really don’t apply to us at all. Sometimes we’ll meet and discuss and say, “really can’t do much about hypothesis, testing hypotheses,” and some of the other things that really don’t apply to how we teach or what we teach.

In the final example, Laura shared,

Honestly, it doesn’t matter what I do. I’m always going to get a certain amount of applying and maybe one innovating and one beginning because the county doesn’t want to see somebody with all innovating because then it’s more explanation. I feel that the evaluations are not really that accurate of a reading tool.”

The next major subtheme for Theme 1 was Teacher impact on students is what matters, which refers to the perception that rather than focus on pay for performance, a teacher’s personal influence on students is what matters. This subtheme was mentioned 13 times during five of the interviews.

David stated,

To me right now as a third year teacher, this is the most I ever made in the job. I’ve been in the workforce for 13 years. For me, I’m not necessarily concerned about how much I
get paid. I’m more personally worried about how I’m impacting the kids. In that sense, I believe I’m highly effective.

Nina stated,

I think you should also base it on how well the students are doing in your class as far as grades, as far as what they’ve learned. I definitely disagree with the whole standardized testing and focusing just on that, but I do feel it is partly necessary only because there are teachers that really depend on teaching to the test. If I’m working and I’m teaching and students are learning, that is good enough for me. There are people sitting at their desk, not doing anything, not teaching and we are all going to be paid the same. I don’t think that’s fair.

Anne explained that impact on students is what mattered:

The pay for performance piece does not influence the way I teach my students because I feel I’m not eligible for it, because I’m on the grandfathered scale. I teach my students and I do the best that I can do with them because of how I feel about teaching and what I want to do for my students and what I want to do for my school. The pay for performance itself does nothing for me.

Eunice also felt that teachers were more focused on impacting students than on the evaluation tool:

I don’t find that people in my particular department are necessarily looking for the extra money. They want kids to learn and they want the best situation for the kids to learn. They want to be able to do a good job without a lot of stress.
Sophie felt similarly, as evident in the following quote: “Obviously I’m accountable to my students’ learning gains and at the end of the day that’s the whole point of education: to get these kids to learn, whether or not my evaluation score is high.”

Another commanding subtheme for Theme 1 was that *The whole teacher is not evaluated*. This refers to the perception that the new evaluation system is rigid and does not assess the whole teacher. This subtheme was mentioned eight times during four of the interviews.

Eunice felt the evaluation was subjective and rigid. She stated,

It is subjective because first of all a lot of people come through and look at what you’re doing, but they mainly might be there for just five minutes. They have no idea what you did yesterday. They don’t know what your relationship is with these students. They may not even know you. When you come into my classroom because you’ve been there many times, you have a sense of what you can expect but when another administrator comes in they don’t know me. They’ve never been very much involved with what goes on in my classroom or who I am or what I’m even teaching. The evaluator who supervises math for example, may come in to a foreign language classroom and not even speak Spanish. They may not understand the word I’m saying unless on the day when I’m speaking English. It gives me pause to think why they are even here because they don’t really know what I’m talking about. They don’t know what I’m doing.

Laura felt that the evaluation only provided “snapshots” and not an overall picture of the teacher’s effort:

No, I don’t know that is exactly accurate or I don’t like the word fairness. I don’t think it’s necessarily fair because when somebody comes in here, she does not know what had
just happened or what’s going to happen after she leaves. So the problem with that is sometimes this snapshot isn’t going to capture the whole story.

Sophie said,

I feel like I personally know that I can achieve a lot more but because we have this iObservation tool I feel like I have to almost be in this box and fit this mold of exactly what this tool is looking for. And for example with the learning goals, I feel like it’s this way whereas I may have another way to incorporate that but if you don’t see that when you come into my room then I feel like it takes away from that and I just feel I could do a lot more on my own. I feel when I’m being evaluated that I have to fit in this mold.

Again, I like to teach a lot of things out of the box and when I wasn’t so concerned about test scores and evaluation and all that stuff, my kids were learning at a really fast pace. Because all of these teachers are evaluated on the same tools, we feel like we all have to teach the same way at the same time just like each other and it creates kind of a robot model instead of letting teachers step out. Sometimes you all have activities that fail and I think the that failure is how you learn, but at the same time we’re scared to see that failure because our scores are tied to that.

The next subtheme for Theme 1 was *Consistency is needed*. This denotes the perception that consistency is needed in the use of the new evaluation system as a tool for determining teachers performance-related pay. This subtheme was mentioned six times during three of the interviews.

David said, “The only thing that I would say is that there needs to be someone to just make sure that there is consistency across the board.”
Eunice provided a detailed example of why consistency was lacking in the evaluation tool itself. Eunice voiced,

I think the bottom line here is that I don’t think that the evaluation tool is useful for my area and therefore every time I read it and every time I try to comply with it, I think I must be doing something wrong because you have to have a pretty good imagination to see how some of these things are going to be fit in this area in this particular area. I think one of the problems with it is that when you take one tool and apply it to multiple areas of teaching, you keep the right answer. The teacher that’s sitting there many times in our meetings – well there’re PLC meetings this year, but in our department meetings we talk a lot about what a particular goal on the evaluation tool means in our classroom. It’s unfortunate that all of the administrators who might see us haven’t been able to sit there and listen to us talk about what we think our goal is here and what it means in my classroom for my level of student. When you look at it for example – I don’t know just talking in general about something compare and contrast. You look at that and what that means for a fourth year Spanish student’s completely different than what it means for a first year Spanish student, not to mention what it means in math or social studies or any of the other areas. So I think that tool is too detailed in some ways but it’s not detailed enough in others. You have to have a tool that is specific to the subject area.

In a final example of this theme, Laura mentioned a lack of consistency among evaluators and raters:

So I don’t know that I necessarily think that it can be accurate and I don’t think it could necessarily be fair because administrator A may be more critical, and administrator B who comes in right after him may be a little bit more flexible. I only see that in English
classes, we see that with subjective grading on essays. Two people can read the same essay and come up with two different scores. It’s the same thing with this.

Another common subtheme for Theme 1 was *The new evaluation system is fair.* This implies the perception that teachers feel that the new evaluation system is fair as a tool for determining their performance-related pay. This subtheme was mentioned three times during two of the interviews.

David affirmed,

I think it is a fair process and a lot of people have different ideas or comments. For me, I use the tool as a way to improve my teaching; at the end of the day, I look to improve my teaching to help my students to learn and retain more information.

Nina proclaimed,

I think it is a good tool as far as teachers that come to the point where they don’t want to teach or they don’t care anymore. It gives them more motivation to work harder and to develop a new teaching philosophy or if they don’t mind just teaching to the test. This tool can encourage new ideas and thoughts in some of the older teachers ready to retire.

The final crucial subtheme for Theme 1 was *The new evaluation system is useful if feedback is given.* This conveys the perception that the new evaluation system is useful as a tool for determining performance-related pay if feedback is provided to teachers from the evaluators. This theme was mentioned two times during two of the interviews.

For example, David explained the importance of feedback post-evaluation. He stated,

In my personal experiences I have some great points of getting feedback. It’s been really good. I used those comments to make sure I’m positively moving forward. I do different things and mix things up. It’s very important that some of the administrators are also in
the classroom for 20 minutes, 30 minutes to see what is going on and evaluate subjectively what they see. It’s really up to the teacher during that time of feedback, to have a conversation about what your intentions were and what you were trying to get across to make sure that everybody’s on the same page. Feedback for me has been very good. I just make sure that when I’m in there talking to my administrator that they understand exactly what the point was on what I was trying to do.

Similarly, Anne revealed the importance of post-evaluation feedback:

The only time that I think that that conversation is open is during the formal observation because what we’ve been told in the past is the formal observation is not finalized until you sit down and have that post conference. At the post conference, which is mandatory, you sit there and you have that conversation and you’re able to talk about what you were planning, what you were thinking about, and what you saw in your lesson as compared to what the observer saw. Then you can come to agreement or at that point, if you disagree about something, you can voice your opinion at that time and then adjustments could be made to the evaluation.

**Synopsis**

Theme 1 generated the question, “Do teachers perceive the new evaluation system as a valid, equitable, and useful tool for determining their performance pay?” After honest and thorough discussion with the participants, it was found that the system is not fair and is inaccurate. The participants believed that this tool is subjective and that much of the observation is based on the mood of the evaluator. The thought was that if the evaluator does not like the teacher or if the evaluator had a bad day, then their observation would be effected by the emotional state of the evaluator. While the participants believed in the inaccuracy of the tool,
they felt that teacher impact is what matters. They perceived their own treatment of their students as imperative in supporting student achievement. The participants feel that their job goes beyond the teaching position. They feel that their impact on students involves being so much more than what they do in the classroom. Another area of contention with the participants is that the whole teacher is not evaluated when being observed in the classroom. Many of the participants feel that their extra-curricular involvement in the school including advising and sponsorship of clubs and sports, and the outside work they complete to plan field trips and competitions is not included in their observation, thus not speaking to their efforts outside of the classroom. The participants believe that consistency is necessary to evaluate their performance in the classroom. They stated that the fact that they may be observed by more than one evaluator, at any given time, creates inconsistency in their observation and/or evaluation. One evaluator might visit at the end of the lesson where another might visit at the beginning; one evaluator might be more familiar with the subject content rendering a more strict observation; or one evaluator just might not like the teacher being observed. Finally, the participants believed that feedback is paramount to their improving instructional practices. Quite often, they are observed and not told what they did well or what they need to improve on. With this being said, they reiterated that feedback was often not given to them after a particular observation. Without feedback, they felt that they could not improve because they were not told what they needed to improve on.

Theme 2. Theme 2 speaks to whether self-efficacy affects teachers’ perceptions of the new performance-based evaluation tool. The four primary subthemes related to this theme are summarized in this section. As reflected in Table 4 (see Appendix H), the primary subthemes were (a) Teachers are self-efficacious and confident, (b) Teachers want to self-improve, (c)
Teachers lack confidence with the tool, and (d) Teachers need more training, development and knowledge to feel confident. Table 5 (Appendix I) shows the frequency with which the subthemes appeared across interviews and across the data. Participants were very much intent upon discussing their own experiences with self-efficacy and their belief in their own ability to complete tasks and reach goals. This was a real concern in the majority of the interviews, as most of the participants understood that self-confidence and self-worth were integral to ensuring both personal success and student achievement.

The most frequent subtheme for Theme 2 was Teachers are self-efficacious and confident, which refers to the perception that teachers are self-efficacious and confident about the new performance-based evaluation tool. This subtheme was mentioned 15 times during four of the interviews. Several examples of this subtheme are presented below using quotes.

David expressed self-efficacy when revealing,

Personally, I try to keep up to date on what I have there, what I need to do to make sure that I get important time whenever I’ve been observed on the past. How can I improve myself, but just generically by making sure I understand what is expected of me as a teacher as far as those criteria are concerned? It’s up to me to be the one to improve upon those things whether by analyzing my data marks and seeing where I need to improve or taking my feedback from my administrator and using it to come up on new ways to improve. It’s all helping me just improve so it definitely increases my confidence for sure.

Nina asserted,

I actually take what we learned on campus to better myself as a teacher in the classroom and hopefully I can utilize a lot of those best practices in my room and that will help with
the standardized testing. As far as being confident, I am very confident in my curriculum.

The same teacher later stated, “I am a confident teacher. I know I have good skills. I have various lesson plans that I’ve come up with and or gotten from other people that have been successful.”

Anne expressed self-efficacy as well:

I’m effective. I’m not going to get those highly effective marks because I really do think the evaluation tool is set up so that most teachers will be effective and not highly effective. Now, I just focus more on good teaching, which I think falls into this evaluation tool. A couple of the things that I’ve changed and that I continue to use in my classes is that I have put more emphasis on using learning goals and skills in my lessons and providing students more feedback. But other than that, just good teaching, I think if you’re doing good teaching, you’re hitting most of those elements.

In a final example of this theme, Laura expressed confidence in her ability to do what is needed to appear competent to any administrator:

Personally, I need to be aware, every moment, that somebody could have come in and I had to ask myself the question all day, is what I’m doing something that would be perceived positively if an administrator or walks in to evaluate me.

The next prime subtheme for Theme 2 was Teachers want to self-improve. This conveys the perception that teachers want to improve their own teaching to meet the requirements and standards of the new performance-based evaluation tool. This subtheme was mentioned 13 times during four of the interviews. There were a variety of examples of teachers taking steps to
improve their teaching in light of the new performance-based evaluation tool; a few are shared below.

David shared,

I choose my professional development activities in two ways. Number one, I previously mentioned about how I constantly like to look. Where can I improve within the iObservation system? Number two, if I’m choosing professional development I want to choose the things that I need to work on because that makes the most sense. Anything that I have a beginning mark on which I’ve had two of in my two years of teaching, I can apply what I’ve learned as a classroom teacher to get to the innovating mark, I might do that. I also sign up for professional development that I know I’m already personally interested in because I’m a firm believer that if I do this in my classroom, I will be teaching something and learning something that students need and want to learn. I’m only going to go so far to pay attention in a training if I want to be there and if I want to understand that particular topic. Therefore, I choose professional development that I need to improve upon and I know I need to improve upon and take a self-reflection on and also things that do interest me.

Nina disclosed, “I’m also very open minded to learning new things and finding new ways to teach it because if it has been working for 10 years, it doesn’t mean it’s going to work now.”

Anne mentioned,

I looked not only to resources that are school district provided but I looked at resources from other school districts to become more knowledgeable on the standards of the evaluation system. I want to know what they looked like, what it meant, how to
implement it, and how it was rated to really get that greater understanding. I volunteered myself to go to five-day training from the makers of the evaluation system to really get that accurate information and get it from the source, so I could have a greater understanding. This way, I could know what I do need to do in my classes to get those higher ratings.

Finally, Melissa provided the following example of how teachers work together to improve their teaching:

Well, one positive thing as a result of it is when we work together as a department in our PLCs and in general, and store and share best practices, we do create more activities to reflect the activities that are considered good on the Marzano benchmarks tool to achieve a higher placement rating. We’re always looking to improve achievement based on what we’re doing. Is there a better way to teach the lesson and/or curriculum? Maybe we can break it up into better chunks. We kind of look at that and so we do use that to help restructure our lessons that helps kids.

The next considerable subtheme for Theme 2 was Teachers lack confidence, referring to the perception that teachers lack confidence in the new performance-based evaluation tool. This subtheme was mentioned 11 times during five of the interviews.

When asked how the evaluation process, support, and feedback influenced your confidence in your ability to teach students and does the process itself lend itself to a little bit of lack of confidence when you feel you might not be doing what your supposed to be doing, Nina said, “A little bit.”

Anne explained how confidence can decrease when the new performance-based evaluation tool is used:
I think at first. I think you became worried. Oh, what do you mean I’m not innovating. I can hear I’m not getting the innovating. Is anybody else not getting innovating. And you start to experience a little bit of self-doubt. But then when you grow to really understand the evaluation tool and the ratings, you realized that it’s not about you not succeeding or being at the highest level. This is just how it is and you come to accept that and as long as you’re doing your best and reaching the majority of students, then you have to take away from that you’re doing a good job.

Eunice proclaimed how the tool led her to question herself:

I think that the evaluation tool itself makes me question whether or not I am doing my job and because it isn’t well written, specifically for my job, I think that it doesn’t really help me much to know whether I should go on, whether I should go back, and whether I have done the job that I needed to do. Just to give a little bit of an example, I’m not exactly sure I’m answering this question but I try to say to the kids, “Do you feel like you know more now? Are you more comfortable with this now than you were when you came in today?” Sometimes I used something as general as… I think it has made me think yes I need to check on these things. But it is also made me think if I don’t know whether the kids understand this, what have I been doing for the last 38 years.

In another example of this theme, Laura indicated, “there are people that feel that they’re not comfortable enough with the process and the tools, for sure.”

The final subtheme for Theme 2 was Teachers need more training, development, and knowledge to feel confident. This reflects the perception that teachers need more training, professional development, and knowledge to feel confident about the new performance-based evaluation tool. This subtheme was mentioned six times during two of the interviews.
Anne shared,

I personally have felt that as the evaluation system was rolled out, accurate and timely information was not provided. I spent five whole days looking at the entire evaluation tool. I really think that a full five-day training is what teachers needed from day one to truly understand, not only what each element means and what it looks like and what they need to do to meet those standards, but to understand how they all tie in with each other and really get the global picture.”

Later, this same teacher reiterated, “I still think that the majority of the professional development has been pretty superficial.

Laura had a similar perception. When the researcher asked her “Okay, have you found the need or have you heard of anybody who felt that they needed to attend their own professional development to help them with this evaluation tool and the evaluation process?” Laura replied, “Yes.” The researcher probed further by asking, “You have? Okay?” and Laura elaborated, “Yes, there are people that feel that they’re not comfortable enough with the process and the tools, for sure.”

**Synopsis**

Theme 2 asked, “Does self-efficacy affect teachers’ perception of the new performance-based evaluation tool?” The study participants perceived self-efficacy was a trait that they all have and that they are confident in their abilities to motivate students and assist them in student achievement. The participants want to self-improve. They did not discount the need for self-improvement and they did feel that they could gain insight into new strategies and curricula from making an effort to self-improve on their craft. They did confer that they lacked confidence in the evaluation tool, as a whole, because they don’t know enough about it to feel comfortable with
its contents. In addition, the participants don’t feel that there has been enough district training and professional development on the tool itself in order to understand the domains and elements involved in each rating. In order for the participants to fully believe in this evaluation system and to fully understand the ramifications to performance pay, they must be offered advanced follow-up, consistent training.

**Theme 3.** The main question that developed Theme 3 was whether teacher performance evaluations and performance-related pay caused stress on teachers. The two primary subthemes related to this theme are summarized in this section. As reflected in Table 6 (see Appendix J), the primary subthemes were (a) *Teacher performance evaluations and performance-related pay cause stress on teachers* and (b) *It has caused competition among teachers.* Table 7 (see Appendix K) shows the frequency with which the subthemes appeared across interviews and across the data.

The first major subtheme for Theme 3 was *Teacher performance evaluations and performance-related pay cause stress on teachers.* This refers to providing support to a variety of stakeholders. The data for this subtheme were further classified into three topics: (a) *The classroom observation evaluation causes stress*, (b) *The connection to pay is stressful*, and (c) *The overall performance evaluation is stressful.* The topics are detailed below.

*The classroom observation evaluation causes stress* was the most frequently mentioned topic under the overall subtheme of *Teacher performance evaluations and performance-related pay cause stress on teachers.* This topic refers to the perception that classroom observation and evaluation specifically causes stress for teachers; it was mentioned 23 times during all seven interviews.

In her interview, Nina explained,
I think regardless of how many years you have, how confident you are, at least with me, whoever walks into your room whether it’s an administrator or another teacher, you are a little nervous. I don’t really change a lot of what I do in the class and or almost always teach and I’m not going to just put on a fake show or be fake when someone comes in. I do want them to see how authentic I am.

Eunice divulged,

I can see where with many teachers it would inspire fear and insecurity. I think sometimes it depends on the evaluator, but I think sometimes evaluators don’t realize how stressful it is for a teacher to be evaluated especially when there are issues with who’s doing evaluation, whether they actually know anything about what you're teaching, whether that particular evaluator has a reputation for being easy to get along with, and understanding versus somebody who isn’t. So I think that there’s a lot of opportunity for fear and stress to be there. I think it is a very stressful thing for whatever reason it is because I don’t think the tool is addressing the things that I’m supposed to be doing in my classroom to be a good educator in Spanish or whether it’s stressful because I feel like people who are evaluating me don’t know what I’m supposed to be doing or because of personalities or whatever.

Laura expressed the same perceptions:

I know the new teachers do feel a great deal of stress, large degrees of stress. I remember being stressed my very first year of teaching because I was going to be evaluated. I didn’t have this tool, but it was just the fact that it’s somebody going to be coming into my classroom and evaluating and observing me. So I think that there is that degree of stress, but more so now I feel more veteran teachers are feeling the stress and it’s not that
they are doing a bad job, it’s just that they’re not sure if it’s going to be perceived the same way that they think it should be perceived. So there’s that stress; it’s all about that subjectivity and the idea that we don’t know if an administrator who comes will think my lesson is going to be great. You sit there going well I’d be okay when administrator A comes in. But administrator B comes in, I’m going to be in big trouble. So now because I don’t really know, and it’s that if I’m doing anything wrong or different it’s just that you never know how this going to be perceived.

Finally, Sophie affirmed,

It’s definitely really stressful because I feel like if I want to step back, it’d happen to be a day that an administrator came in and my lesson kind of fell apart. My evaluation might be affected, where obviously if it went well I would be rated well. But at the same time, it’s like I have that fear of if I do something and it fails that now my evaluation will be a reflection. Even though that’s necessarily true, it’s just that kind of feeling that is viewed by the new teachers.

The connection to pay is stressful was the next most frequently mentioned topic under the overall subtheme of Teacher performance evaluations and performance-related pay cause stress on teachers. This topic alludes to the perception that the connection of the performance evaluations to teacher pay and salary is stressful for teachers; it was mentioned 10 times during five of the interviews.

During his interview, David shared,

I totally see that there’s a lot of stress on pay for performance because a lot of teachers really do hang their head on the fact that they need to be highly effective and do depend,
in that way, to make that extra money for themselves, for their families or whatever it is.

There is a lot of stress, in length, to the pay for performance for sure.

Nina explained this stress as follows:

Personally, it is very stressful knowing that at some point down the road that my evaluation is going to be based on test scores and if my students don’t do well, it’s a like trigger effect. I’m stressed out about the test and that stress pushes out on my students.

Now, my students are stressed about the test. Again, wanting to do well or having those students that don’t care, affect your score and your test scores too causes that dreadful anxiety.

Anne divulged,

There is stress in terms of whether or not a teacher should stay on the grandfathered scale? Should they move to the pay-for-performance? What’s going to ensure the greatest chance of them getting a raise and getting paid for their hardworking commitment to the school?

Melissa (who is not on the pay-for-performance scale) said,

Now I think it would be very stressful because trying to reach that higher level of pay on the pay-for-performance scale requires that your kids have to do really well and that is if you’re teaching those end of course exam (EOC) courses.

In a final example of this topic, Sophie shared,

I think theoretically it’s a good idea; however, there’s just so many compounding variables when you’re talking about evaluation and high stakes testing that the fact that my salary is now tied to the fact that I could have varying groups of kids. I mean it just adds a lot of stress and just makes it a lot harder to enjoy the profession. The fact that I
could be making the same amount of money in 10 years from now because I always teach
the same level students makes me anxious and stressed. That’s really stressful. This
would not happen in any other profession. The fact that it happens in teaching is really,
really stressful.

_The overall performance evaluation is stressful_ was the final heavily discussed topic
under the subtheme _Teacher performance evaluations and performance-related pay cause stress
on teachers_. This topic was mentioned seven times during five of the interviews and it refers to
the perception that in general, the teacher performance evaluations are stressful.

While being interviewed, Anne disclosed,

_There is stress in teacher performance evaluation. For me personally, I think it is because
I have such a greater understanding of a rating system. A lot of times, I disagree with the
scores that I get and you’re in a position where there’s nothing that you can do. That in
itself is stressful because you knew you felt you should have done better on something.
You know you should’ve been scored better on something and you’re stuck with it;
you’re at the mercy of whatever somebody else perceives._

Eunice revealed, “_It’s stressful for me on a number of levels because I know how much
it’s going to get in my way of accomplishing my classroom goals and my subject area goals. It’s
really a big stressor._”

Melissa voiced,

_Each year I had to talk to a few of my teacher friends who really have been so stressed
out by it. To be honest what I basically tell them is this: Look you’re going to be an
effective teacher. None of us are ever going to be highly effective teachers because we
always rely on all the school’s scores and so we’re never going to get there so why stress_
yourself out? Be the good teacher that you are and just try not to worry about it because at the end of the day, you're going to be an effective teacher according to the observation system.

In the final example of this topic, Sophie indicated,

Teachers are constantly curious to all of these other teachers. For example when I give a test, if my average is much lower than all the other teachers and I’m grading on the same scale, now I feel like maybe my teaching is not as good as others. Maybe I’m not as competent or whatever and so the stress kind of falls back on the teacher versus the student gains.

The final subtheme for Theme 3 was *It has caused competition among teachers*. This explains the perception that teacher performance evaluations and performance-related pay has caused competition among some teachers. This subtheme was mentioned 10 times during six of the interviews.

David asserted the potential competition among teachers in the following manner:

I feel that maybe the evaluation tool prohibits a lot of teachers from being confident enough to ask best practices. They’re not really comfortable with somebody because at the end of the day they want to be viewed as a better teacher than another English teacher in the department or whatever it is. To do that they need to make sure that they don’t necessarily share all the things that they are doing; I guess that being a problem for sure. Number one, I definitely think teachers hold back because they don’t want to share everything that they’re doing because, again, they want to do something better than anybody else to be that highly effective.

Anne shared,
I think there might be at some point some strategic request for scheduling of specific classes with specific students. But other than that, I can’t see how teachers wouldn’t continue to work together because in so many cases the overall score is what goes toward a particular teacher’s evaluation.

Laura proclaimed,

I know there are concerns and I’ve heard there are some teachers who are actually very competitive. I’m not competitive yet, but I know that there are some teachers who have been adamant about wanting and asking for higher level classes because they don’t want classes where they’re going to end up in a situation where they can’t do anything to move the children and it does happen.

In regard to the evaluations and ratings, Sophie made known, “There is competition on that end.”

Additionally, Eunice explained how there can be competition among teachers from different schools. She disclosed,

I definitely think so because if you’re on a pay-for- performance scale and you're trying to prove yourself in a classroom setting for an administrator, whoever the evaluator might be, all of these things that you’re doing, you think it’s going to give you that innovating mark. In fact, I would say the one experience that I’ve had is a certain amount of, I don’t know exactly if I would call it jealously. It’s more like annoyance that somebody, but it hasn’t happen so much here at the school. But I have a friend who’s at another school and we’re all on the same scale in that way. I have friend who’s at that other school and she would tell me all the time, “Oh, we all got innovating.” Whatever it was that we’re being evaluated on, it was really annoying to me that I wasn’t getting those innovating
scores because depending on the administrator, the innovating rating causes terrible competition among colleagues not only at our school, but other schools too.

**Synopsis**

The Theme 3 question asked, “Have teacher performance evaluations and performance-related pay caused stress on teachers?” The overwhelming response to this theme was both the evaluation tool and the idea behind performance pay are stressful. The participants spoke to the piece that the idea they are being observed in their own classroom setting causes them to be stressed and anxious. The thought of an unannounced visit observation leaves teachers with constant state of stress. Moreover, the participants perceived that the idea of their pay-for-performance salary being affected by their final evaluation score from the classroom observations is also extremely stressful. Tying performance pay to “outside” factors on an observation is stressful and often causes anxiety on the teacher. Another area discussed under this theme was that the evaluation tool does play a factor in competition among teachers. Even though the participants believe in sharing best practices and offering support to their colleagues, they worry that what they share may be copied by another teacher whose final evaluation rating may offer a higher performance pay salary scale than the teacher sharing that best practice. Teachers are now leery of sharing their knowledge and experiences.

**Theme 4.** The question that created Theme 4 addressed other factors teachers perceive as influencing performance evaluations and performance-related pay. The five primary subthemes related to this theme are summarized in this section. As reflected in Table 8 (see Appendix L), the primary subthemes were (a) *Test scores matter*, (b) *Student behavior matters*, (c) *Teacher collaboration matters*, (d) *What you do outside of the classroom should be considered*, and (e)
Additional approaches and tools matter. Table 9 (see Appendix M) shows the frequency with which the subthemes appeared across interviews and across the data.

The most frequent subtheme for Theme 4 was Test scores matter. This conveys the perception that students’ standardized test scores influence performance evaluations and performance-related pay. This subtheme was mentioned 22 times during six of the interviews.

For example, Nina shared,

Now, I feel that the focus has been taken away, being that we’ve added course exams and things like that. It’s definitely, I feel it is detrimental to the students. Again, I know that they have to learn to take tests because that’s part of post-secondary training and whatever. This might help students know what they will do after they graduate, but I don’t think we’re teaching the students what we should as far as like at least in English literature, things like that. I think we’re again very focused, all on the testing to the test, which we weren’t before.

Anne said,

Well, basically if you teach a subject area that is tested with the state-wide standardized test, the scores that your students achieve on those tests are used to calculate your VAM scores. But where I think it becomes a little unfair is that when teachers don’t directly teach students who are being tested on those standardized tests and the overall school scores are being used for their VAM scores. These teachers may have never spent five minutes with any of the students whose data is being used to be included in their evaluation. Especially take 11th and 12th grade teachers whose scores are being based on 9th and 10th grade data or taken to prepare those students for that particular
standardized test. The scores that the students get on the test are being counted towards their evaluation.

Melissa voiced, “Professionally it’s just testing, testing, so much testing. When I think – you just teach, just teach my content. Have the kids learned what I want them to learn, but this is the world we live in.”

Laura added, “I don’t know if I have enough students where it is going to be my test scores that are going to impact my performance pay or the school wide score.”

In a final example, Sophie indicated,

I'm aware of some of it. I know that we focus on students learning gains, now, as well as their actual scores. So obviously we want students to be passing, but I know that the learning gains are now also factored into the scores. However, I feel that with teachers who have a lot of low achieving students may have learning gains in there but it is still the perception among teachers that they’re still not going to get those scores that the district is looking for.

The next most frequent subtheme for Theme 4 was Student behavior matters. This embodies the perception that students’ behavior during the classroom observation and evaluation influences performance evaluations and performance-related pay. This subtheme was mentioned 12 times during all seven interviews.

David explained how student behavior can impact an evaluation:

Most of the time, I will tell my class that an administrator is coming. But a lot of the times, I don’t say I get distracted often. I just want to make sure that they’re 100% behaviorally prepared, but I know that’s not really the point on what it would really look like. I don’t want to say it. If somebody misbehaves or they do like that, I take a step
back and I maybe react in a different way that I normally would. I try to the best of my ability to keep students on task to make sure that, especially during an observation, I feel like at the end of the day the whole point is showing what I am teaching and how I am teaching. Obviously the classroom management is a big part of it, but nobody is perfect in the classroom. Everybody is going to have those kids who are issues regardless of how great you are with implementing discipline and implementing different structures. I do the best that I can to try to make sure, that at the end of the day, what I’m trying to teach gets across.

Nina revealed,

I think when an administrator or anyone for that matter walks in your room and your students aren’t behaving, it definitely is a reflection of you, not to say it might be a student or two having a bad day or what have you. I think that does affect your lesson; again it’s a trickledown effect. You have one student who is acting up. You’re trying to teach and the negative behaviors affect the other students in the room. If you’re not careful, it could spread like wildfire. But I will say, “Thank God I have really good classroom management,” so I can stop that immediately.

Continuing on this topic, Eunice communicated,

Most of the kids in my class - and I’m saying their parents have raised them to be polite, to be attentive, to be quiet while somebody is talking to respond when somebody is asking – they have those manners. When I was teaching Spanish 2, for example, where we notoriously have students who (a) don’t know what they were supposed to know from Spanish 1 and then (b) aren’t really well behaved anyway, that’s a kind of a terrible combination. If I’m teaching Spanish 2 class and an administrator comes in and kids
were all walking around, have their cell phones out, and I haven’t really got them organized yet, that’s going to say something negative about me. You can come to the conclusion that an administrator, evaluating me at the time, can tell something about me as a teacher. I think that those behaviors do impact my evaluation.

Melissa affirmed,

Well once again depending upon what you teach, I’m very fortunate and that I really don’t have any classes that I would consider who have a lot of discipline problems. I might have an occasional kid who talks too much or whatever, but basically this could be a serious thing if you have a group of kids who are low level kids and/or disrespectful and I know some teachers at the school that deal with this on a daily basis. It’s serious stuff; you’re trying to teach a lesson and sometimes the kids don’t pay any mind. Well you have other things, you have other parts of that Marzano iObservation tool that kind of lend themselves to maybe getting negative marks as a result of all those things mentioned.

Laura said,

Yeah, I think one of the factors could be that the person coming in may know how to teach because hopefully the administrator is better in the classroom which I know wasn’t always the case. But the other situation is if that person doesn’t understand the subject matter and doesn’t understand how it is disseminated or the best way to disseminate it, I think that another big factor that could influence my evaluation. We know our students, but it doesn’t mean that every administrator who walks in and evaluates, knows our students. So what may work for my first period isn’t necessarily going to work for my second period and we have to adjust. For instance, somebody comes in and evaluates me
during the period and I’m reading with the students and it is going beautifully. Then a different evaluator or even the same evaluator comes in and it’s my next period and I’m reading the same modern text with the students, he or she might wonder why isn’t this like the previous class? Well because this is a very low academic group and they don’t really grasp the material as easy as an honors group, the actual evaluation can be very different. So those factors come in to play and I don’t know that the evaluator necessarily has a handle on the climate of the class.

Finally, Sophie added,

I’ve had students who have been usually pretty well behaved when an administrator comes in, but again I’m constantly worrying about managing my class with an iron fist because if someone walks in and there’s a kid out of line, I'm going to get in trouble for it and no one wants that.

The third most frequent subtheme for Theme 4 was Teacher collaboration matters. This refers to the perception that teachers’ ability to collaborate with other teachers influences performance evaluations and performance-related pay. This subtheme was mentioned 11 times during five of the interviews.

To begin, Nina confirmed,

There is also a really good component that is outside the classroom. We talked about teacher collaboration, which I think helps the teachers not only in the room but outside to show that they can, not only teach in the room. They can also take those practices and share them with their fellow colleagues.

Anne revealed,
I think we have a great team of teachers at the school. I think we are big collaborators. We all engage in sharing best practices and we try and improve each other’s practice because in the end it impacts a greater amount of students.

Eunice shared,
I talked about this a little bit before because some of us are on this pay-for-performance and some of us are not. We tend to help each other quite a lot. We do whatever we have to do to make it simple for our colleagues to do their job. We tend to share material and if I make a test in a subject that somebody else is teaching, I share it. If I make an activity, I share it. If somebody needs to know how to do something, I’m perfectly willing to show them how because there is no negative impact on me for them knowing. Melissa added, “We’ve managed to all get together and work together to come up with activities we can use to meet the standards to support adjustable trends, all of those different things that administration looks for.”

Sophie posited,
Definitely, once we get our evaluations a lot of my colleagues talk about the results because we are professionally and personally close. We talk about those things including activities that someone did in class that were really seen as an innovating rating in the administrator’s eyes. So then we kind of share those, but we also talk about where we went wrong and how other people handle that situation. Therefore, we do communicate; we do use some of those tools to communicate which is a positive thing about the iObservation system and tool.

The next subtheme for Theme 4 was What you do outside of the classroom should be considered. This speaks to the perception that what you do outside of the classroom should be
considered in performance evaluations and performance-related pay. This response was mentioned six times during three of the interviews. David expressed that the things that happen outside the classroom should also be considered:

I know that for me particularly, I make sure to invite others any time I have an event or a thing after school or whatever it might be. I make sure to invite the principal, the administrators, anybody really is invited to just kind of see because our kids deal with a lot, you know. Obviously I teach them business skills, but a lot of them learn leadership skills, putting on events, event planning, stuff like that—stuff, being inside the classroom from a book. So I feel like having those aspects and having people see firsthand the stuff that I do, outside of the school, reflects that the kids are still learning in a just different way. Just making sure that people are invited and people come out and see what’s going on is extremely important. I do a lot of work off the clock so many times every day. I’m not saying that I necessarily deserve extra money for that. There’s something that goes to saying there are teachers out there who really do go above and beyond for their kids that might not be the best in the classroom. I do feel that while I understand the need for some structure, I do feel that maybe it needs to be looked upon to see other elements that might help. Because at the end of the day, I feel what I do outside the classroom for the kids is just as important as what I do inside the classroom. It’s a mixture of both.

Nina said,

In my opinion this might be just part of it, but how you conduct yourself in the room, outside the room on campus, and outside in the community I believe plays a big factor. I’m not saying it’s everything, but I know maybe a little bit of it is based on that.

In the final example for this theme, Sophie asserted,
Definitely performance-related evaluations are affected by things that you do outside of the class along with helping out a lot around the school and sponsoring and overseeing clubs and things like that. There is a much higher or better perception of you than if you’re a teacher that doesn’t want to volunteer to do anything.

The final subtheme for Theme 4 was *Additional approaches and tools matter*. This suggests the perception that additional approaches and tools, used in the classroom and with the curricula, impact the quality of teaching and influence performance evaluations and performance-related pay. This subtheme was mentioned two times during two of the interviews.

David felt that additional approaches and tools that impact the quality of teaching do influence performance evaluations and performance-related pay:

I definitely think it does because I try to relate everything that I do vice versa both inside and out of the classroom, but I teach very one-sided in the classroom. I try to do different things, but I add points in group projects because those are my experiences from when I worked in the business world. So for me, it’s a different dynamic from seeing somebody that I’m teaching out of a book or a PowerPoint as opposed to seeing me help kids lead something. If I do a group project in the classroom, there are students who are taking leads on the project and I am not necessarily teaching. But again it’s within a confined requirement and a confined rubric and stuff like that. Whereas, a group project event allows students to perform and shows my level of instructing students to perform, it provides a different dynamic. I think if you see a good grasp of both, I think that you can evaluate my qualities as a teacher within the classroom. I just think having that extra thing just helps supplement the things that I’m already doing well.

In another example of this theme, Laura said,
Well I feel like I’m wasting time doing things that are not going back to support learning. When this happens, I’m doing things that are not good benefit to students. The whole idea of rating students’ understanding by asking their level of learning as one, two, three, four is insignificant. When I spend an inordinate amount of time putting together the scales so that students could hold up their fingers with one, two, three, four, they look at me and they laugh when I say it. It isn’t that they laugh when I say it because I am a veteran teacher standing at the front of the room and by raising their hands and/or fingers up, doesn’t help me to understand what’s wrong, how do they understand the material and how to I get them to know it. So I go over it again and it’s just because of my comfort level in the classroom and the students comfort with me that this is what we can do. We don’t need this little tool and strategy because it doesn’t make sense.

On the other hand, Sophie confirmed,

However, when I use specific strategies to enhance the lesson, I feel that my students are not only interested, but they are also engaged in what we are trying to learn as a class. I can honestly say that if students are offered unique and different techniques, they are more apt to understand the material and are more interested in it.

**Synopsis**

Theme 4 noted the attention of the participants in the question, “What other factors do teachers perceive to influence performance evaluations and performance-related pay?” The participants’ answers communicated that test scores matter and that student behaviors matter. Teachers have so much pressure to prepare students for the Florida Standards Assessment (FSA) test that they no longer have the flexibility to teach other important material and curricula. All of this testing preparation led the participants to say that a great deal of their evaluation rating is
dependent upon test scores and little else. The participants also relayed that the behavior of students during an observation plays a major role in the final observation result. What if the class is disruptive or if one or two students cause a distraction during the observation? All of this has a significant effect on the observation rating. As discussed earlier, the participants believe that teacher collaboration and outside involvement should make a difference in the observation rating and final evaluation score. The participants feel that so much more goes on within their professional lives that influence their whole person that it is unfair to just observe them in their classroom.

**Summary**

The purpose of this study was to conduct an in-depth investigation of the performance-based evaluation model and its relationship to teacher pay-for-performance in one selected school district. The participants were extremely supportive of this study. They all were very open, honest, and willing to share their experiences.

These findings were the result of accepted qualitative and case study research protocol for an IPA approach. Interviews were conducted to gather evidence to address the research question regarding the performance-based evaluation system in the Broward County Public Schools. Issues of validity and reliability were addressed through triangulation of the sources of evidence. A summary of findings, discussion, and implications for practice and recommendations for further research are included in the next chapter.

The overarching research question, what are teacher understandings and perceptions about the new Broward County, Florida, evaluation tool, was addressed via the exploration of four themes and their various subthemes.
The primary subthemes for Theme 1 regarding teachers’ perceptions of the evaluation tool on validity, equity, and usefulness included: (a) *It is not fair or equitable*, (b) *It is inaccurate*, (c) *Teacher impact on students is what matters*, (d) *The whole teacher is not evaluated*, (e) *Consistency is needed*, (f) *The new evaluation system is fair*, and (g) *The new evaluation system is useful if feedback is given*.

Theme 2 about self-efficacy affecting teachers’ perceptions of the new performance-based evaluation tool had four subthemes. They were: (a) *Teachers are self-efficacious and confident*, (b) *Teachers want to self-improve*, (c) *Teachers lack confidence with the tool*, and (d) *Teachers need more training, development and knowledge to feel confident*.

Theme 3 regarding the performance evaluations and performance-related pay causing stress on teachers showed two primary subthemes. These subthemes were: (a) *Teacher performance evaluations and performance-related pay cause stress on teachers* and (b) *It has caused competition among teachers*.

Finally, Theme 4, speaking to other factors that teachers may perceive as influencing performance evaluations and performance-related pay, generated five primary subthemes. These subthemes were: (a) *Test scores matter*, (b) *Student behavior matters*, (c) *Teacher collaboration matters*, (d) *What you do outside of the classroom should be considered*, and (e) *Additional approaches and tools matter*.

**Conclusion and Personal Findings**

Analyzing the qualitative material using the IPA framework in this study was an inspiring activity, although complex and time-consuming. I was able to immerse myself in the data and even stepped into the participants’ shoes. J. A. Smith et al. (2013) described the IPA approach as giving evidence of the participants’ making sense of phenomena under investigation, and at the
same time documenting the researcher’s sense making. Going into this study, I was anxious to understand the perceptions of various teachers as they struggled through their own “demons” regarding the new evaluation tool. While my initial impressions and beliefs appeared to be correct based on the data researched, there were several “Ah-ha” moments during the process.

When interviewing the group, it became apparent that most of the participants were not necessarily against performance reviews, but they were wary about the process. I was surprised to hear that five of the seven felt that administration used this tool more subjectively than objectively. They were concerned that some evaluators (administrators) used this tool to punish the teacher and not to objectively evaluate classroom practices as the tool was designed to.

Another specific take-away was that the interviewed teachers believed that the student VAM scores are unequally and unfairly weighed against them when they received their final annual evaluation score and rating. This pertained to student scores being attached to them from students who they don’t teach or even see. Unless the teacher had a significant number (60%) of their own classroom students taking the state-mandated assessment test, the teacher’s scores were based on the school average score. Every teacher, to a person, felt that this was unfair, unethical, unprofessional, and detrimental to the morale of the school. These teachers felt that they were not being judged by how well they performed their duties in the classroom, how much they positively moved student learning gains, how well they taught their curricula, or how they motivated students. They felt they were unjustly rated based on students they never saw.

This study has provided a true and honest account of Broward County School District teachers’ perspectives and views on the iObservation evaluation tool and its components. All of the study participants contributed thoroughly to the interview protocol and ensured that their perceptions were heard, communicated, and documented.
Chapter 5: Discussion, Conclusion, and Recommendations

Introduction

Chapter 5 discusses the results of this study and conclusions drawn from it. The chapter begins with a review of the problem, the purpose, and the significance of the study. A review of the research question, a discussion of the study’s methodology, and a summary of the results follow those sections. The chapter concludes with presentation of the limitations of the study, conclusions, recommendations for school districts and policy makers, and recommendations for further study. Final thoughts are presented in the concluding section of this chapter.

The purpose of this study was to determine the perceptions of teachers about the new district performance evaluation tool and its impact on pay-for-performance. Interviews were conducted with seven participants at one high school in the district to gauge awareness of the correlation between their own performance evaluations and performance pay. Having a better understanding of how teachers perceive the district evaluation tool will help educational leaders determine if implementing a pay-for-performance structure is beneficial to all educators in the district.

The findings of this study have implications in the following areas: (a) providing the school district with vast information about the effects of performance evaluations on teacher pay, (b) raising awareness about the desired and undesired effects of the new teacher evaluation system, (c) providing suggestions for mitigating negative influences of the new evaluation system, and (d) gaining powerful information that can be used to guide new change initiatives within the district.

This study provided extensive information about the effects of the evaluation tool on teachers’ pay. The central research question, what are teacher understandings and perceptions
about the new Broward County, Florida, evaluation tool, is answered through these major themes:

- Is the evaluation a valid, equitable, and useful tool for determining teacher performance-related pay?
- Is the evaluation tool affected by individual teacher self-efficacy?
- Do teacher performance evaluations and performance-related pay cause stress on teachers?
- Do other factors influence performance evaluations and performance-related pay?

**Discussion of Findings**

**Evaluation as a valid, equitable, and useful tool in determining performance pay.**

Most participants felt that the evaluation tool is not fair or equitable. This raised several questions regarding the impartiality of the tool, including that an administrator’s evaluation is subjective in that it is based on knowledge of the subject, experience, and personal attitudes toward the teacher. Research suggests that teacher acceptance of an evaluation process increases when teachers see it as being relevant to their situation, being less dependent on a single individual, having some teacher control, and being fair (K. D. Peterson, 2006). Complaints also arose about subjectivity and the potential for favoritism. A number of concerns were raised. Questions around this theme surprisingly totaled more than those regarding student performance:

What if I have a bad day? What if the principal is late and misses the best part of my lesson? What if the principal is called to an emergency on his or her way to observe my class? How can I be natural when my evaluator is sitting in the back of the room? What kind of a distraction will the principal be to the students when I am trying to conduct a lesson? What if the principal does not like me? What if I get nervous and it is not a natural lesson? These are just a few examples
of the many questions each of the participants shared. Although it takes a lot of time per teacher, from the teacher’s perspective each evaluation is just a snapshot or slice of the teacher’s work (Matula, 2010).

The participants also believed that the Broward County evaluations are inaccurate, as they don’t necessarily measure students’ real academic gains and achievement levels. Because standardized achievement tests count so heavily in the scope of the total evaluation, and because not every teacher’s students are at the same academic level or are motivated to succeed, this is an unfair advantage to those teachers with less focused and motivated students. Thus, teachers with motivated and higher achieving students will receive better and stronger evaluations, rendering these evaluations inaccurate. B. G. Davis (1993) shared “motivation is, perhaps, the most uncontrollable variable when using student performance as an evaluative criterion. Students are not dependable machines that can be counted on to perform with the same consistency day after day. They have feelings and moods and can, on the spur of the moment, decide not to try. Student moods are uncertain at any grade level, sometimes caused by an intentional desire to sabotage the system and sometimes caused by an unintentional immaturity due to the child’s age. Sometimes children give up on themselves and sometimes children are mad at the teacher for even giving them a test in the first place. This can be true especially when they are inundated with assessments that are needed by the teachers but may not be of value to the students” (p. 193-195).

This issue speaks to reliability. When students’ emotions and motivation are relied upon when taking assessment tests, the results are not reliable. Teachers have control of classroom management, lesson planning, and subject area curricula, but they often cannot control the mood
a student will have when walking into the classroom. Therefore, it is unlikely the exact same results would be obtained each time.

Finally, the teacher’s impact on students is what matters, not the rating on the teacher evaluation. Most of the participants saw teacher impact as being paramount to student achievement. Most of the teachers believed that their respect for the students and empowerment are the most important tools for ensuring that students are successful in class. To most, it was not about the actual score on the teacher evaluation, but actions and interest taken in their students that are all-important. Sanacore (2012) stated, “children and their teachers are more likely to achieve success when genuine caring is connected to learning. This connection is important because it increases the chances of success across the curriculum and through the grades” (p. 189).

The evaluation tool affected by individual and independent teacher self-efficacy. Self-efficacy is defined as people’s beliefs about their ability to produce designated levels of performance that influence events that affect their lives. Self-efficacy beliefs determine how people feel, think, self-motivate, and behave. Education is a profession in which the professionals believe that what they do, for the most part, affects how children learn. This was true of the participants of this study.

The majority of the participants noted that they felt self-efficacious and confident when in front of their students and were often in their curricular comfort zone when being observed by administration. However, several of the participants felt a certain lack of confidence during any evaluation observation because they felt that they were being judged on their classroom management and their ability to move and motivate their students. In several study findings, the instructional leadership behaviors of school principals and administrators had a positive and
significant effect on teachers’ self-efficacy. This finding was demonstrated through the review of related literature. For instance, Weisel and Dror (2006) found that school principals’ instructional leadership behaviors predicted approximately 15 percent of teachers’ self-efficacy. Similarly, Howard (1996) discussed a causal link between these two variables. Moreover, Ross (1994) expressed that leadership is an important variable in determining teachers’ self-efficacy. Weisel and Dror (2006) noted, “there is a positive and significant relationship between supportive and non-threatening leadership and teacher self-efficacy. These studies concluded that instructional leadership and honest and fair evaluative tools helped teachers with their self-efficacy and their need to feel that they can affect their students’ lives positively” (p. 162).

Another important finding from the interviews was that the majority of the participants wanted to self-improve and felt that they can and will with proper feedback and input from their evaluators. Khachatryan (2015) wrote, “teachers perceived the majority (60 percent) of feedback comments as validating or affirming of their teaching practices. Teachers appreciated getting detailed information about particular instructional practices. The more novice teachers described feedback as providing new information about their teaching” (p. 165). In addition, Khachatryan’s (2015) research confirmed that teachers talked about feedback comments, prompting them to reflect on their practice and plans to change it. As teachers read their feedback comments aloud, they talked about being prompted to reflect, pay more attention, or think differently or deeply about certain aspects of their practice.

Some of the participants felt that school-wide training and professional development workshops were helpful and supportive in creating confidence within the classroom. Because professional development is intended to positively impact instruction and enhance student outcomes, teachers who support their curricula welcome it, and thus support the performance
evaluation as well. Not only does this generate confidence and self-worth, but it also assists teachers who may not be entirely comfortable with their subject area knowledge and/or may need additional help in classroom management. Lin, Cheng, & Wu (2015) wrote, “most research suggests that professional development improves teachers’ knowledge and pedagogy and enhances teachers’ confidence in facilitating a positive attitude about student learning. Through professional growth, teachers can apply their new knowledge and skills to design and implement curriculum and enhance students’ learning ability as well” (p. 70). As Giraldo (2014) indicated, “teachers improved their classroom performance after a professional development program as their teaching became more communicative, organized, attentive to students’ needs, and principled” (p. 64).

Stress from teacher performance evaluations and performance-related pay.

Teaching can be a rewarding profession involving meaningful, influential, and important work. However, teaching can also be challenging due to the complex nature of the job. Indeed, up to one third of teachers are stressed (Geving, 2007), and similar numbers leave the profession within the first 5 years of teaching (Ingersoll & May, 2012).

The findings from this study show that classroom evaluation does cause undue stress on teachers. In addition, the connection of evaluative scores to performance pay is similarly stressful, and this stress has caused some competition among faculty and staff.

In North Carolina, nearly 30 percent of educators reported significantly high stress related to high-stakes tests and evaluations in 2013, with the majority of stress stemming from administrative (75 percent) and parent (36 percent) pressure (von der Embse, Kilgus, Bowler, Solomon, & Curtiss, 2014). These percentages mirror the concern that participants had regarding the evaluation tool. Teachers claimed that information on the evaluation tool keeps
being updated, so they never fully know the expectations when an observer comes to visit the classroom. The angst of not knowing what the teacher was being evaluated on (there are 41 elements that may be evaluated during each visit) caused stress to most of the participants in this study. Additionally, the feeling or knowledge that “big brother is watching” while teaching caused anxiety because each participant wanted to score as high as possible on the evaluation tool. Moreover, due to the number of domains available for observation during any given evaluation, most participants felt that they did not understand which observation element they are being judged on. While they wanted to do a good job and know that they were doing what was right for students, many often had a sinking feeling that they were not truly being rated on how they were teaching.

The overall evaluation performance and score, including student assessment scores at the end of the school year, is unnerving and caused undue stress to most of the participants in the study. Hutter (2004) posited, “High-stakes standardized testing is an area where many teachers are asked to meet heavy demands. Many teachers are under pressure to meet the demands placed upon them by the state. It appears that high-stakes, state-mandated testing is a mechanism of stress for teachers” (p. 27). Costigan (2002) wrote, “As a stressor, testing perhaps has become as significant as issues of survival and classroom control” (p. 9). Hutter (2004) continued, “Teacher stress due to high-stakes testing is becoming a frequent headline in the media all across the United States. High-stakes testing could guide teachers and principals in finding weaknesses or areas needing improvement in their current teaching method or education system” (p. 28). In some related cases from other school districts, these tests alone are being used to evaluate a teacher’s ability and being used to fire teachers and for pay-for-performance salary scales (Kohn, 2000). Tying teacher performance scores to high-stakes student assessment test scores (VAM)
was a stressor for the study participants.

Because a teacher’s salary may be dependent upon a high evaluation rating at the end of the school year, the evaluation process was extremely taxing and stressful for the participants. Five of the seven participants spoke specifically about this issue. The participants in this study were concerned that their financial livelihoods would be contingent upon the outcome of their evaluation scores. In a prior study conducted in Missouri by Routh (2014), participants were asked about the perceptions of teachers, administrators, and superintendents in the public school systems toward performance pay and its link to teacher evaluation. Nearly 68 percent of respondents answered, “Disagree or strongly disagree” to basing performance pay on administrator evaluations of teachers. When asked about the fairness of using performance pay to reward teachers, nearly 73 percent of respondents answered “disagree or strongly disagree” (Routh, 2014). The results from these five questions clearly indicated that participants were in line with the findings from the Routh study, as they did not support tying teacher performance pay to evaluations.

Study participants also expressed concern that tying performance evaluations to pay has caused competition among teachers. These teachers shared concerns that other teachers were not as willing to share their best practices with each other for fear that their classroom pedagogies may be used by others. Study results show that tying teacher pay to performance evaluations can create a lack of teacher/peer trust, fear of ideas being stolen, and isolation among faculty and staff.

These findings on teacher competition are consistent with previous research findings by Engstrom (1982) and others, who found that school organizations don’t give material rewards or recognition to teachers for sharing and collaboration, or for improved effectiveness that may
result from sharing. Teachers have learned not to expect rewards for sharing best practices beyond the positive feelings of collaboration with their colleagues (Engstrom, 1982). As noted by study participants in Engstrom (1982), “teachers may have little desire to share best practices or teaching strategies with others due to the fear that some of their practices might be used against them or for selfish reasons by their colleagues, which would affect their evaluation scores” (p. 33).

**Other factors that teachers perceive to influence performance evaluations and performance-related pay.** This study produced several other findings on performance evaluation that were not addressed in previous chapters. The participants were concerned that so much more goes into the evaluation than just classroom observation; each one pointedly discussed what was absent from the evaluation process.

**Test scores matter.** Never before in the history of the United States have we based so many key education policy decisions on test score outcomes. Across the United States, high stakes testing policies have caused a trickle-down effect in which politicians put pressure to increase standardized test scores on school boards and superintendents, superintendents put pressure on principals, principals on teachers, and teachers on students—all to little or no avail (Croft, Roberts, & Stenhouse, 2016). The participants in this study felt that test scores play a significant role in the final score on their performance evaluations and that, regardless of how they perform in the classroom, their students’ scores or the average school score will be the final determinant. This suggests that there is too much emphasis placed on student scores; even though the classroom teaching might be above par, the student VAM scores will often lower the teacher’s final evaluation score because of the formula used to establish it.

**Student behavior matters.** The participants strongly believed that student behavior and
classroom management play large roles in a teacher’s evaluation. Teachers are responsible for creating a classroom with a positive environment that reinforces certain values, such as respect and fairness, and makes students feel welcome and successful. Although creating a positive environment sounds simple, it is much more challenging than just being nice to students. Today, students bring many complicated issues to school. There are many students coming from home environments where they are not receiving adequate support in developing the social skills for interacting appropriately in school. Additionally, students are exposed to an increasingly violent and unpredictable society that causes stress (Brownell & Thomas, 2001). Brownell & Thomas (2001) go on to say, “The students who are coming from such environments may be apt to receive less familial support and can act out or behave inappropriately in their classes. Participants in this study felt that an evaluation is negatively impacted when students are uncooperative, unfocused, and non-participatory in class even with teacher intervention. In addition, dismissive classroom behaviors from students are often spontaneous, and students may act out just when an administrator enters the room to conduct an evaluative observation” (p. 31-32).

*Teacher collaboration matters.* The participants strongly conveyed that collaboration and sharing best practices are imperative and paramount to excellent teaching and student learning growth. They all indicated that teacher collaboration leads to stronger evaluation scores for everyone. Amato, Anthony, and Strachan (2014) wrote, “every day in classrooms, there are miracles happening, absolute miracles. Teachers are doing fantastic things. And the teacher in the classroom next door has no idea about the miraculous things that the teacher next to him or her is doing. How do we take those miracles and share them” (p. 49)?
In conjunction with the participants’ beliefs about the importance of teacher collaboration, Campbell, Lieberman, & Yashkina (2013) wrote:

“The evidence from our research is clear: These teachers and teacher leaders will do amazing things; they will initiate, innovate, implement, and share a wide range of topics and activities which can develop professional learning, improve practice, and support student learning; they will navigate personal, interpersonal, and practical challenges as their leadership is tested and grows; they will learn how to collaborate and share to spread knowledge and sustain changes; and they will demonstrate the professional, educational, and financial value of self-directed, teacher-led innovative and effective practices” (p. 59).

**Outside the classroom considerations.** While the evaluation protocol is largely dependent upon observations inside the classroom, three of the study participants discussed additional factors that should be included in the final evaluation. One factor is service performed outside the classroom, such as sponsoring afterschool activities and teaching in afterschool tutoring programs. Furthermore, these teachers travel with students to district and out-of-county tournaments and competitions that are required as part and parcel of the program and curricula. This is in addition to their responsibilities teaching students inside of the classroom. Laura stated,

In order for a fair and overall evaluative balance, it is necessary for these outside factors to be included in the evaluation process, so to provide a real-world overview of what I do as a teacher in my professional and personal work.

**Conclusion.** Recent national policy, including the 2010 federal Race to the Top (RTTT) competition, has been responsible for a dramatic overhaul of teacher evaluation systems. State
and local education agencies and districts responded by replacing traditional evaluation approaches with new systems that incorporate multiple methods of assessing teachers and formal processes to provide teachers with feedback. In recent years, districts across the nation have begun implementing these new evaluation systems with few challenges.

This study examined teachers’ perceptions of the iObservation evaluation program in a school district that has used it for the past five years. This phenomenological study may help district leaders understand teacher perceptions of new evaluation programs. The intent of this study was not to evaluate or measure the merit of the new evaluation program in use in BCPS, but rather to provide insight into the perceptions of those most affected by the new evaluation program. Similar studies of perceptions found teachers feared implementation of solely summative evaluations, feared that new evaluation systems would contribute to control and sanctions, and were opposed to the implementation of evaluation programs (Finnegan, 2016). Finnegan (2016) reports, “Teachers were very resentful that student value-added-model scores comprise half of their overall evaluation rating. Understanding teachers’ perceptions of the new evaluation program provides information relevant for designing, implementing, and maintaining effective teacher evaluation practices” (p. 57). Finnegan further states, “Understanding how teachers perceive an evaluation process that incorporates domains for teacher classroom strategies and behaviors, planning and preparing, reflection on teaching, and collegiality and professionalism with student VAM scores can help district leaders achieve greater success with evaluation practices and lead district staff to support teachers in their evaluation process so they can receive effective scores and performance pay” (p. 58-59).

**Implication of the Results for Practice**

The following implications were drawn from the findings, conclusions, and
recommendations found in this study:

- Teacher evaluations are vital to the continuous development and growth of educators across the United States. Providing professional development opportunities, relevant and adequate feedback from the evaluators, and continual classroom observations improves the professional evaluative process.

- Teacher evaluations are essential to increasing student achievement and academic performance. Teacher evaluations can drive meaningful instructional practices that will lead to an increase in student achievement while creating a positive classroom environment.

- The new teacher evaluation system in the district is used to effectively aid in the overall developmental growth of each educator through the evaluations of the 41 instructional performance standards. It is imperative that teachers are astutely aware of the dynamics of the evaluation process, so that they can be current on the practices they will be rated on during their classroom observations.

- This research implies that teachers are not specifically knowledgeable about the standards that they are evaluated on.

- Participants from this study indicated that their professional performance plays an equal part, along with student value-added-model scores, in the outcome of the evaluation and that the evaluator is documenting their performance in a subjective rather than objective manner.

- Sustainable and ongoing professional development and training should be given continuously to the teachers to aid in improving their instructional skills, along with providing refresher courses yearly to evaluators that provide updates about the
iObservation evaluation process.

- The results of this study will be useful in creating a formal and accurate documentation of administrators’ and teachers’ perceptions of the iObservation evaluation tool and also to analyzing the pros and cons of the evaluation based on the data. Based on the findings of this study and the review of this information, district school leaders will be able to make necessary adjustments to continue with this initiative.

- The results of this study can be used to support the use of iObservation. However, results from this study demonstrate a need to modify the evaluation process to better reflect a holistic assignment of scores. In addition, the iObservation tool can provide feedback to current administrators, teachers, and school leaders in the District.

Implications for Policymakers and School Districts

The results of this study indicate that this sample of Broward County, Florida, teachers generally had negative attitudes toward performance evaluations and their relationship to performance pay. The participants strongly expressed their feelings in agreement with each other, deploring the replacement of structured salaried pay scales with pay-for-performance and stating that it is hard to link the work done in schools to individual teacher performance. The participants’ negative responses to questions on these statements should inform policymakers and school administrators about the challenges the pay-for-performance model presents.

If policymakers and school districts continue implementing pay-for-performance scales, negative teacher attitudes and perceptions must be addressed. The outcomes of this study, which focused on teacher production and evaluations, included resentment among faculty, decreasing staff morale due to stress and anxiety, and discouraging collaboration. These findings differ from what the school district hierarchy currently focuses on, which is tying teachers’ evaluation
scores to performance pay. In order to move forward and establish a strong framework, proponents of pay-for-performance will have to lay strong foundations and articulate a clear rationale for continuing to promote this compensation practice. District senior management must continue to monitor the challenges that continue to stymie any and all positive teacher perceptions about tying the evaluation tool to performance pay. Furthermore, they will have to convince those in the education arena why it is good for this school district.

Senior district administrators must also understand that teachers believe pay-for-performance implementation does not improve their productivity and motivation. This study presents policymakers with evidence that teachers believe that their classroom productivity and their professional obligation to student achievement should not be tied to their salary because of the human factor in the equation. BCPS policymakers must be made aware of these findings to better understand why teachers strongly oppose any pay-for-performance compensation model.

Therefore, to build future support for models of teacher compensation that are tied to teacher evaluations, it will be critical to involve teachers in the continuing discussion and monitoring of these models. In discussing and addressing the concerns identified in this study, it will be imperative that school-based faculty positively accept what is being implemented in the district. Much can be learned about how and what to communicate during such discussions: issues to address, challenges to meet, changes to make, and attitudes of teachers to understand.

**Recommendations for Future Research**

This study produced important feedback from teachers that can be used to assist district personnel in their quest to streamline and enhance the teacher evaluation tool that is currently used in Broward County.

A deeper and richer understanding of the current situation in BCPS could be achieved
through a follow-up quantitative study that would analyze and use numerical data. A larger sample population that provides more measureable data might expose a greater number of teachers to the questions and elicit a broader range of responses.

Because this study was conducted during the early years of a new district initiative, there may not have been adequate time for teachers to have become comfortable in their exposure and experience with the new evaluation system. Repeating the study in a few years would determine if trends associated with this study change over time as both educators and supervisors gain more experience with the new evaluation process.

This research suggests that teachers’ attitudes are based on what they have independent and individual control over, and that they don’t want a program that bases incentives on evaluation factors that are out of their control. Finally, it is important to explore whether or not teachers have low levels of self-efficacy or feel a high external loss of control. Additionally, it would be beneficial to investigate if teachers find it fair that all students in the school, whether they teach them or not, should be included as part of their final evaluation score as opposed to only having their final score be based on students they have in their classrooms. Thus, future research could address the bias inherent in teacher performance evaluation ratings that are based on the ability levels of the students in this pay-for-performance salary structure.

**Personal Reflections and the Authentic Voice**

Through this research, the researcher has gained insight into how teachers perceive the evaluation process and their current role as educators. Education is a difficult and stressful profession regardless of the roadblocks that are thrown at teachers every day. There are many outside influences that affect the school environment including, but not limited to, student attitudes and behaviors, parent involvement or lack thereof, and community intervention. These
external influences don’t take into account all of the in-school requirements and regulations. Teachers face a great deal of pressure and, without mindfulness, could jeopardize their self-efficacy and professionalism. For these reasons, the study participants admitted their reluctance to buy into this pay-for-performance pay structure based on their evaluations.

The current evaluation process should receive strong teacher input and continue to be revised based on concerns and feedback from all interested parties. District personnel must be aware of teachers’ perceptions in order to gain acceptance of the pay-for-performance model. Galloway (2009) advocated for research into why teachers felt that student outcomes were not important to include in a performance pay model if that was to be the measure of how teachers excel in the classroom. “If the intention is to measure teachers who excel in the classroom, then policymakers need to take into account the teachers’ negative attitudes toward evaluations and the pressure placed upon teachers by including them” (p. 133).

While conducting this study, I struggled yet evolved through analyzing the many hours of data from participant interviews. I was pleased to hear that most of the teachers very much enjoyed their jobs and appreciated the opportunity to explore their experiences with me. I was interested to see that teachers felt burdened by the current focus on assessments as the primary measure of the success of teachers, students, and schools. While they said that they wanted their teaching quality to be validated through good test scores, they expressed concern that the current level of emphasis on testing can be overwhelming and at times discouraging.

Stiggins and Duke (1988) posited, “the paradox of teacher evaluation is that it holds the potential to help nearly every teacher improve, yet in actual practice, it helps almost no one” (p. 1). Doherty (2009) went on to say, “as school districts look for the formula that will create sustainable schools that help students achieve at their highest level, they sometimes overlook the
power that a teacher evaluation system can have on making that vision a reality. Teachers play a
critical role in the academic, social, and emotional development of students. However, in order
for them to succeed, they need the support to professionally grow as educators and to be able to
deal with the challenges that students bring to the classroom today” (p. 203). In the Broward
County Public Schools’ District, teachers want and need evaluation systems that are formative in
nature and include differentiated evaluation options, a set of clearly defined standards,
meaningful training of administrators and teachers, multiple sources of data, a supportive cultural
context, and adequate resources. Such a system can be an untapped asset that has the potential to
make substantial contributions to school and district improvement initiatives, the continuous
professional growth of educators, and student learning. It is abhorrently apparent that forces
outside of the evaluation tool that stand to damage their reputation, professionalism, and
livelihood negatively affect teachers.

In reviewing my original positionality statement, I have to say that the research
conducted and the data culled from this project did little to change my impressions of the
perceptions of Broward County schoolteachers about the evaluation process and all that is
attached to it. Teachers remain the one constant in a student’s life, and how the teacher
addresses his or her classroom ultimately affects the success of the student. Going into this
study, I was aware of how teacher evaluations influence the classroom teacher. Along with the
stress of student performance, teachers experience stress when being evaluated in the classroom
and realizing that the results of the evaluation may affect their salary. Not only was this the basis
and rationale for this study, but it also will provide valuable and tangible data to teachers,
administrators, and district personnel who must continuously strive to advance the iObservation
evaluation practices throughout the district.
Throughout this interview protocol, it had become extremely coherent that the voices of
the participants were honest, truthful, and real. Not only did the participants share personal
stories and interests, but also they addressed issues that were not comfortable for them to discuss.

As this practitioner challenges himself to segue into more of a scholar and researcher, he
can look back at these authentic voices that affirmed their truthfulness, honesty, and sheer
commitment to make a difference in discussing their concerns about performance evaluations
and performance pay. As the interviews progressed, it became more and more apparent that the
participation in this study was not only welcomed, but also needed. Each participant in this
study brought his experiences and concerns to the interview. Without considering ramifications
of what was said, the participants relied on their voices to communicate their issues surrounding
their anxieties of performance evaluations and to dispel that all is okay with the process.
Because of this research, and the results and findings of this paper, this researcher will continue
the quest of these authentic voices to conduct and maintain communication with district
personnel to revise and rework the performance evaluation process in order for it to be palatable
for all parties.
References


doi:10.1016/j.tate.2007.02.006


Harwell, S. H. (2003). *Teacher professional development: It’s not an event, it’s a process*. Waco, TX: CORD.


Sawchuk, S. (2014). Florida teacher-evaluation system found unfair, but legal; Union vows to continue fight despite ruling. *Education Week, 33*(31), 2.


Appendix A

In-Depth Interview Protocol

This interview protocol was based on the following research question:

- What are teacher understandings and perceptions of the new Broward County, Florida, evaluation tool?

I. Questions relating to the methods and purposes of evaluation.

What is your understanding of the standards on the evaluation system?

What do you need to do to meet each of the standards on the iObservation Learning Map?

What are your perceptions about the accuracy and fairness of the evaluation process you receive?

What other factors do you perceive to influence performance evaluations and performance-related pay?

Talk to me about the district’s current policy to evaluate teachers based on standardized test scores.

What role does competition between teachers play in your building?

II. Questions asking about how the evaluation system relates to your teaching practices and work behaviors.

How have your choices for professional development activities been affected by the evaluation system?

How have your teaching and work behaviors been affected by the evaluation system?

How is your own self-efficacy (the extent or strength of your belief in your own ability to complete tasks and reach goals) affected by iObservation?

Talk about how this evaluation tool affected student achievement and improvement.

What role does high-stakes testing have in effecting you professionally? Personally?

Discuss the iObservation teacher performance evaluation in regard to stress.

In what ways has the evaluation process, support, and feedback influenced your confidence in your ability to teach students?
How do classroom discipline and student behaviors impact your evaluation?

**III. Questions that speak to your perceptions of performance-based pay structures.**

What are your thoughts about teacher evaluation and pay-for-performance?

Where do you see your accountability in reference to performance-based pay structures?

How does pay for performance influence the way that you teach your students?

In regards to the evaluation tool, tell me about your sharing of best practices with your colleagues.

Talk about a relationship between pay-for-performance and stress.

What, if any, other intended or unintended impacts of tying teacher evaluation results to pay-for-performance can you address?

Is there anything else you would like to tell me about the evaluation system and performance pay?
Appendix B

INFORMED LETTER OF CONSENT

December 2015

Dear Potential Research Participant:

Before agreeing to participate in this research, you are strongly encouraged to read the following explanation of this study. This statement describes the purpose and procedures of the study. Also described is your right to withdraw from the study at any time. The Institutional Review Board of Northeastern University in Boston, MA has approved this study.

Explanation of Procedures
This study is designed to examine the perceptions of teachers in regard to the iObservation evaluation tool and its relationship to performance pay. This study will be conducted to gather information about your feelings and beliefs since this topic has not been studied much in the past. Participation in the study involves an in-depth, one-on-one interview, between the researcher and you that asks you to speak about your experiences with the evaluation tool and your feelings about the potential challenges that may come with it. With the results of this research, we hope to inform the School Board, teacher organizations, and school leaders in the development and/or revision of policies and procedures for awarding performance pay based on teacher evaluations.

Time
This research will require approximately 2 hours of your time. During this time, you will be interviewed about your experiences with the iObservation evaluation tool. The interviews will be conducted at a location of choice, by mutual agreement, and will be audio tape-recorded.

Risks, Discomforts, and Benefits
There are no risks or discomforts that are anticipated from your participation in this study. However, a potential risk or discomfort may include possible emotional feelings when asked questions during the interview.

You may also find the interview to be very enjoyable and rewarding, as many teachers do not get to share their experiences with a skilled and nonjudgmental interviewer, as you will. By participating in this research, you may also benefit others by helping teachers to better understand the rationale and vision behind this educator evaluation tool.

Confidentiality
Several steps will be taken to protect your anonymity and identity. The information gathered during this study will remain confidential in secure premises during this project. Only the
researcher will have access to the study data and information. There will not be any identifying names on the interview transcripts; they will be coded and the key to the code will be kept locked away. Your names and any other identifying details will never be revealed in any publication of the results of this study. The tapes will be destroyed at the completion of the study. The results of the research will be published in the form of a doctoral paper (dissertation) and may be published in a professional journal or presented at professional meetings. It may also be published in book form.

Withdrawal without Prejudice
Participation in this study is voluntary; refusal to participate will involve no penalty. You are free to withdraw consent and discontinue participation in this project at any time without prejudice or penalty. You are also free to refuse to answer any question(s) asked of you.

Further Questions and Follow-Up
You are welcome to ask the researcher any questions that occur to you during the survey or interview. If you have further questions once the interview is completed, you are encouraged to contact the researcher, Paul Satty, at 954-993-4800.

If you have any questions about your rights in this research, you may contact Nan C. Regina, Director, Human Subject Research Protection, 490 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.4588, Email: n.regina@neu.edu. You may call anonymously if you wish.

I have read (or have been read) the above information regarding this research study on the experiences of this County’s evaluation tool, and consent to participate in this study. I freely agree to participate in this study. I understand that I am free to refuse to answer any question and to withdraw from the study at any time. I understand that my responses will be kept anonymous.

__________________________________________________________ (Printed Name)
__________________________________________________________ (Signature)
__________________________________________________________ (Date)
Appendix C

Invitation Letter to Interview Participants

December 2015

Mr. Joe Smith
123 Main Street
Ft. Lauderdale, FL  33333

Dear Mr. Smith,

I am a doctoral student under the direction of Dr. Billye Sankofa Waters in the Education Leadership program of the College of Professional Studies at Northeastern University. I am conducting a research study to analyze the perceptions and understandings of teachers about the Broward County evaluation tool. This study could help Broward District educators, administrators, and policy makers better understand the effect that the evaluation tool is having on pay for performance.

I am requesting your participation, which will involve participating in an in-depth interview. The interview will take approximately one hour. Your participation in this study is voluntary. If you choose not to participate or to withdraw from the study at any time, there will be no penalty. The results of the research study may be published, but your name will not be used and the information collected from the interview will be completely anonymous. It will be used for the sole purpose of identifying ways to improve the quality of Marzano’s iObservation evaluation tool.

Although there may be no direct benefit to you, the possible benefit of your participation is vital to better understand the potential effects of Broward’s evaluation tool to pay for performance. The results of this research will inform the School Board, teacher organizations, and school leaders in the development and/or revision of policies and procedures for awarding performance pay based on teacher evaluations.

If you have any questions concerning the research study, or you wish to volunteer for this study, please email my student address at satty.p@husky.neu.edu. Please do not call my office or email my school address.

Sincerely,

Paul Satty
Northeastern University, Boston, Massachusetts
satty.p@husky.neu.edu
Appendix D

Principal Request Letter

November 2015

Principal
ABC High School
123 Main Street
Ft. Lauderdale, FL  33337

Dear Principal,

My name is Paul Satty and I am a doctoral degree (EdD) student at Northeastern University. I am conducting research on the iObservation evaluation tool under the supervision of Dr. Billye Sankofa Waters. As an assistant principal at Coral Glades High School, I have been afforded the opportunity to work with the evaluation tool for the past five years.

This study is designed to examine the perceptions of teachers in regard to the iObservation evaluation tool and its relationship to performance pay. This study will be conducted to gather information about teachers’ feelings and beliefs since this topic has not been studied much in the past.

My data collection method will include audiotaping teacher interviews. I intend to invite my teacher colleagues to collaborate on their perceptions on the iObservation evaluation tool and its relationship to pay-for-performance. I will guarantee confidentiality of all information obtained from the interviews. I will not reveal anything of a personal or compromising nature. Participation in this research is completely voluntary.

I am requesting your permission to interview teachers at your school for my study. Once I have received your consent to approach educators to participate in the study, I will arrange for informed consent to be obtained from the participants. Thank you for support and consideration.

Educationally,

Paul Satty

______________________________  ______________________________
Principal Signature          Principal Name
Appendix E

Table 1

*Study participants by grade level, subject, and years of service*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Sex</th>
<th>Grade Level</th>
<th>Subject ParticipantTeaches</th>
<th>Years of Service in Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>(David)</td>
<td>Male</td>
<td>9th, 10th, 11th, 12th</td>
<td>Marketing/DECA (CTE)</td>
<td>3</td>
</tr>
<tr>
<td>(Nina)</td>
<td>Female</td>
<td>11th</td>
<td>English/Reading</td>
<td>13</td>
</tr>
<tr>
<td>(Anne)</td>
<td>Female</td>
<td>10th</td>
<td>Social Studies</td>
<td>17</td>
</tr>
<tr>
<td>(Eunice)</td>
<td>Female</td>
<td>11th, 12th</td>
<td>Spanish</td>
<td>38</td>
</tr>
<tr>
<td>(Melissa)</td>
<td>Female</td>
<td>9th, 10th, 11th, 12th</td>
<td>French</td>
<td>34</td>
</tr>
<tr>
<td>(Laura)</td>
<td>Female</td>
<td>10th, 12th</td>
<td>English</td>
<td>19</td>
</tr>
<tr>
<td>(Sophie)</td>
<td>Female</td>
<td>9th, 10th</td>
<td>Math</td>
<td>1</td>
</tr>
</tbody>
</table>
## Appendix F

Table 2

*Subthemes and Definitions for Theme 1*

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is not fair or equitable</td>
<td>The new evaluation system is not fair or equitable as a tool for determining their performance-related pay.</td>
</tr>
<tr>
<td>It is inaccurate</td>
<td>The new evaluation system is inaccurate as a tool for determining their performance-related pay.</td>
</tr>
<tr>
<td>Teacher personal influence on students is what matters</td>
<td>Rather than focus on pay for performance, a teacher’s influence on students is what matters.</td>
</tr>
<tr>
<td>The whole teacher is not evaluated</td>
<td>The new evaluation system is rigid and does not assess the whole teacher.</td>
</tr>
<tr>
<td>Consistency is needed</td>
<td>Consistency is needed in the use of the new evaluation system as a tool for determining their performance-related pay.</td>
</tr>
<tr>
<td>The new evaluation system is fair</td>
<td>The new evaluation system is fair as a tool for determining their performance-related pay.</td>
</tr>
<tr>
<td>The new evaluation system is useful if feedback is given</td>
<td>The new evaluation system is useful as a tool for determining their performance-related pay if feedback is provided to teachers.</td>
</tr>
</tbody>
</table>
## Appendix G

### Table 3

*Frequency of Subthemes for Theme 1*

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Number of participants mentioning this subtheme</th>
<th>Total exemplar quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is not fair or equitable</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>It is inaccurate</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Teacher impact on students is what matters</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>The whole teacher is not evaluated</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Consistency is needed</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>The new evaluation system is fair</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>The new evaluation system is useful if feedback is given</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 4

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers are self-efficacious and confident</td>
<td>Teachers are self-efficacious and confident in regard to the new performance-based evaluation tool.</td>
</tr>
<tr>
<td>Teachers want to self-improve</td>
<td>Teachers want to self-improve to meet the requirements and standards of the new performance-based evaluation tool.</td>
</tr>
<tr>
<td>Teachers lack confidence</td>
<td>Teachers lack confidence in regard to the new performance-based evaluation tool.</td>
</tr>
<tr>
<td>Teachers need more training, development and knowledge to feel confident</td>
<td>Teachers need more training, professional development, and knowledge to feel confident in regard to the new performance-based evaluation tool.</td>
</tr>
</tbody>
</table>
Appendix I

Table 5

Frequency of Subthemes for Theme 2

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Number of participants mentioning this subtheme</th>
<th>Total exemplar quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers are self-efficacious and confident</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Teachers want to self-improve</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Teachers lack confidence</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Teachers need more training, development and knowledge to feel confident</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
Appendix J

Table 6

Subthemes and Definitions for Theme 3

<table>
<thead>
<tr>
<th>Subtheme/Topic</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher performance evaluations and performance-related pay cause stress on</td>
<td>The classroom observation and evaluation specifically causes stress for teachers.</td>
</tr>
<tr>
<td>teachers.</td>
<td></td>
</tr>
<tr>
<td>The classroom observation and evaluation causes stress</td>
<td>The connection of the performance evaluations to teacher pay and salary is stressful to teachers.</td>
</tr>
<tr>
<td>The connection to pay is stressful</td>
<td></td>
</tr>
<tr>
<td>The overall performance evaluation is stressful</td>
<td>In general, the teacher performance evaluations are stressful.</td>
</tr>
<tr>
<td>It has caused competition among teachers</td>
<td>Teacher performance evaluations and performance-related pay has caused competition among some teachers.</td>
</tr>
</tbody>
</table>

Note. Topics are italicized.
Table 7

*Frequency of Subthemes for Theme 3*

<table>
<thead>
<tr>
<th>Subtheme/Topic</th>
<th>Number of participants mentioning this subtheme/topic</th>
<th>Total exemplar quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher performance evaluations and performance-related pay cause stress on teachers.</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>The classroom observation evaluation causes stress</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>The connection to pay is stressful</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>The overall performance evaluation is stressful</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>It has caused competition among teachers</td>
<td>6</td>
<td>10</td>
</tr>
</tbody>
</table>

*Note.* Topics are italicized.
## Appendix L

### Table 8

**Subthemes and Definitions for Theme 4**

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test scores matter</td>
<td>Students’ standardized test scores influence performance evaluations and performance-related pay.</td>
</tr>
<tr>
<td>Student behavior matters</td>
<td>Students’ behavior during the classroom observation and evaluation influences performance evaluations and performance-related pay.</td>
</tr>
<tr>
<td>Teacher collaboration matters</td>
<td>Teachers’ ability to collaborate with other teachers influences performance evaluations and performance-related pay.</td>
</tr>
<tr>
<td>What you do outside of the classroom should be considered</td>
<td>What you do outside of the classroom should be considered in performance evaluations and performance-related pay.</td>
</tr>
<tr>
<td>Additional approaches and tools matter</td>
<td>Additional approaches and tools that impact the quality of teaching influence performance evaluations and performance-related pay.</td>
</tr>
</tbody>
</table>
Appendix M

Table 9

*Frequency of Subthemes for Theme 4*

<table>
<thead>
<tr>
<th>Subtheme</th>
<th>Number of participants mentioning this subtheme</th>
<th>Total exemplar quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test scores matter</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Student behavior matters</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Teacher collaboration matters</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>What you do outside of the classroom should be considered</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Additional approaches and tools matter</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Appendix N

Domain 1: Classroom Strategies and Behaviors
Domain 1 is based on the Art and Science of Teaching Framework and identifies the 41 elements or instructional categories that happen in the classroom. The 41 instructional categories are organized into 9 Design Questions (DQs), and further grouped into 3 lesson segments to define the Observation and Feedback Protocol.

Notes: DQ refers to Design Question in the Marzano Art and Science of Teaching Framework. The 9 DQs organize the 41 elements in Domain 1.

The final Design Question, DQ10: Developing Effective Lessons Organized into a Cohesive Unit, is contained in Domain 2: Planning and Preparing.