

University of Central Florida

STARS

Electronic Theses and Dissertations

2017

An Investigation of Low Performing Central Florida Teacher Evaluation Feedback and Improvement Plans as Related to Value-Added Model Scores and Instructional Practice Scores

Tara Butler

University of Central Florida



Part of the [Educational Administration and Supervision Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Butler, Tara, "An Investigation of Low Performing Central Florida Teacher Evaluation Feedback and Improvement Plans as Related to Value-Added Model Scores and Instructional Practice Scores" (2017). *Electronic Theses and Dissertations*. 5660.
<https://stars.library.ucf.edu/etd/5660>

AN INVESTIGATION OF LOW PERFORMING CENTRAL FLORIDA TEACHER
EVALUATION FEEDBACK AND IMPROVEMENT PLANS AS RELATED TO
VALUE-ADDED MODEL SCORES AND INSTRUCTIONAL PRACTICE SCORES

by

TARA LYNN BUTLER

B. A. University of Connecticut, 2010

B. S. University of Connecticut, 2010

M. A. University of Connecticut, 2011

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the School of Teaching, Learning, and Leadership
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Spring Term
2017

Major Professor: Barbara A. Murray

© 2017 Tara Lynn Butler

ABSTRACT

The purpose of this study was to determine if low performing Central Florida teachers, according to Florida VAM scores, had related instructional practice evaluations that were being used to determine placement on improvement plans in the 2013-14 school year. Additionally, the feedback held within the instructional practice evaluations was analyzed to discover the levels of feedback most frequently rendered by administrative evaluators to the lowest performing teachers. Deidentified data from a population of 528 Central Florida teacher evaluations and improvement plans within the lowest 10% of VAM scores for the 2013-14 school year were gathered from a Florida Department of Education database and public record requests. Data were analyzed to determine if any significant relationship existed between VAM scores and instructional practice scores. A very weak relationship existed between these two variables. Tenured teachers were rated significantly more favorably on summative instructional practice evaluations than nontenured teachers. Within the population, fewer than 1% of low-performing teachers (two total) were prescribed an improvement plan, regardless of tenure or nontenure status. Finally, evaluation feedback was largely low-level without reference to student growth or achievement.

Evidence of administrative barriers within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults were inferred from the literature and subsequent findings. To improve upon the current evaluation system, administrators must be aware of, and well-prepared for the demands of evaluating, remediating, and providing feedback to teaching professionals

relative to student achievement and growth in a manner that is simultaneously respectful, candid, fair, timely, and actionable.

ACKNOWLEDGMENTS

It has become apparent throughout this journey that not one person earns any achievement alone. There are a great many people who have motivated me, mentored me, been my greatest friends, and remained my most steadfast companions. To all who have had a hand in this journey, who have marked on my map, and who have helped direct the steering wheel, who carried me when I was weary- I am forever grateful to you.

To my parents, Roger and Lynn Butler: who always have been my greatest champions, who have made me into the woman I am, who have shaped and molded my character, who have believed in me and my choice to become a public educator- despite your initial caution in my choosing of this profession- I love you.

To my dissertation chair, Dr. Barbara Murray: who I attribute to allowing me to become the educational leader I was always meant to be, for showing me what it means to be brave, for showing me how to stand up for what is true- I am inspired by you.

To my dissertation committee- Dr. Lee Baldwin, Dr. Walter Doherty, Dr. Cynthia Hutchinson, and Dr. Kenneth Murray: your guidance and thoughtful examination of my work on one of the most important journeys of my life has been the opportunity of a lifetime- I am in awe of you.

To my principal, Dr. Cheryl Salerno: for providing me the forum and opportunity to initiate thought into practice, for being a trailblazer in the educational community of Volusia County, for being “the woman in the arena,” for being concurrently bold, passionate, and kind- and for always considering the prime interest of students above all else- I admire you.

TABLE OF CONTENTS

LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS AND ACRONYMS	x
CHAPTER 1 INTRODUCTION	1
Background of the Study	1
Statement of the Problem.....	5
Purpose of the Study	6
Significance of the Study	6
Definition of Terms.....	9
Theoretical Framework	14
The Social Systems Theory of Administrative Behavior	14
Nomothetic and Idiographic: Competing Dimensions	15
Role-Personality, Role, and Personality Conflicts.....	17
Effectiveness, Efficiency, and Satisfaction.....	20
Leadership and Followership Styles	22
Morale.....	23
Research Questions and Hypotheses	24
Limitations	25
Delimitations.....	27
Organization of the Study	28
CHAPTER 2 REVIEW OF LITERATURE	29
Introduction.....	29
Understanding the Context of Teacher Evaluation in the Current Era of Reform	30
The Lake Wobegon Effect.....	32
The Case for Using Value-Added Modeling to Rate Teacher Performance	35
Barriers to Candid Teacher Evaluation Practices	37
Individual Conflict Avoidance.....	38
Bureaucratic Procedural Interferences	43
Administrative Procedural Faults	48
Summary	53
CHAPTER 3 METHODOLOGY	54
Introduction.....	54
Purpose of the Study	54
Research Questions	55
Selection of Subjects and Population.....	56
Instrumentation	58
Florida's Value-Added Model (VAM)	58

Florida Instructional Personnel Evaluation Requirements	60
Florida Instructional Personnel Contract Types	62
Levels of Feedback	64
Data Collection	65
Data Analysis	66
Summary	68
CHAPTER 4 PRESENTATION AND ANALYSIS OF DATA.....	69
Introduction.....	69
Overall Descriptive Statistics.....	70
Research Question 1	71
Research Question 2	79
Research Question 3	81
Research Question 4	83
Summary	85
CHAPTER 5 SUMMARY, DISCUSSION, AND RECOMMENDATIONS.....	90
Introduction.....	90
Summary of the Study	90
Discussion of the Findings.....	93
Research Question 1	93
Research Question 2	97
Research Question 3	99
Research Question 4	102
Implications for Policy and Practice	106
Recommendations for Future Research	109
Summary	112
APPENDIX A UCF INSTITUTIONAL REVIEW BOARD EXEMPTION OF RESEARCH.....	114
APPENDIX B SAMPLE LETTER FOR PUBLIC RECORDS REQUESTS.....	116
APPENDIX C SAMPLE RUBRIC FOR FEEDBACK DETERMINATION RATING	118
LIST OF REFERENCES	120

LIST OF FIGURES

Figure 1. Florida Value-added Example of Score Determination	59
Figure 2. Histogram of Value-added Model (VAM) Score Frequencies Observed	73
Figure 3. Histogram of Instructional Practice Score Frequencies Observed	75
Figure 4. Scatterplot of Value-added (VAM) Scores as Related to Instructional Practice Scores	78
Figure 5. Pie Chart Representing Percentage of Teachers Prescribed an Improvement Plan	82

LIST OF TABLES

Table 1 Characteristics of Study Population.....	57
Table 2 Research Questions, Research Methods, Variables, Data Sources, and Data Analysis.....	67
Table 3 Pre- and Post-Data Collection Frequencies of Study Teachers by County: 2013-14.....	71
Table 4 Frequencies and Percentage of Teachers by Instructional Practice Rating Category	74
Table 5 Pearson Correlation Coefficients: VAM Scores and Instructional Practice Scores	76
Table 6 Frequency and Percentage of Teachers by Contract Status	79
Table 7 Mean and Standard Deviation for Instructional Practice Scores of Tenured versus Nontenured Teachers	80
Table 8 Independent Samples Test Comparing Mean Instructional Practice Scores for Tenured vs. Nontenured Teachers	81
Table 9 Frequencies and Percentages of Highest Feedback Levels Observed	85
Table 10 Research Questions, Variables, Data Sources, Methods of Analysis & Results	88
Table 11 Comparison of Prior Research as Related to Teacher Effectiveness Ratings....	95

LIST OF ABBREVIATIONS AND ACRONYMS

AC:	Annual Contract
CC:	Continuing Contract
CFPSBC:	Central Florida Public School Board Coalition
FDOE:	Florida Department of Education
LIFO:	Last In First Out
NCLB:	No Child Left Behind
NGA:	National Governors Association
PC:	Probationary Contract
PSC:	Professional Service Contract
RTTT:	Race to the Top
SAT:	Stanford Achievement Test
VAM:	Value-Added Model

CHAPTER 1 INTRODUCTION

Background of the Study

The National Commission on Teaching and America's Future proposed "Within a decade- by the year 2006- we will provide every student in America with what should be their educational birthright: access to competent, caring, and qualified teaching in schools organized for success" (National Commission on Teaching and America's Future, 1996, p. 10). The commission found that the teaching profession had been greatly neglected and required restructuring at its foundation, deeming that effective and quality teachers, the most important ingredient in education reform, was most often overlooked (Darling-Hammond, 1996). Three years after this goal was supposed to be realized, Race to the Top required in 2009 that participating states begin rewarding both highly effective teachers and administrators using a value-added model (VAM) of student achievement in order to define and quantify teacher effectiveness (Harris, Ingle, & Rutledge, 2014, p. 74). In the state of Florida, Senate Bill 736 (also known as the Student Success Act, which became Florida Statute §1012.34) was passed in 2011 and required VAM scores to account for 50% of the overall teacher evaluation, with local instructional practice scores determined by school administrators to comprise the remaining 50% (Florida State Senate Bill 736, 2011). These initiatives were passed with the understanding that student achievement measures are important factors when evaluating teachers and making high-stakes decisions about teachers' careers (Harris et al., 2014). Florida Statute §1012.34 was subsequently revised and, at the time of the present study, required that at least one-third of the overall teacher performance evaluation be based upon data indicators of

student performance via VAM scores. At least one-third was required to be based on an instructional practice score decided by school-based evaluators. The final one-third may be compiled considering other indicators of performance. These indicators included professional and job responsibilities recommended by the State Board of Education or identified by the district school board, peer reviews, objectively reliable survey information from students and parents based on teaching practices that were associated with higher student achievement, and other valid and reliable measures of instructional practice (Florida Statute §1012.34, 2015).

The use of VAM scores to determine teacher effectiveness is a fairly new idea grounded in the belief that teacher effectiveness is directly and significantly correlated to student achievement. Although instructional practice evaluations have been in place for decades, there is little evidence to support that these evaluations reliably identify ineffective teachers with respect to student achievement outcomes. For example, in a study conducted in Pennsylvania, 98.2% of teachers received a perfect evaluation score when districts used the Pennsylvania State Education Department's standard rating form (Langlois & Colarusso, 1988). Specific to Florida, Mela (2013) found, within a population of 1,138 teachers in Brevard County, that 99% were rated effective or highly effective. More recently, Pace (2015), in a similar Brevard County study, found that administrators rated 92.5% of all teachers as effective or highly effective on the professional practices portion of the county's state-approved teacher evaluation instrument. These local teacher evaluation scores were in stark contrast to mounting research indicating the average number of incompetent teachers to be between 5% and

15% (Tucker, 2001). In May 2012, nine California public school students filed a statewide lawsuit that struck down five laws governing tenure, dismissal, and layoff procedures that were said to protect approximately 3% of “grossly ineffective teachers” (*Vergara v. California* Final Judgment, 2014, p. 8). The final judgment in this case surmised “...the number of grossly ineffective teachers has a direct, real, appreciable, and negative impact on a significant number of California students, now and well into the future, for as long as said teachers hold their positions” (*Vergara v. California* Final Judgment, 2014, p. 8). This contrast has led researchers to suspect that inflated teacher evaluations are a common and strong indicator that teacher supervision and evaluation are dysfunctional systems that do not appropriately address poor teacher performance (Langlois & Colarusso, 1988).

The Lake Wobegon Effect is characterized by “a phenomenon in which most individuals or groups perform above average” (Wheeler & Haertel, 1993, p. 82). In this case, the Lake Wobegon Effect is manifested when nearly all teachers are deemed effective or highly effective by their performance evaluation despite contradictory evidence displayed by low student achievement measures (Tucker, 1997; Wheeler & Haertel, 1993). Three key dimensions have surfaced from the literature as contributors to the Lake Wobegon Effect of inflated teacher performance evaluations: (a) individual conflict avoidance, (b) bureaucratic procedural interferences, and (c) administrative procedural faults.

First, individual conflict avoidance has been defined as an administrative personality flaw or lack of will that causes an avoidance of conflict and discomfort when

conducting face-to-face performance critiques necessary to achieve the institutional goal of increasing student achievement (Donaldson, 2010; Getzels & Guba, 1957). Second, bureaucratic procedural interferences occur when state and local district policies and personnel do not fully support school-based administrative actions and decisions intended to tackle ineffective performance (Donaldson, 2010; Getzels & Guba, 1957). Finally, administrative procedural faults are credited for instances in which administrators do not accurately and appropriately document teacher instructional practice with fidelity, leading to grievances and costly legal union battles when administrators do take action to dismiss ineffective teachers. It is important to note that the key difference between individual conflict avoidance and administrative procedural faults is that, though individual conflict avoidance translates to a lack of will to relay high-quality and candid feedback on performance, administrative procedural faults account for a lack of skill in the delivery of high-quality feedback (Donaldson, 2010).

When school administrators fail to make the required decisions to eliminate ineffective teachers, the administration becomes the critical factor behind poor student achievement measures (Fuhr, 1993). Generally, across public school systems, accountability is one-directional: from the administrator to the teacher (Futernick, 2010). However, Elmore (2002) proposed a more promising approach, stating that administrators should be equally and reciprocally responsible for teacher performance accountability, effectively raising the bar for all school personnel to develop capacity for increased human capital:

For every increment of performance, I demand from you, I have an equal responsibility to provide you with the capacity to meet that expectation.

Likewise, for every investment you make in my skill and knowledge, I have the reciprocal responsibility to demonstrate some new increment in performance. This is the principle of 'reciprocity of accountability for capacity.' It is the glue that, in the final analysis, will hold accountability systems together. (p. 5)

The primary role of evaluation is to provide evidence for administrators in making a binary decision: to recommend for teacher retention or non-retention (Range, Duncan, Scherz, & Haines, 2012). It is the responsibility of school-based evaluative administrators to ensure teacher effectiveness within schools and take actions to remediate or dismiss ineffective teachers using student achievement data. When the role of evaluation is compromised, evaluations become far less meaningful and impactful to the central purpose of a school as an institution: cultivating student achievement. The three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults acting in tandem manifest the Lake Wobegon Effect within a school system. They reinforce ineffective teaching and administrative behaviors that contribute to a sub-par teacher workforce with low student growth and achievement outcomes.

Statement of the Problem

To date, limited research has been conducted on how Florida value-added model (VAM) student achievement data, instructional practice evaluation ratings, feedback, and improvement plans have been used to determine and remediate teacher performance. Additionally, there is little known regarding how barriers to candid and appropriate feedback, (i.e., individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults), contribute to administrative decisions to assign teachers

with low VAM scores to improvement plans and subsequently pursue the dismissal of such ineffective teachers.

Purpose of the Study

The purpose of this study was to determine if grossly ineffective teachers, according to Florida VAM scores, had related instructional practice evaluations that were being used to determine placement on improvement plans. Additionally, this study served to determine if there was a difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teacher instructional practice scores and assignment to improvement plans for teachers within the lowest 10% of VAM scores in Central Florida school districts. Finally, the levels of feedback held within the local district performance evaluations and improvement plans were reviewed to search for evidence of administrative barriers within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults that exist and inhibit administrative evaluators from providing candid and actionable teacher performance evaluations relative to student growth and achievement.

Significance of the Study

Federal legislation mandated in Race to the Top that participating states use VAM measures of student achievement to determine a portion of overall teacher and administrative effectiveness ratings (Harris et al., 2014). Although teachers have the greatest impact on student achievement, administrators also have an immense effect on student success with regard to making decisions about which teachers are hired and

allowed to maintain their position in the classroom (Range et al., 2012). As a result of the accountability movement, there has been an increased need for specific research on the leadership actions of school administrators and the use of teacher evaluation systems (Ovando & Ramirez, 2007). Tucker (1997) stated: “Principals have the ultimate responsibility to ensure minimal standards of teacher competency” (p. 105). Following this logic, administrators have been charged with the instructional and ethical leadership task of building human capital by hiring and firing instructional personnel, and facilitating remediation methods for teachers to improve their practice (Range et al., 2012, p. 303).

However, Weisberg, Sexton, Mulhern, Keeling, & New (2009) proclaimed: “A teacher’s effectiveness- the most important factor for schools in improving student achievement- is not measured, recorded, or used to inform decision making in any meaningful way” (p. 31). As VAM scores in Florida have been released since the 2011-12 school year and used to contribute to overall individual teacher evaluation scores, a more quantitative measure has been introduced with respect to teacher effectiveness, providing more insight as to who the teachers are that may be considered grossly ineffective according to student growth and achievement outcomes.

Historically, instructional practice evaluations have been conducted with little distinction “...between great teaching from good, good from fair, and fair from poor” based on short, infrequent classroom observations (Weisberg et al., 2009, p. 3). The outcome has been, “...on paper, almost every teacher is a great teacher” (Weisberg et al., 2009, p. 3). Many researchers have sought to find out why ineffective teachers remain in

the classroom long after they should have been rightfully dismissed. Fuhr (1993) declared, "Principals don't like to talk about the marginal or incompetent teacher. No one likes to admit these [teachers] exist in a school" (p. 26). Moreover, taking action against grossly ineffective teachers who exist within school systems requires "courage, honesty, knowledge, and hard work" on behalf of all administrative personnel involved (Staples, 1991, p. 142). Bridges (1993) found in a study of school principals "About 30-45 percent of administrators will not confront a bad teacher and tell them they are doing a bad job" (p. 36). Furthermore, Mock and Melnick (1991) surveyed principals and found that 50% would, if permitted, replace between 3% and 10% of their staff. However, 94% of the principals surveyed reported giving less than 2% of teachers an unsatisfactory rating. More compelling were the findings of Tucker (1997) who determined a comprehensive teacher evaluation did not ensure that a principal would address and respond to incompetence. According to Brieschke (1986), most of the principals included in a study reported that they "...tried to avoid involving themselves in the lengthy, time-consuming, complicated, and tension-producing procedure of removing a teacher from the school because the process often did not end in the desired result" (p. 244).

Ineffective teachers remaining in the classroom is, in part, a result of a dysfunctional system that permits, and even compels, the leadership within schools to continue providing low-level, non-confrontational drive-by feedback on teacher performance evaluations (Langlois & Colarusso, 1988). This leaves accountability systems ineffective, empty, ritualistic practices that rarely mark teachers as ineffective,

and less often prescribe improvement plans or dismissal despite mounting student growth and achievement evidence (Langlois & Colarusso, 1988).

The present study was significant as it aimed to provide insight as to how incompetence observed by Florida teacher VAM scores was correlated to local district instructional practice evaluations conducted by school administrators. The researcher also sought to provide deeper understanding and recommendations as to what could be done to ensure administrators were using evaluative tools appropriately to increase teacher performance and accountability, thereby providing a system and environment that ensured a competent teacher is in every classroom. Overall, the goal of this study was to discover if lack of leadership was an indicator of a dysfunctional system that permits ineffective teachers to remain in the classroom.

Definition of Terms

Annual Contract: A type of teaching contract in the state of Florida, which are the only contracts that may be offered to instructional personnel hired after July 1, 2011 for a period of one school year, in which the district school board may choose to award or not to award without cause (Florida Statute §1012.335(1), 2015). For the purpose of this study, the term “annual contract” may be used interchangeably with “nontenured.”

Dismissal: Refers to “...the termination for cause of any tenured teacher or a probationary teacher within the contract period” (Cambron-McCabe, McCarthy, & Thomas, 2004, p. 286).

Classroom Teachers: Staff members assigned the professional activity of instructing students in courses in classroom situations, including basic instruction,

exceptional student education, career education, and adult education, including substitute teachers (Florida Statute §1012.01, 2015).

Continuing Contract: A type of teacher contract that may be held by any employee in the state of Florida who held continuing contract status prior to July 1, 1984 (Florida Statute §1012.33(4)(a), 2016). Any Florida teacher holding continuing contract status prior to July 1, 1984 shall be entitled to retain such contract and all rights arising therefrom as prescribed by the State Board of Education, unless the employee voluntarily relinquishes his or her continuing contract (Florida Statute §1012.33(4)(a), 2016). Any member of the district administrative or supervisory staff and any member of the instructional staff, including the school principal, who is under continuing contract may be suspended or dismissed at any time during the school year; however, the charges against him or her must be based on immorality, misconduct in office, incompetency, gross insubordination, willful neglect of duty, drunkenness or being convicted or found guilty of, or enter a plea of guilty to, regardless of adjudication of guilt, any crime involving moral turpitude, as these terms are defined by rule of the State Board of Education (Florida Statute §1012.33(4)(c), 2016). For the purpose of this study, the term “continuing contract” may be used interchangeably with “tenured.”

Feedback: “Information about how we are doing in our efforts to reach a goal” (Wiggins, 2012, p. 11). Furthermore, feedback must “...provide information specifically relating to the task or process of learning that fills the gap between what is understood and what is aimed to be understood” (Hattie & Timperley, 2007, p. 82).

Grossly ineffective: Term used in the *Vergara v. California* (2014) final judgment to describe roughly one to 3% of teachers in the state of California that substantially undermine the ability of a child to succeed in school. Furthermore, the final judgment of this case declared “...Considering the effect of grossly ineffective teachers on students, as indicated above, it therefore cannot be gainsaid that the number of grossly ineffective teachers has a direct, real, appreciable, and negative impact on a significant number of California students, now and well into the future, for as long as said teachers hold their positions” (*Vergara v. California* Final Judgment, 2014, p. 8). For the purpose of this study, grossly ineffective teachers are classroom teachers with the lowest 10% of VAM scores in 13 Central Florida school districts, as observed by the Florida Department of Education database for 2013-2014.

Instructional Leadership: The primary role by which school leaders impact student success: by helping teachers improve their practice (Range et al., 2012). A series of administrative behaviors that include: making suggestions, giving feedback, modeling effective instruction, soliciting opinions, supporting collaboration, providing professional development activities, and rendering praise for effective teaching (Blase & Blase, 2000).

Instructional Practice Scores/Ratings/Feedback: Performance evaluations that are conducted by school-based administrative personnel using a system that is state-approved for rating teacher instructional performance (Florida Statute §1012.34, 2015). Each evaluation system must be designed to support effective instruction and student learning growth, and provide appropriate instruments, procedures, timely feedback, and criteria for continuous quality improvement of the professional skills of instructional personnel

(Florida Statute §1012.34, 2015). These performance evaluations require a summative rating of (a) highly effective, (b) effective, (c) needs improvement (or, for instructional personnel in the first three years of employment who need improvement, developing), or (d) unsatisfactory (Florida Statute §1012.34, 2015).

Probationary Contract: A type of teaching contract which may be offered for a period of one school year to instructional personnel upon initial employment in a school district (Florida Statute §1012.335 (1)(c)). For the purpose of this study, the term “probationary contract” may be used interchangeably with “nontenured.”

Professional Service Contract: A type of teaching contract which was only offered to instructional personnel hired prior to July 1, 2011 and shall be renewed each year unless the district school superintendent, after receiving recommendations required by Florida Statute §1012.34, (a) charges the employee with unsatisfactory performance and notifies the employee of performance deficiencies as required by the statute; (b) the employee receives two consecutive annual performance evaluation ratings of unsatisfactory within a three year period; or (c) three consecutive performance evaluations of needs improvement or a combination of needs improvement or unsatisfactory (Florida Statute §1012.33 (2)(c), 2015 & Florida Statute §1012.33 (3)(a)(b), 2015). For the purpose of this study, the term “professional service contract” may be used interchangeably with “tenured.”

School Administrators: School principals or school directors who are staff members performing the assigned activities as the administrative head of the school and to whom have been delegated responsibility for the coordination and administrative

direction of the instructional and non-instructional activities of the school (Florida Statute §1012.01 (3)(c)(1), 2015). Assistant principals who are staff members assisting the administrative head of the school are also included in this definition. This classification also includes career center directors (Florida Statute §1012.01(3)(c)(2), 2015). For the purpose of this study, school administrators are those who perform the managerial and instructional leadership responsibility of directing employees' work, planning the work schedule, controlling the flow of work or materials, train employees, handle complaints, authorize payments, and appraise productivity and efficiency of employees (Florida Statute §1012.01(7), 2015).

Teacher Effectiveness: Quantitatively measured in two ways: using value-added measures of student achievement and growth via standardized test scores, and using evaluation ratings rendered by administrative personnel who supervise schools (Torff & Sessions, 2009, p. 127). Effective teachers are those whose students experience high academic growth, while the students of less effective teachers experience less academic growth (Stronge, Ward, & Grant, 2011, p. 339). In Florida, teacher effectiveness is currently measured by using at least one-third of the total evaluation according to the teacher's value-added model score, at least one-third of the evaluation comprising an instructional practice score deemed by administrative observations, with the remaining one-third to be determined by other performance indicators (Florida Statute §1012.34, 2015).

Value-Added Model: "... a class of statistical procedures that use longitudinal test score data, i.e., data collected over a period of time, to measure the change in a student's

performance during a specific period of time” (Doran & Izumi, 2004, p. 3). According to the Florida Department of Education, the value-added score represents a positive or a negative percentage figure that compares the difference between predicted student performance and actual student performance as a result of an individual teacher (Florida Department of Education, 2015). For example, a value-added model score of zero means that a student’s performance was exactly as predicted. A negative value-added model score means that the student’s performance fell short of the prediction, and a positive value-added model score means that the student’s performance exceeded the prediction.

Theoretical Framework

The Social Systems Theory of Administrative Behavior

To understand the relationship between low student achievement measures and school leadership response to ineffectiveness, the researcher reviewed numerous writings, theories, and concepts that set the foundation for this study. From this research, Getzels and Guba’s (1957) social systems theory of administrative behavior provided the best blueprint in recognizing and defining the social behaviors in the administrative process grounded within this study. Getzels and Guba’s research rested on the premise that there are multiple components that contribute to a behavior exhibited by an administrator, and balancing these components is necessary in order to achieve institutional goals (Getzels & Guba, 1957). The groundwork for social systems theory is the innate conflict and need for balance between the nomothetic and idiographic dimensions of a social system (Getzels & Guba, 1957).

Nomothetic and Idiographic: Competing Dimensions

Getzels and Guba (1957) asserted that every institution is a social hierarchy encompassing both nomothetic and idiographic dimensions. Nomothetic dimensions are the dimensions in which the observed behavior and the goal are completely role-referenced. That is, the most direct route to achieving a goal is by having an established hierarchy and clear-cut expectations of behavior in order to reach a desired state. In the case of a school, the simplified hierarchy is the role of the administrator as the superordinate and the role of the teacher as the subordinate. Teachers are held accountable for goal achievement and direction toward purpose by their supervisors who are normally the school principals or assistant principals.

Furthermore, every institution has a purpose that resonates from the bottom to the top of the structural hierarchy (Getzels & Guba, 1957). In the case of a school system the institutional purpose is to educate students. Further, in the current age of accountability with regard to public education systems, the purpose is stretched from educating students to increasing student achievement. In Florida, student achievement accountability measures include individual and quantitative teacher effectiveness ratings known as VAM scores. Therefore, the goal of administrators and teachers alike is to increase student achievement as observed in VAM scores. There is, however, another component to teacher effectiveness: instructional practice scores deemed by administrators (often principals and assistant principals). If the function, purpose, and goal of any educational institution is to educate students and increase student achievement, behaviors within each role must be observed that directly lead to realization of the goal on behalf of both

teachers and the administrators. In the case of the school, the administrator holds the role of supervisor and evaluator, and the teacher holds the role of the subordinate and evaluatee. In terms of leadership and followership, the administrator's role is to "write the book," (Getzels & Guba, 1957, p. 436) and the teacher's role is to do things "by the book" (Getzels & Guba, 1957, p. 436). The nomothetic dimension is very black and white--there are behaviors that are required, and conversely, there are behaviors that are prohibited within roles of the social system (Getzels & Guba, 1957). With reference to a school institution, the behaviors that are required are ones which meet the goal of increasing student achievement, and the prohibited behaviors are any behavior that does not increase student achievement. The administrator must supervise, evaluate, and give performance feedback. The teacher, on the other hand, must achieve results.

Alternatively, there is the idiographic dimension associated with social systems theory. Whereas the nomothetic dimension places importance on the purpose of the institution as a machine-like system that achieves a goal, the idiographic dimension places importance on the understanding that all institutions are peopled, and the people are the key in realization of a goal (Getzels & Guba, 1957). The idiographic dimension deals exclusively with the individual needs and personalities of the people within the institution (Getzels & Guba, 1957). This dimension recognizes that the people within an organization are the "flesh-and-blood" that realize goals, and that no two people are exactly alike in terms of personality and needs. The needs of administrators are not only to achieve the goal of increasing student achievement, but also to create a sense of belonging and high morale between all stakeholders within the institution to develop

human capital. One might refer to the need of administrators as creating an institution that keeps teachers happy and feels natural and harmonious, thus making their own roles enjoyable without exerting a significant amount of strain and psychic energy (Getzels & Guba, 1957).

Here lies the dilemma for administrators: the conflict between balancing the nomothetic and the idiographic dimensions of a social system. Getzels and Guba (1957) wrote a general equation for observed administrative behavior: $B = f(R \times P)$, where B represents behavior observed, R represents the given institutional role defined by expectations (or nomothetic dimension), and P represents the personality of the particular role incumbent (or idiographic dimension). According to this formula, the school administrator is viewed to be caught between an individual and institutional conflict of roles and expectations (supervision, evaluation, and feedback toward the goal) versus his or her own personality and needs (harmony and belongingness), and must achieve a balance of both role and personality in order to exhibit behavior conducive to achievement of the institutional goal (Getzels & Guba, 1957). This notion was illustrated by Getzels and Guba (1957):

The unique task of administration, at least with respect to staff relations, is just this: to integrate the demands of the institution and the demands of the staff members in a way that is at once organizationally productive and individually fulfilling. (p. 430)

Role-Personality, Role, and Personality Conflicts

Getzels and Guba (1957) explained social systems theory by elaborating on three different but interconnected sources of conflict that an administrator may encounter while

trying to merge and balance the nomothetic institutional and idiographic individual dimensions: role-personality conflicts, role conflicts, and personality conflicts.

Role-personality conflict occurs when administrators feel they must choose between the nomothetic expectations and fulfillment of individual needs, creating a lopsided equation $B = f(r \times P)$ or $B = f(R \times p)$. If administrators favor the nomothetic, $B = f(R \times p)$, and become entirely role-oriented in pursuit of the institutional goal, they are said to be “autistic” in the sense that they completely ignore the personal integration aspect of administration and do not communicate with one another (Getzels & Guba, 1957). However, if administrators choose to maximize the personality aspect of the equation, $B = f(r \times P)$, they become “...liable for unsatisfactory role adjustment” (Getzels & Guba, 1957, p. 431) and considered unable to perform the role expectations. For the present study, role-personality conflicts has been referred to as “individual conflict avoidance.”

The second source of conflict is role conflict which manifests in three distinctive ways. First, role conflict can appear when there is disagreement within groups in defining the administrator’s role, as best exemplified by Getzels and Guba (1957):

For example, the principal of the school may be expected by some teachers to visit them regularly for constructive help and by others to trust them as professional personnel not in need of such supervision. (p. 432)

Additionally, conflict becomes apparent when there is disagreement among several power-rendering groups regarding the right each has to define the expectations for the same role, thereby representing contradictions in the expectations of two or more roles held by the same administrator (Getzels & Guba, 1957). For a school-based

administrator, examples of the groups that have the right to define that administrative role are the school board, the superintendent, and the teachers' unions, to name a few. Each of these groups not only has the right to define the role of the administrator, but each group also holds enormously different expectations for the same role.

Furthermore, another type of input resulting in conflict is contradiction within administrators when they are required to fulfill more than one role (Getzels & Guba, 1957, p. 432). For principals or assistant principals, being both mentors and evaluators to teachers create role conflict because administrators assume the role of the caring and helpful advisors to teachers who are struggling.. Also, however, they understands that they may be faced with the unsavory task of conducting a summative evaluation for a struggling teacher with the end result being possible dismissal. This type of conflict is referred to in this study as "bureaucratic procedural interferences."

The final sources of conflict are personality conflicts. Personality conflicts occur due to unique personality characteristics that detach the individual from the institutional purpose and goal, leaving the person "... to work out personal and private needs and dispositions, however inappropriate these may be to the goals of the social system as a whole" (Getzels & Guba, 1957, p. 432). This creates incongruence between the nomothetic and idiographic dimensions; the two ends simply do not meet, leading to administrative failure and a loss in institutional productivity (Getzels & Guba, 1957). When school administrators exhibit personality conflicts, they cannot reconcile their own needs as a person with the roles they are expected to perform. They may see no reason for the expectation placed upon them to supervise and evaluate staff because, due to their

own personalities, they find it difficult to call attention to the shortcomings of their subordinate teachers. For the purpose of this study, personality conflicts are henceforth referred to as “administrative procedural deficiencies/faults.”

Effectiveness, Efficiency, and Satisfaction

Getzels and Guba (1957) further elaborated on the nomothetic and idiographic dimensions in terms of effectiveness, efficiency, and satisfaction. Effectiveness is directly linked to the nomothetic role-centered dimension, whereas efficiency is linked to the idiographic side of the equation, and satisfaction transpires when the role expectations and the needs of the individual meet (Getzels & Guba, 1957). That is, one can be effective without being efficient, and vice versa, and one can also be satisfied within both the nomothetic and idiographic dimensions without ever being effective or efficient (Getzels & Guba, 1957).

Effectiveness is defined as reaching the goal or institutional purpose by behaving in such a way that is commensurate with role expectations (Getzels & Guba, 1957). Conversely, ineffectiveness is the opposite--the behaviors exhibited by the person to not meet the role expectations as defined by the institutional goal (Getzels & Guba, 1957). For the purpose of this study, school administrators can be said to be ineffective when they do not exhibit evaluative behaviors that move instructional staff closer to reaching the goal of increasing student achievement.

Efficiency is directly linked to the needs and behavior or the idiographic dimension (Getzels & Guba, 1957). In terms of an administrator, the needs dimension

must match the behavior in order for the administrator to maintain the efficiency needed to supervise a school and the “flesh and blood” teachers that reside within it:

When behavior conforms to the needs dimension, it appears “natural,” even pleasurable, and is forthcoming with a minimum of strain or expenditure of psychic energy. In this sense, the behavior is efficient. When the behavior conforms to the expectations dimension and there is a gap between expectation and needs, behavior is “unnatural,” even painful, is if forthcoming with a maximum of strain and expenditure of psychic energy. In this sense, the behavior is inefficient. (Getzels & Guba, 1957, p. 434)

When an administrator must evaluate a subordinate (i.e., teacher) and bring forth shortcomings via critical performance feedback directed to the teacher, it is a very unnatural behavior. It is one that is inherently difficult for the administrator, causing stress and inefficiency even though the feedback might be quite effective. In contrast, giving positive performance feedback is very natural and pleasurable for the administrator, thus rendering it efficient. It may, however, be ineffective in terms of reaching the institutional goal of increased student achievement.

Finally, the balance between effectiveness and efficiency results in overall satisfaction of the administrator must be considered (Getzels & Guba, 1957, p. 435). One is said to reach maximum satisfaction when the nomothetic dimension of the role and its expectations complements the idiographic individual needs dimension perfectly (Getzels & Guba, 1957). When this occurs, administrators are able to fulfill their roles in terms of their personalities without any conflict.. They are able to meet their personal needs and the expectations of the institution with great satisfaction, rendering the administrator both effective and efficient. This further results in behaviors that will reach the institutional goal (Getzels & Guba, 1957).

Leadership and Followership Styles

The next dimension of social systems theory elaborates on all the aforementioned components as related to leadership-followership styles (Getzels & Guba, 1957).

According to Getzels and Guba (1957), there are three types of leadership-followership styles--the nomothetic, the idiographic, and the transactional. The nomothetic style is based on role expectations without regard for need-dispositions by means of clear “black and white” expectations as the direct route to social behavior. The idiographic style is based on the need-dispositions without regard for the role expectations via deliberately vague and informal expectations (Getzels & Guba, 1957). The transactional style emerges as a balanced approach that satisfies both the role and personality factors in the behavior equation (Getzels & Guba, 1957).

In the transactional style of leadership-followership, there are clear-cut expectations set forth by leaders for their followers, but there is also high regard for the “flesh and blood” aspect--that the institution’s ability to reach its goal is incumbent upon those who comprise the institution (Getzels & Guba, 1957). The transactional style is the middle ground of leadership style, satisfying both the nomothetic and the idiographic dimensions, but it is also the most difficult for an administrator to navigate:

“Expectations are defined as sharply as they can be but not so sharply as to prohibit appropriate behavior in terms of need-dispositions” (Getzels & Guba, 1957, p. 438). It is important to note that school administrators must be transactional in their leadership craft. All teachers are different; no two are alike. Therefore, expectations must be within

the Goldilocks zone: sharp without stifling the individual art of teaching, but not too dull as to release the teacher from accountability for low student achievement measures.

Morale

It is natural for school leaders to wish to maintain high levels of morale among staff and within themselves. Morale is an important component to the social systems theory as it evokes feelings of identification and belongingness between people, institutions, and goals (Getzels & Guba, 1957). However, morale can only be achieved when the role expectations for the people within the institution are logically appropriate and the people within the institution experience a sense of identification of their personal needs within the goal (Getzels & Guba, 1957). When these two ends meet to a great degree, the morale within the institution is high. However, when one of these factors is zero, according to Getzels and Guba, morale can never be high, and motivation to reach said goal will be nonexistent.

In terms of school administrators and teacher evaluations, administrators need to realize the rationality of their role expectations within the evaluation system as a relevant component toward reaching the goal. If they do not, their morale will be low and evaluations will become more of a process to be completed than a process critical to the goal of student achievement. Similarly, if administrators do not identify the goal of student achievement within their own needs and values, motivation will be low and morale will also suffer.

In summary, the framework for this study was built using social systems theory as set forth by Getzels and Guba (1957). Administrators must reconcile between the

nomothetic and the idiographic dimensions of a social system so that their roles and personalities are balanced with those of their subordinates. Furthermore, administrators must recognize and appropriately react to various forms of conflict, including role-personality conflicts, role conflicts, and personality conflicts, (i.e., individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural deficiencies/faults), to maintain an effective, efficient, and satisfying work environment for the institution and the people held within it. Finally, leaders must develop a transactional leadership style that extends balance of their role expectations and the need-dispositions against those of subordinates in order to create a sense of belongingness for all constituents, hence maximizing morale. Administrators walk a tightrope, moving toward the goal of increased student achievement while also candidly evaluating subordinates in ways that do not diminish the individuality of teaching and learning but hold subordinates accountable for their behaviors that do not coincide with reaching the overarching institutional goal. Within this theoretical framework, the problem of balance between the nomothetic and the idiographic inherently exists as a contributor to the observed disparity among value-added measures of teacher performance and instructional practice scores specified by administrators.

Research Questions and Hypotheses

Following are the research questions that were used to guide this study:

1. What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?

H₀₁: There is no observed relationship between the lowest 10% of teachers' VAM scores as correlated to local instructional practice scores within Central Florida school districts.

2. What difference, if any, exists between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts?

H₀₂: There is no observed difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts.

3. What percentage of professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers identified with the 10% of lowest VAM scores in Central Florida school districts have been placed on improvement plans?
4. Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida school districts?

Limitations

The researcher identified the following limitations for this study:

1. The Florida Department of Education (FDOE) VAM score data for 2013-2014 excluded some teacher names. According to an information file held within

the documents delivered to the researcher, some individual teacher names were masked for any teacher who is exempt from public record disclosure per Florida law. The FDOE cited the following Florida Statutes for these exemptions: §119.071 (2)(j), (4)(d), and (5)(i); §265.605; and §267.17.

2. The researcher relied on local school districts for accurate and complete school district evaluation and improvement plan data. Some quantitative and qualitative data may not be present in the information provided by the local school districts if this information was not held within the personnel files of individual teachers at the district-level offices.
3. Research question two was asked to determine if there was an observed significant difference between tenured and nontenured teachers' instructional practice scores. The only two variables that were analyzed and investigated to respond to this question were contract status, and the instructional practice score of each teacher in the population. Extraneous variables, such as years of teaching experience and corresponding VAM scores for each teacher were not included in the data findings.
4. Charter school employees were removed as selected subjects due to charter school employee personnel records not being housed within local school district offices. Thus, this study was limited strictly to traditional public school teachers and did not include any subjects who worked in a charter school for the 2013-14 school year.

5. Of the 758 total teachers identified in the population, 230 teachers were removed from this study due to a variety of limiting factors, such as: (a) incomplete personnel files held within school district central offices, yielding missing information for subjects; (b) termination of employment prior to completion of the 2013-14 school year summative instructional practice evaluations; (c) retirement of employees prior to the completion of the 2013-14 school year summative instructional practice evaluations; and (d) duplicates of teacher names within the original VAM score data from the FDOE due to part-time employment at multiple schools.
6. Marion County School District and Lake County School District teachers were eliminated from the population due to lack of contract statuses being produced by these school district central offices via public record request.

Delimitations

This study was delimited by the following factors:

1. This study was restricted to the 13 counties represented by the Central Florida Public School Board Coalition (11 counties represented once Marion and Lake County school districts were eliminated from the total population).
2. The personnel records examined for the purpose of this study were those of public school teachers with the lowest 10% of VAM scores for each of the Central Florida school districts described according to FDOE data sources for the 2013-14 school year.

3. The VAM data, local instructional practice evaluations, and improvement plan data from the 2013-2014 school year were accessed under Florida Statute §119.07 via public record requests made to the FDOE Division of Accountability Research and Measurement and each individual school district. Any data that was exempt from this statute could not be obtained for this study.

Organization of the Study

This research study is presented in five chapters. Chapter 1 includes the background of the study, statement of the problem, purpose of the study, significance of the study, definition of terms, theoretical framework, research questions, limitations, and delimitations. Chapter 2 contains a review of the literature which includes individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults. Chapter 3 describes the methodology used in this research study. Chapter 4 presents the findings and data analyses from the five research questions. Finally, Chapter 5 provides a summary of the entire study, discussion of the findings, implications of the findings for theory and practice in school administration, and recommendations for further research.

CHAPTER 2 REVIEW OF LITERATURE

Introduction

At the time of the present study, the use of Valued-added Model (VAM) scores to determine teacher effectiveness was a fairly new idea grounded in the belief and recent legislation that student achievement was directly and significantly correlated to teacher effectiveness. *A Nation at Risk*, published in 1983, drew interest in reforming teacher evaluation in connection with merit pay to address “the rising tide of mediocrity” (Donaldson, 2009, p. 4). More than 20 years later, United States Secretary of Education Arne Duncan proposed that districts begin reporting the percentage of teachers rated in each evaluation performance category in response to the growing concern that students, and teachers by extension, are underperforming despite years of progress monitoring (Donaldson, 2009).

Although instructional practice evaluations have been in place for decades, there has been little evidence to support that evaluations conducted by principals reliably identify ineffective teachers with respect to student achievement outcomes (Weisberg et al., 2009). Moreover, there has been even less evidence to suggest that the instructional practice framework evaluation models used by school districts was being used with fidelity to make human resource personnel decisions and provide responsive and candid feedback to improve teaching and increase capacity for building human capital. Though there has been little public disagreement that ineffective teachers should not remain in the classroom, few have agreed as to how administrators should identify ineffective teachers for remediation and dismissal (Winters & Cowen, 2013).

This chapter presents the rationale for conducting further research on the use of Florida VAM scores for personnel decisions as correlated to local instructional practice scores. The review of literature has been organized around (a) understanding the context of teacher evaluation in the current era of reform; (b) an explanation of the observed Lake Wobegon Effect in prior studies conducted throughout the United States; (c) the barriers of candid teacher evaluation, including three key dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults; and (d) a summary of prior research related to the present study.

Multiple sources were used to search the literature. These included ERIC, ProQuest, prior dissertations and theses, Internet sources, and references including journals, periodicals, books, published reports, and the Florida Statutes.

Understanding the Context of Teacher Evaluation in the Current Era of Reform

According to the National Council on Teacher Quality, only four states required teacher evaluations in 2009, and no states used evaluations to make tenure or dismissal determinations (DeNisco, 2014). Traditionally, teacher evaluations systems relied solely on principal observations of instructional practice and did a poor job at defining the most effective, least effective, and middle-range teachers (Weisberg et al., 2009). There has been little evidence supporting historical evaluation systems as a solid link between teacher evaluation and student achievement (Steinberg & Donaldson, 2014). After the passage of No Child Left Behind (NCLB) and Race to the Top (RTTT), all 50 states updated evaluation policies and statutes so as to increase the use of student achievement outcomes in teacher evaluation practices such as VAM scores (Hazi & Arredono

Rucinski, 2009; Steinberg & Donaldson, 2014). According to Steinberg and Donaldson (2014), states with new evaluation systems have required, on average, approximately two formal observations and two informal observations for both beginning and veteran teachers. In most cases, a formal observation was about 30 minutes in length, and observers gathered and recorded evidence of a teacher's instructional practice guided by an observation rubric (Steinberg & Donaldson, 2014). A formative evaluation typically ranged from 15 minutes to a short walkthrough where data collection and associated feedback were not always required (Steinberg & Donaldson, 2014). However, the frequency and intensity of an evaluation has often varied according to a teacher's contract status. A veteran teacher with a professional service contract or continuing contract (often known and referred to as tenured) typically has been evaluated less frequently and intensely, if at all (Donaldson, 2009).

The reasoning for major teacher evaluation reform is twofold: first, to use solid methods of teacher assessment to increase teacher instructional skills via professional development, coaching, and training, and, second, to aid in high-stakes personnel decisions such as non-reappointment, termination, decision to award merit pay, and removal from tenure status (Steinberg & Donaldson, 2014). In the state of Florida, at the time of the present study, teacher evaluation in accordance with Florida Statute §1012.34(3)(4) was comprised of at least one-third student achievement outcomes as measured by a VAM, at least one-third local instructional practice evaluation scores, and the remaining one-third of other indicators of performance. These indicators were outlined by professional and job responsibilities as recommended by the State Board of

Education or identified by the district school board, peer reviews, objectively reliable survey information from students and parents based on teaching practices that are consistently associated with higher student achievement, and other valid and reliable measures of instructional practice.

These statutory changes occurred in the wakes of NCLB and RTTT, in addition to a call by the National Governors Association (NGA) to target teacher evaluation as a means to increase student achievement via a highly qualified teacher in every classroom (Hazi & Arredondo Rucinski, 2009). Among the policy goals proposed by the NGA for improving student learning, Florida specifically implemented required annual teacher evaluations and a focused approach on improving teaching practice through peer review and portfolios. Broadening evaluation participation to include teachers, administrators, and parents, using a state-approved teacher evaluation system with an increased focus on student progress (learning gains), and the use of teacher peers when conducting evaluations were also included in proposed system revisions (Hazi & Arredondo Rucinski, 2009).

The Lake Wobegon Effect

The Lake Wobegon Effect is a term characterized by “a phenomenon in which most individuals or groups perform above average” (Wheeler & Haertel, 1993, p. 82). Within the educational realm, the Lake Wobegon Effect was first used in a survey report proposed by Cannell in 1987 where in all 50 states, 90% of school districts reported that students tested above average despite evidence of poor literacy and graduation rates (Wheeler & Haertel, 1993). Wheeler and Haertel cited that the reasoning for the

manifestation of the Lake Wobegon Effect in Cannell's study was due to "...dated norms, use of nonsecure tests, selection of curricula and programs that are closely aligned to the tests, teaching to the tests, inappropriate test administration and scoring procedures, and selection of those individuals to be tested" (p. 82). In the case of local teacher instructional practice scores, the Lake Wobegon Effect is realized when most teachers are deemed effective or highly effective despite contradictory evidence displayed by low student achievement measures (Tucker, 1997). Three key dimensions have surfaced from the literature as contributors to the Lake Wobegon Effect on inflated performance evaluations: individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults. These three dimensions encompass barriers to meaningful evaluative practice such as... "poor evaluation instruments, limited district guidance, lack of evaluator time, lack of evaluator skill, lack of evaluator will, absence of high-quality feedback for teachers, and few consequences attached to evaluation" (Donaldson, 2010, p. 55).

An illustration of the Lake Wobegon Effect in teacher evaluation practice was documented in 2001 when Patricia Hopkins became the superintendent of two school districts in Maine. According to Hopkins, when conducting a review of the summative evaluations of all teachers in the two districts, the performance reviews were more akin to valentines than evaluations; the summative ratings were full of vague, meaningless praise, and overall, were absent of constructive criticism and actionable feedback (Donaldson, 2010). In a study of 15,176 teachers across 12 districts, 74% of teachers surveyed reported that they had not received specific feedback in their evaluations on

how to improve instructional practice (Weisberg et al., 2009). Moreover, in the same study, half the districts surveyed did not dismiss a single non-probationary teacher within a five-year span; more than 99% of teachers receiving a satisfactory rating on a binary scale; and 94% of teachers received one of the top two ratings on a multi-tiered scale (Weisberg et al., 2009). In the same year, Donaldson (2009) cited a Chicago study in which 56% of veteran principals conceded that they assigned teachers a higher evaluation rating than their evaluation warranted. She also noted that although it is possible that all teachers are effective in some schools, it is more likely that variation in teacher effectiveness exists within schools than between them.

According to Weisberg et al. (2009), not only do school districts fail to acknowledge differences between teachers in terms of teacher performance and student outcomes, but most teachers' evaluations appear to be highly inflated and skewed at the top of the rating scale: over 60% of teachers received the highest rating, 10% received a middle rating, and only 2% received the lowest rating possible. Even with a state-prescribed instrument for principal evaluation of teacher performance, there is no guarantee of a more stratified distribution of performance scores; Langlois & Colarusso (1988) found that 98.2% of teachers received a perfect evaluation score when districts used the Pennsylvania State Education Department's standard rating form to assess teacher performance. Mela (2013) found within a population of 1,138 teachers in Brevard County, Florida, 99% were rated effective or highly effective (p. 138). More recently, Pace (2015), in a similar Brevard County study, found that administrators rated 92.5% of all teachers as highly effective or effective on the professional practices portion

of the state-approved teacher evaluation instrument. These instances displayed a continued failure of school administrators to use evaluation ratings and observation results to accurately distinguish between levels of teacher performance despite increased evaluator training initiatives. Furthermore, these findings have not been confined to individual states: Tucker (2001) estimated the average number of incompetent teachers to be between 5% and 15% overall. In like manner, a random sample of principals nationwide determined that only 46% of all principals gave their school an excellent rating:

Thus, *any* school--low- performing or high-performing, wealthy suburban or under-resourced urban--is more likely to employ more under-performing teachers than its evaluations ratings suggest. In fact, principals *and* teachers believe that teachers are less effective than evaluations ratings would indicate. (Donaldson, 2009, p. 2)

The Case for Using Value-Added Modeling to Rate Teacher Performance

It has been established that individual teachers are the most important school-related factor in student achievement and growth, but differences in teacher effectiveness are not well predicted by historical evaluation practices and measures (Daley & Kim, 2010, p. 1). Differences between teacher effectiveness provide the best available explanation for differences in achievement and growth between students once student background is controlled for using VAM analysis (Daley & Kim, 2010). Moreover, although it has been established that principals can accurately identify the best and worst teachers in a school, VAM estimates of teacher effectiveness have been found to generally better predict future student achievement than have principal ratings, particularly when conducting evaluations of the broad middle-range of teachers in terms

of effectiveness (Jacob & Lefgren, 2008). When considering local district evaluations as compared to VAM scores, a scan of prior research provided evidence of low to moderate correlations (Donaldson, 2009). In Cincinnati, Ohio, combined VAM estimates and teacher evaluation scores for a sample of teachers in Grades 3-8 correlated at .43 for mathematics, .32 for reading, and .27 for science, $p < .001$ (Milanowski, 2004). Additionally, in a study of a Los Angeles, California elementary charter school, Gallagher (2004) indicated a strong, positive, and statistically significant relationship between teacher evaluation scores and student achievement using VAM teacher effects correlated with teacher evaluation scores in reading (.50 correlation, $p = .01$). Papay (2011) found correlations within a large Northeastern United States school district that ranged from .15 to .58 when correlating VAM scores with various state standardized assessments for mathematics, reading, and English-language arts, as well as the Stanford Achievement Test (SAT) subtests for reading and mathematics. Specifically, in Florida, Mela (2013) correlated each of the eight components of the School Board of Brevard County Instructional Personnel Performance Appraisal System Instrument to individual teacher VAM scores, and found weak to moderate positive relationships spanning between .089 and .231, $p < .05$.

The Lake Wobegon Effect manifested through teacher instructional practice scores has serious and lasting ramifications on student progress and achievement over time. Although the instability of VAM scores of teacher effectiveness has been acknowledged in the literature, there are data that provide for validity in practice when multiple assessments and years of teacher performance have been considered (Papay,

2011; Winters & Cowen, 2013). For example, Winters and Cowen, in their 2013 study in Florida), found that students assigned to teachers who one or two years earlier would have been dismissed according to a VAM-based policy of ineffective teacher dismissal made considerably smaller academic improvements than did students assigned to effective teachers denoted by VAM measures. More specifically, these data provided evidence that students assigned to teachers at or below the fifth percentile with consecutive years of ineffective VAM score indicators yielded an average 0.188 standard deviation decrease in student achievement one year later compared to students with teachers who were above the fifth percentile according to VAM measures (Winters & Cowen, 2013).

Barriers to Candid Teacher Evaluation Practices

School administrators face many challenges that typically work against their decision to recommend contract non-renewal for teachers (Nixon, Packard, & Dam, 2011). Among these barriers are individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults, embedded within the competing nomothetic and idiographic dimensions of the social systems theory of administrative behavior (Getzels & Guba, 1957). Getzels & Guba posed this theory in terms of the following equation: $B = f(R \times P)$, where B represents behavior observed, R represents the given institutional role defined by expectations (the nomothetic dimension), and P represents the personality of the particular role incumbent (idiographic dimension). According to this formula, school administrators are viewed to be straddling individual and institutional conflict of roles and expectations (supervision, evaluation, and feedback

toward the goal) versus their own personality and needs (harmony and belongingness) (Getzels & Guba, 1957). School administrators must achieve a balance of both their roles and personality dimensions to exhibit behavior conducive to achievement of the institutional goal. It is important to note that individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults all contribute to the same equation, are interconnected in nature, and often overlap in professional practice.

Individual Conflict Avoidance

Individual conflict avoidance is defined as an administrative personality flaw or lack of will to achieve the educational goal that causes an aversion to conflict and discomfort when conducting face-to-face performance critiques. Getzels and Guba (1957) referred to this as role-personality conflict. Individual conflict avoidance occurs when administrators believe they must choose between nomothetic expectations and fulfillment of individual needs, creating a lopsided equation between the administrative role and personality domains $B = F(r \times P)$ or $B = f(R \times p)$ (Getzels & Guba, 1957). When administrators favor the nomothetic, $B = (R \times p)$, they behave in a way that is entirely role-oriented in the pursuit of the institution goal while simultaneously ignoring the personal integration aspect of administration and leadership. In contrast, if administrators choose to maximize the personality aspect of the equation, $B = (r \times P)$, they become unable to perform the role expectations due to an acute awareness of the “flesh-and-blood” needs of those within the organization (Getzels & Guba, 1957). Evaluation systems that reflect both the accountability and the personal growth

dimensions are critical to the balancing of individual and institutional goals within an educational leader (Stronge, 1995). Hain and Smith (1966) stated:

Every principal holds in his hands the career of a significant number of teachers. Hopefully, principals exercise wise judgments and view their role as one helping their teachers do a better job with students. (p. 1)

Although these words date back to the mid-1960s, the notion of principal as caregiver to his or her teachers still rings true today. Principals have a role to fulfill in increasing student achievement, but they also have a role in the human capital development of staff, including non-retention of staff members who are not fulfilling their responsibility toward reaching the institutional goal:

As educational leaders, principals are held responsible for not only the educational quality at the building level but the morale and general welfare of all those who work and study in their buildings. They must be able to verbalize their vision of ethically responsible conduct and demonstrate through their actions that they take seriously the moral obligations that come with the position no matter how onerous the task. (Pratt, 1996, p. 30)

When faced with difficult decisions such as non-retention of an underperforming teacher, the educational leader must come to a morally defensible decision that balances the demand of the institutional goal against the competing demands of protecting students, teachers, and the organization (Pratt, 1996). This can be an extremely lonely time for a principal, creating an uncomfortable divide between the leadership and the subordinates within the school (Randklev & Lemon, 1990). A major contributor in the dismissal of incompetent teachers is the personal stress a principal may endure during the process: “The psychological consequences of threatening another human being’s self-worth becomes a critical issue for principals when they are face-to-face with teachers

who are not performing satisfactorily” (Mock & Melnick, 1991, p. 6). Additionally, administrators may experience difficulty with confrontation and hostility, stemming from the personality-role conflict of first offering assistance, and rendering the necessary final judgment on performance (Tucker, 1997). In situations where principals must decide on whether to recommend or non-retain a teacher, they must deal with their own underlying personal issues such as disillusionment, distrust, stress, and fear of failure, combined with the multiple roles they hold within the organization while considering how to act and be perceived in their roles as educational leaders (Pratt, 1996; Stronge, 1995). Phillips and Young (1997) surmised that educational leaders take on a role in an organization as the keeper of justice and caring, simultaneously balancing an attitude of caring and justice for students with caring and justice for individual teachers whom they regard as incapable of meeting students’ educational needs. In an interview, a superintendent stated that the role of an educational leader must be to grow human capital at any cost:

Well, I guess the one thing that a good teacher really can’t afford to do is to give up on any students...I would take that analogy and apply it to what we are doing with teachers...that those who are on the team were good enough to make it. And if they’re not playing as well as you want them to be, or they’re affecting the performance of the team, then you’ve got to do all you can to help them. (Phillips & Young, 1997, p. 112)

The superintendent continued:

Well, how many chances do you give a person? That’s a difficult one... it may mean that we are going to have to be in that pattern of working with that person on an on-going basis, and we may have to accept that. (Phillips & Young, 1997, p. 112)

However, the means of action for an educational leader in the current era of reform simply cannot be to accept that a teacher is underperforming and enter into a holding

pattern of assistance. President Obama recently spoke about improving teacher evaluation in instances of teacher performance issues stating:

If a teacher is given a chance or two chances or three chances but still does not improve, there is no excuse for that person to continue teaching. I reject a system that rewards failure and protects a person from its consequences. The stakes are too high. We can afford nothing but the best when it comes to our children's teachers and the schools where they teach. (Weisberg et al., 2009, p. 2)

However, dismissal of incompetent teachers can be surmised as a problem inherent within the culture of school organizations, where teachers expect to be rated effective or highly effective (Weisberg et al., 2009). The discomfort associated with confronting ineffective teachers, particularly when there has been a long history of evaluative predecessors who had avoided confronting poor performance in the past, can be daunting for an educational leader (Phillips & Young, 1997). If educational leaders tend to imbalance the role and personality dimensions in favor of individual conflict avoidance, the ramifications can have lasting negative effects when nearly all teachers are rated good or great (Weisberg et al., 2009). In a 2009 New Teacher Project report, a Chicago public school teacher revealed, "Many teachers are accustomed to receiving a 'superior' rating and simply do not accept anything lower. It also seems to be an easier way out for administrators, rather than have a confrontation with the teacher" (Weisberg et al., 2009). In another study, an educational supervisor was interviewed, revealing how some predecessors had responded to a nonperforming teacher in favor of individual conflict avoidance:

They had, in fact protected this teacher over the years--the school and the school district- in not giving this particular teacher [courses for which there were province-wide examination] so that there wasn't a check on achievement and that kind of thing.... Why wasn't this addressed in the first 20 years of this man's

career? It's just sad that he was allowed to carry on.... Here's a man who taught in this district for 20 years. And if he wasn't provided with assistance to the job that we wanted to be done- there's some responsibility on our part- over 20 years. There is something wrong with what we have been doing as a district.... Kids were dropping out of his classes, no one ever told him why...the parents...would phone the principal. Things were changed quietly. I think no one was every up-front with the guy. It's not fair. (Phillips & Young, 1997, p. 113)

Other means of individual conflict avoidance on behalf of educational administrators is transferring an underperforming teacher to another school to avoid facing a difficult decision where there are no completely satisfactory solutions (Bridges, 1985). Some administrators may move underperforming teachers to another school in the hope that a change in environment and a fresh start may spark improvements in the teacher's performance, though little evidence exists to suggest this is a solution to the overarching problem (Fuhr, 1993). This practice is referred to within the education community as "the dance of the lemons...pass the turkey" (Bridges, 1985, p. 21). Fuhr (1993) criticized this practice, stating that it was "...simply the coward's way out. Remember that whatever we do must be based on what is best for our students. Therefore, marginal performance in School A usually will mean marginal performance in School B" (p. 28).

Working with an underperforming teacher requires discipline and commitment on the part of an educational leader. This includes explicitly articulating to a teacher the area that needs improvement, interpreting concerns and expectations, gathering data on how to assist deficiencies, helping the teacher make the needed improvement outlined within a strategic plan, and setting a timeline for expected improvement within the confines of the plan (Randklev & Lemon, 1990). "You must hope for improvement,

expect improvement, and work for improvement” (Randklev & Lemon, 1990, p. 44). Good teachers respect administrators who are not afraid to confront and correct poor performance (Fuhr, 1993). To the contrary, if management fails to make the required decisions to eliminate poor performance and reward truly exceptional teachers, overall teacher performance and morale will decline (Fuhr, 1993).

Bureaucratic Procedural Interferences

Bureaucratic procedural interferences occur when state and local district policies and personnel do not fully support school-based administrative actions and decisions intended to tackle ineffective performance. Contributing factors to bureaucratic procedural interferences are: poor evaluation instruments, limited district guidance, lack of evaluator time, and few consequences attached to evaluations (Donaldson, 2010). Getzels and Guba (1957) referred to this as role conflict, because it exclusively impacts the role factor (R) in the equation $B = f(R \times P)$. Role conflict becomes apparent when there is disagreement among several power-rendering groups that each has the right to define the expectations for the same role though they represent contradictions in the expectations of two or more roles held by the same administrator (Getzels & Guba, 1957). For a school-based administrator, these groups may be the superintendent, the school board, or the teachers’ union. Each of these groups not only has the right to define the role of the administrator, but each group typically holds enormously different expectations for the same role (Getzels & Guba, 1957, 432).

Langlois and Colarusso (1988) stated that these competing demands for defining the role of the principal can dramatically alter how evaluation practice is carried out.

Furthermore, drastic change is necessary on behalf of education organizations to focus more time on prioritizing, developing, and recognizing human capital through solid administrative supervisory practices:

This kind of change requires courage--any big change does--and it requires rethinking the way school executives spend their time. But how are we to pursue this ideal amid the realities of the frantic busyness of the central office, the demands of superiors, the ire of parents, the bravado of students, the complaints (or worse yet, the complacency) of teachers, and the grumbling of unions? The loudest cry is, of course, about the principal's lack of time. (Langlois & Colarusso, 1988, pp. 13-14)

In addition to time constraints within the role, principals are acutely aware of the political context within educational organizations and feel pressure from multiple power-rendering parties that might affect decisions to dismiss ineffective teachers. Teacher supervision, development, and evaluation have reinforced the top-down nature of school governance (Cooper, Ehrensall, & Bromme, 2005). For example, in New York City, regulations of the Department of Education combined with contract obligations from the United Federation of Teachers have limited the Department's power to review and remove teachers (Cooper et al., 2005). The bureaucratic policies and procedures have translated to limited actions to be taken by school principals due to elaborate processes and third-party interventions. Furthermore, union representatives are specially trained in the grievance process. As a result, principals who believe they have a strong case against an incompetent teacher become disillusioned when the arbitrator cites technical errors and the case is dismissed:

For example, if the principal wishes to give a teacher a negative performance evaluation, he or she must write a letter that conforms in format to past arbitration decisions. An incorrectly written letter will be removed from the teacher's

personnel file and will not be used as future evidence of poor teacher performance. (Cooper et al., 2005, p.118)

When making the decision to dismiss an incompetent teacher, Brieschke (1986) found that most principals cited that they tried to avoid involving themselves in the lengthy, time-consuming, complicated, tension-producing procedures because the process often did not end in the desired result. One principal stated:

The procedure is ridiculous. You just can't keep up with all the fine points. If a teacher is unsatisfactory on the 50th day she's going to be unsatisfactory on the 51st day, or the 55th or 60th. It's a tremendously involved situation and many times you're not successful. I just don't bother with it anymore. (Brieschke, 1986, p. 244)

However, because local school boards and districts, not individual schools, enter into legal contracts with teachers, the district controls the hiring, transferring, and dismissal processes (National Council on Teacher Quality, 2010). Since the implementation of RTTT, districts have largely made a concerted effort to improve evaluation processes. Nevertheless, even in districts and states that require annual evaluations, districts lack systems capable of recording evaluations electronically, rendering the district unable to monitor teacher progress as noted by evaluators (National Council on Teacher Quality, 2010). It has been reported in a study spanning school districts in the states of Arkansas, Colorado, Illinois, and Ohio that only one in 12 districts studied centrally tracked or recorded any evaluation data (Weisberg et al., 2009). Other times, district processes allow for incompetent teachers to circumvent the dismissal process by taking a leave of absence or some other measure (Brieschke, 1986). One principal noted:

I had one teacher who did everything to try and wiggle out of being E-2'd. She just dropped out--didn't show up, totally tried to confuse me with the days. Of course, she was nowhere to be found on the 50th or 51st days (when the E-2 letter had to be delivered). But I just kept counting and watching the days, and I finally got her. (Brieschke, 1986)

Moreover, teacher seniority, rather than teacher performance, has been a long-accepted means of deciding which teachers to dismiss when student enrollment declines in school districts (National Council on Teacher Quality, 2010). In a study conducted by the National Council on Teacher Quality (2010), only six of the 76 districts surveyed explicitly allowed for performance to be a factor in excessing decisions. Among the districts in Florida within the study, Broward County allowed exceptions to this rule only for teachers with extracurricular positions, those who were union representatives, and those who were resource teachers, counselors, librarians, and reading teachers (National Council on Teacher Quality, 2010). Dade County made exceptions for union representatives, bilingual teachers, gifted and talented/international baccalaureate/advanced placement teachers, and teachers with special training or unique skills (National Council on Teacher Quality, 2010). Duval County made exceptions for union teachers with extracurricular positions, union representatives, bilingual teachers, and teachers with special training and or unique skills (National Council on Teacher Quality, 2010). Finally, Palm Beach County made exceptions only for teachers with extracurricular positions (National Council on Teacher Quality, 2010). This bureaucratic means of decision-making means that new teachers are always the first to be dismissed, no matter how effective they are (National Council on Teacher Quality, 2010). Therefore, this practice leaves principals with little discretion as to who teaches in their buildings. It

also inequitably distributes underperforming staff members to schools that primarily serve poor and minority student populations (National Council on Teacher Quality, 2010).

Though school districts may be to blame for producing a bureaucratic system that does not include student performance measures for personnel decisions, state laws also contribute to the problem. The landmark education reform case, *Vergara v. California* recently challenged and struck down both tenure laws and last-in-first-out (LIFO) policies within the state (Davis, 2015; Robertson, 2015). The outcome of the case lent momentum to efforts across the country in favor of using student performance to measure quality teaching and eliminating seniority as a primary criterion for making personnel decisions (Davis, 2015). The plaintiffs argued that existing teacher job-protection laws caused a disproportionately high number of ineffective teachers to be placed in poor and minority schools, essentially violating the equal protection clause of the California Constitution with regard to students' fundamental rights to education (Davis, 2015).

Furthermore, Futernick (2010) surmised that both educators and policy-makers must better understand the challenge to administrative personnel regarding the problem of unaddressed, underperforming teachers. According to both Futernick and Elmore (2002), accountability tends to move in one direction: people with less authority are accountable to those with more. Accountable communities are those in which members have moved beyond simply working well together; they take responsibility for monitoring the community's own actions and for calling others on behaviors and stances that are not helpful to the mission (Platt & Tripp, 2008). When those at the bottom of the

authority hierarchy are squeezed by competing role definitions and “noise” from multiple stakeholders, they are unable to adequately fill any role expectation (Futernick, 2010). To better improve the evaluation system, Elmore (2002) and Futernick (2010) suggested implementing policies of reciprocal accountability, where people with authority would not just monitor performance and impose sanctions, but be responsible for ensuring that those being monitored have the tools and backing they need to succeed.

Administrative Procedural Faults

Administrative procedural faults account for when administrators do not accurately document teacher instructional practice with fidelity. Although individual conflict avoidance is due to lack of administrator will, administrative procedural faults are due to lack of administrator skill in evaluation combined with the absence in delivery of high-quality feedback (Donaldson, 2010). Moreover, Weisberg et al. (2009) asserted that the problem with teacher evaluation is not only with regard to underperforming teacher feedback and dismissal, but also with administrators not formally identifying those teachers who are truly exceptional:

In a world where all teachers are rated good or great, the truly outstanding teachers--those who are realizing life-changing academic success for their students--cannot be formally identified. And if they are not formally identified, schools cannot prioritize their retention or leverage them to develop and improve their colleagues. (p. 13)

Brieschke (1986) referred to this lack of skill as “educational mistakes” (p. 238) which are hallmarked by errors in action, judgment, perception, or impression as to what is required in the school situation to achieve intended goals. Getzels and Guba (1957) referred to these faults as personality conflicts that detach the individual from the

institutional purpose and goal, leaving the person "... to work out personal and private needs and dispositions, however inappropriate these may be to the goals of the social system as a whole" (p. 432). Personality conflict deals exclusively with the personality factor (P) in the equation $B = f(R \times P)$ (Getzels & Guba, 1957). This leads to incongruence between the nomothetic and idiographic dimensions; the two ends simply do not meet, leading to administrative failure and a loss in institutional productivity (Getzels & Guba, 1957). When school administrators exhibit personality conflicts, they cannot reconcile their own needs as persons with the roles they are expected to perform. Administrators may see no reason for the expectation placed upon them to supervise, evaluate, and deliver high-quality feedback to teachers due to their lack of skill in the evaluative process.

A revelation of administrative procedural faults was discovered by Range et al. (2012) in a study conducted to explore differences between principals' and superintendents' perceptions about teacher incompetency, strategies most frequently implemented when working with incompetent teachers, and barriers to dismissing incompetent teachers. Most concerning were the differences observed within a Likert-scale survey (1 = no strength, 4 = high strength) that compared discrepancies between principals' and superintendents' views of barriers to dismissing incompetent teachers (Range et al., 2012). Principals rated required administrative time ($M = 2.86$), protection of employee by professional association ($M = 2.79$), and legal and other expenses ($M = 2.71$) as barriers with strength when attempting to dismiss incompetent teachers (Range et al., 2012). On the other hand, superintendents rated unclear definition of incompetence

($M = 2.60$), protection of employee by professional association ($M = 2.57$), and lack of strength of character by the principal ($M = 2.53$) as the strongest barriers to dismissing incompetent teachers (Range et al., 2012). The data provided a glimpse of the differing perceptions of principals as opposed to superintendents when viewing issues related to dismissing incompetent teachers, but moreover, provided evidence that superintendents agreed significantly more than principals ($t = 2.11$, $p < .04$) and that lack of strength of character of the principal was a major barrier to dismissing incompetent teachers (Range et al., 2012).

Although school administrators have most often cited tenure as a major barrier to incompetent teacher dismissal, opponents of this view cite weakness in skill and ineffective management on behalf of school-based administrators as the root cause of this misconception. Finberg (as cited in Davis, 2015), a representative of the California Teachers Association during the *Vergara v. California* case, surmised:

Statutes don't assign teachers to a classroom...the statutes don't say anything about race or poverty. Districts assign teachers to various schools with various populations...put stronger principals in those schools...put resources into those schools...you won't have that phenomenon. (p. 20)

Furthermore, Zirkel (2010) asserted that the use of tenure as an excuse to not terminate an ineffective teacher was a self-perpetuating scapegoating process. Zirkel (2010) conducted a comprehensive canvas of court decisions, finding that defendants prevailed over plaintiff teachers by a greater than three-to-one ratio, with no significant difference between teachers who were tenured and nontenured:

...contrary to the prevailing perception, teacher tenure doesn't guarantee lifetime employment. Legally, teacher tenure is no more than procedural due process, which means that notice and hearing are needed to ensure fundamental fairness

that a termination must be based on generally accepted reasons, such as incompetency, insubordination, and immorality. If this level of procedural and substantive protection for an individual teacher becomes top heavy, outweighing the interests of students and the rest of the institutional enterprise, the problem isn't tenure, but the lack of will among various players in the tenure process, including those who participated in making state law and collective bargaining. (p. 76)

Moreover, Nettles and Herrington (2007) identified significant relationships in prior literature between selected school leadership practices and student learning, indicating that evidence exists to support the notion that principal behaviors produce a direct relationship to student achievement (p. 724). Among the duties of a school principal as an instructional leader, comprehensively evaluating staff and rendering high-quality feedback are at the top of the list with other areas. These include: maintaining a safe and orderly environment, development of mission and vision, including stakeholders in school communities, monitoring school progress, initiating instructional focus, setting high expectations for student performance, and developing appropriate professional development (Nettles & Herrington, 2007). Ovando and Ramirez (2007) identified common behaviors that principals exhibited in a selected sample of school leaders at exemplary or recognized schools. Behaviors included were: setting clear expectations, monitoring instruction by conducting walkthrough observations, and connecting staff development to teacher performance evaluation. One principal stated:

Obviously, we do walkthroughs. And I'll be the first one to tell you that we don't do enough of them. And- because you're taping I'm going to tell you that it is my formal opinion that walkthroughs is the difference between being a good campus and a not very good campus. (Ovando & Ramirez, 2007, p. 98)

Principal behavior and skill are certainly a factor when tackling incompetence. When improvement efforts fail, principals are faced with the moral dilemma of facing a problem head-on, or ignoring the problem and turning the other way (Blacklock, 2002). Although performance management is seen as a vital component to managing an effective school, it is commonly perceived as much too difficult and tends to be ignored or sidelined by those who manage (Yariv & Coleman, 2005). However, if the school runs a strict procedure of periodical evaluations that includes formative and summative feedback, there is a better foundation to offer assistance or dismiss a teacher, if necessary (Yariv & Coleman, 2005). It can be argued that administrators prefer to give teachers critical feedback outside the formal evaluation process. However, the New Teacher Project found that 47% of teachers reported not having a single informal conversation with their administrators within the past year about improving their performance (Weisberg et al., 2009). Findings by Yariv and Coleman indicated that school principals must be better equipped with knowledge, managerial skills, and sources of assistance to solve difficult personnel difficulties, such as confronting and intervening in instances of staff underperformance. Additionally, Weisberg et al. proclaimed that administrators must marry the institutional goal of increasing student achievement with the personal aspects of differentiating between teacher performance via actionable feedback by holding deep reverence for the teaching profession and evaluation process, stating: “Improved evaluation will not only benefit students by driving the systematic improvement and growth of their teachers, but teachers themselves, by at last treating them as professionals, not parts” (p. 8).

Summary

Although teachers have the greatest impact on student achievement, high-quality administrators who exhibit exemplary leadership and management skills also have an immense effect on student success with regard to making decisions about which teachers are hired and allowed to maintain their positions in the classroom. Additionally, there has been little historical evidence to support that evaluation systems, and the feedback held therein, are used as effective tools to provide a solid link between teacher evaluation and student achievement (Steinberg & Donaldson, 2014). This review of the literature encompassed an understanding of the context of teacher evaluation in the current era of reform, an explanation of the observed Lake Wobegon Effect in prior studies conducted throughout the United States with respect to teacher evaluations, and three dimensions that encompass the various barriers to candid teacher evaluation: (a) individual conflict avoidance; (b) bureaucratic procedural interferences; and (c) administrative procedural faults. Ineffective teachers remaining in the classroom has resulted, in part, from a dysfunctional system and culture present within schools and school leadership. The present study was conducted to discover if lack of leadership is a positive indicator of a dysfunctional system that permits ineffective teachers to remain in the classroom. The findings of prior researchers described in this review of literature were used to guide the investigation.

CHAPTER 3 METHODOLOGY

Introduction

This chapter contains a presentation of the methods and procedures used to conduct the study. Included is a restatement of the purpose as well as the research questions that guided the study and the various instruments were used to measure both quantitative and qualitative research methods used in this study. The methods and procedures that were used in the collection of data and the analysis of the data are also explained in detail.

Purpose of the Study

The purpose of this study was to determine if grossly ineffective teachers, according to Florida VAM scores, had related instructional practice evaluations that were being used to determine placement on improvement plans. Additionally, this study served to determine if there was a difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teacher instructional practice scores and assignment to improvement plans for teachers with the lowest 10% of VAM scores in Central Florida school districts. Finally, the levels of feedback held within the local district performance evaluations and improvement plans were reviewed to search for evidence of administrative barriers within the three dimensions of (a) individual conflict avoidance, (b) bureaucratic procedural interferences, and (c) administrative procedural faults that exist and inhibit administrative evaluators from providing candid and actionable teacher performance evaluations relative to student achievement.

Research Questions

The following research questions and hypotheses guided the investigation of this study:

1. What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?

H₀₁: There is no observed relationship between the lowest 10% of teachers' VAM scores as correlated to local instructional practice scores within Central Florida school districts.

2. What difference, if any, exists between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts?

H₀₂: There is no observed difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts.

3. What percentage of professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers identified with the 10% of lowest VAM scores in Central Florida school districts have been placed on improvement plans?

4. Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida School districts?

Selection of Subjects and Population

Subjects of this study were teachers with the lowest 10% of VAM scores for the 2013-14 school year within the counties represented by the Central Florida Public School Boards Coalition (CFPSBC). The included school districts in the CFPSBC were: Brevard, Highlands, Hillsborough, Lake, Manatee, Marion, Orange, Osceola, Pasco, Polk, Seminole, Sumter, and Volusia.

To determine the subjects to be selected for use in responding to Research Questions 1-3 of this study, a list of VAM scores for the 2013-14 school year was obtained from the Florida Department of Education. The VAM score data were limited by exclusion of some teacher names. According to an information file held within the documents delivered to the researcher, some individual teacher names were masked, (e.g., for any teacher who is exempt from public record disclosure per Florida law). The Florida Department of Education cited the following Florida Statutes for these exemptions: §119.071 (2)(j), (4)(d), and (5)(i); §265.605; and §267.17. Due to some teachers' names being masked, the researcher excluded incomplete data cases to obtain a numerical list of VAM scores from greatest to least. Next, the lowest 10% of VAM scores were obtained from this numeric list and selected for further analysis. Finally, the researcher organized the list by Florida school district and included only teachers from

the 13 school districts within the CFPSBC. Table 1 describes the population of included subjects for Research Questions 1-3. The data file yielded a total of 758 teachers who were selected for personnel file data examination with varying frequencies and percentages held within the data for each of the 13 counties represented.

Table 1

Characteristics of Study Population

Central Florida County	District Size by Student Enrollment	District Grade	Frequency of Teachers within Lowest 10% of VAM Scores	% of Selected Population by Central Florida County
Brevard	71,234	B	24	3.2
Highlands	12,199	C	11	1.5
Hillsborough	203,432	B	195	25.7
Lake	41,789	C	28	3.7
Manatee	46,703	C	20	2.6
Marion	42,107	C	34	4.5
Orange	187,092	B	99	13.1
Osceola	58,203	C	60	7.9
Pasco	68,103	C	54	7.1
Polk	97,957	C	121	16.0
Seminole	64,846	A	41	5.4
Sumter	8,281	B	17	2.2
Volusia	61,237	C	54	7.1
Total	963,183	N/A	758	100.0

Note. VAM = Value-added model.

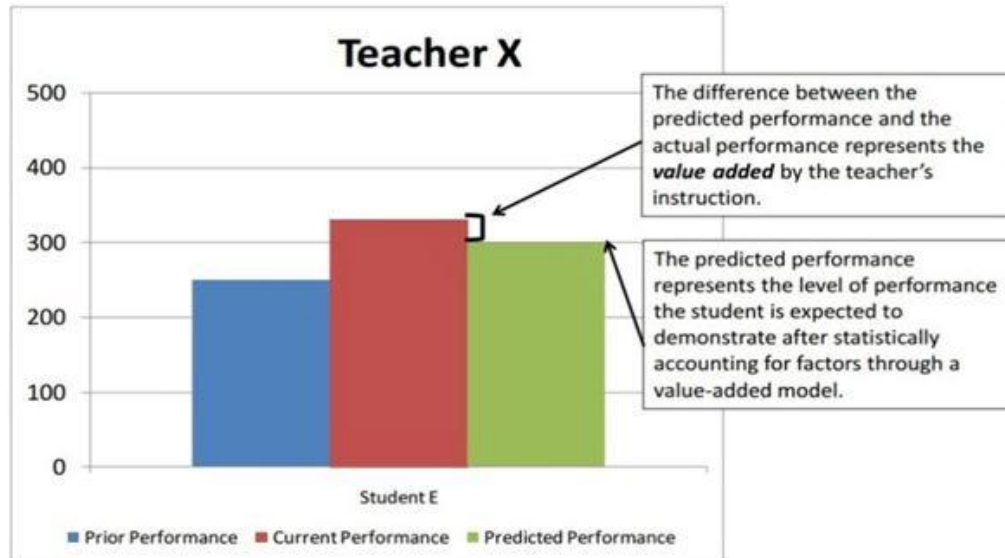
Instrumentation

Florida's Value-Added Model (VAM)

Florida has used a value-added model (VAM) measurement to determine the contribution of a teacher or school to student learning. The VAM measurement determines the difference in student performance on a statewide assessment from one year to the next, and accounts for other factors specific to student, classroom, and school characteristics that impact the learning process. Florida uses VAM scores for teachers who are coded as teaching a course that includes a statewide assessment. Therefore, the participants in this study were only teachers who taught a course or grade level with an associated statewide assessment in the 2013-14 school year.

The VAM score determines the difference between the predicted performance and the actual performance of a student on a statewide assessment. Value-added model scores can be negative, positive, or zero. A score of zero means that the students on a specific teacher's roster scored exactly as predicted. A negative score means that student scores fell below the predicted score. Conversely, a positive score means that student scores exceeded the expectation. For example, if a teacher received a VAM score of negative 0.30, it would mean that on average, the teacher's students scored 30% below the state average growth for that grade and subject. Alternatively, if a teacher received a VAM score of positive 0.30, it would mean that on average, the teacher's students scored 30% above the state average growth for that grade and subject. Figure 1 graphically details how Florida VAM scores are determined.

Value-Added Example



Note. Reproduced with permission from “Florida’s value-added models (VAM) frequently asked questions”, Florida Department of Education, 2015.

Figure 1. Florida Value-added Example of Score Determination

For the purpose of this study, VAM scores were delivered to the researcher in an online database that reflected the values described in this section. Value-added scores were then sorted from greatest to least, and the lowest 10% of VAM scores were identified. Once the lowest 10% of VAM scores by teacher were determined, the researcher selected for teachers only within the 13 Central Florida counties represented by the CFPSBC, yielding a total of 758 subjects for the purpose of investigating Research Questions 1-4, utilizing secondary de-identified data from publicly available databases under Florida Statute §119.07.

Florida Instructional Personnel Evaluation Requirements

Florida Statute §1012.34(3) designates evaluation procedures and criteria for instructional personnel and administrators. According to the statute, instructional personnel and administrator performance evaluations must be based upon the performance of students assigned to their classrooms or schools. Additionally, a school district's performance evaluation system is not limited to basing unsatisfactory performance of instructional personnel and school administrators solely upon student performance, but may include other criteria to evaluate instructional personnel and school administrators' performance, or any combination of student performance and other criteria (Florida Statute §1012.34 (3), 2015). Evaluation procedures must comply with the following: at least one-third of the overall teacher performance evaluation must be based upon data indicators of student performance via VAM scores. At least one-third must be based on an instructional practice score decided by school-based evaluators, and the final one-third may be compiled considering other indicators of performance such as professional and job responsibilities recommended by the State Board of Education or identified by the district school board, peer reviews, objectively reliable survey information from students and parents based on teaching practices that are associated with higher student achievement, and other valid and reliable measures of instructional practice (Florida Statute §1012.34 (3), 2015).

According to Florida Statute §1012.34, performance evaluation system requirements include an evaluation system that has been approved by the Florida Department of Education for the purpose of increasing student academic performance by

improving the quality of instructional, administrative, and supervisory services in the public schools of the state (Florida Statute §1012.34 (1)(a), 2015). Additionally, evaluation system requirements for instructional personnel and school administrators must be designed to support effective instruction and student learning growth; and performance evaluation results must be used when developing district and school-level improvement plans (Florida Statute §1012.34 (2)(a), 2015). Furthermore, appropriate instruments must be utilized, procedures must be in place, timely feedback must be rendered with criteria for continuous quality improvement of the professional skills of instructional personnel and school administrators. Performance evaluation results must be used when identifying professional development (Florida Statute §1012.34 (2)(b), 2015). Moreover, Florida Statute §1012.34 (2)(e) states that performance evaluations must differentiate among four levels of performance as follows: (1) highly effective, (2) effective, (3) needs improvement or, for instructional personnel within the first three years of employment who need improvement, developing, and (4) unsatisfactory. Each of the 13 Central Florida school districts that were selected for this study have a state-approved teacher evaluation system. To fulfill the statutory requirements, Lake, Orange, Osceola, Pasco, and Seminole school districts have used the state model of teacher evaluation based on the research and meta-analyses of Marzano. Highlands, Hillsborough, Manatee, Marion, Sumter, and Volusia school districts have used the Danielson model and may have adopted principles from the state model, including the scoring system and deliberate practice plan. Finally, Brevard and Polk school districts selected another state-approved model to fulfill the statutory requirements and may

include indicators from the state model. Overall, all three instructional practice options for state-approved teacher evaluation models differentiate between the four levels of summative performance ratings as prescribed by Florida Statute §1012.34 (2)(e). For the purpose of this study, and to respond to Research Question 1, evaluation ratings of (1) highly effective, (2) effective, (3) needs improvement or developing, and (4) unsatisfactory were used to determine if the instructional practice score related to the matching VAM score for the selected subjects. Instructional practice scores for selected subjects were obtained to determine if there was a statistically significant difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) instructional practice quantitative ratings to respond to Research Question 1. Additionally, whether or not an improvement plan was included in the instructional practice evaluation for the selected subjects by tenured or nontenured contract status was investigated to respond to Research Question 3. Finally, the highest level of feedback within instructional practice evaluations and improvement plans, as described by the rubric developed by Rafalski (2015), was used to respond to Research Question 4.

Florida Instructional Personnel Contract Types

The Florida Department of Education has been governed by two statutes that describe the different types of instructional personnel contracts that may be offered to teachers. The first type of contract is a probationary contract which may be offered for a period of one school year to instructional personnel upon initial employment in a school district (Florida Statute §1012.335(1)(c), 2015). The second type of contract is defined as

an annual contract which is the only contract that may be offered to instructional personnel hired after July 1, 2011. An annual contract can be offered for a period of one school year, and the district school board may choose to award or not to award without cause (Florida Statute §1012.335(1), 2015). The third type of contract is a professional service contract which was only offered to instructional personnel hired prior to July 1, 2011 and is renewed each year unless the district school superintendent, after receiving recommendations required by Florida Statute §1012.34, charges the employee with unsatisfactory performance. In this case, the superintendent notifies the employee of performance deficiencies as required by the statute; the employee receives two consecutive annual performance evaluation ratings of unsatisfactory within a three-year period; or three consecutive performance evaluations of needs improvement or a combination of needs improvement or unsatisfactory (Florida Statute §1012.33 (2)(c) & Florida Statute §1012.33 (3)(a)(b), 2015). The fourth and final contract type, a continuing contract, is less common. A continuing contract may be held by any employee in the state of Florida who held continuing contract status prior to July 1, 1984 (Florida Statute §1012.33(4)(a), 2016). Any Florida teacher holding continuing contract status prior to July 1, 1984 shall be entitled to retain such contract and all rights arising from it as prescribed by the State Board of Education, unless the employee voluntarily relinquishes his or her continuing contract (Florida Statute §1012.33(4)(a), 2016). Any member of the district administrative or supervisory staff and any member of the instructional staff, including the school principal, who is under continuing contract may be suspended or dismissed at any time during the school year; however, the charges

against him or her must be based on immorality, misconduct in office, incompetency, gross insubordination, willful neglect of duty, drunkenness or being convicted or found guilty of, or enter a plea of guilty to, regardless of adjudication of guilt, any crime involving moral turpitude, as these terms are defined by rule of the State Board of Education (Florida Statute §1012.33(4)(c), 2016). All contracts are subject to the same aforementioned annual performance evaluation requirements. To investigate Research Questions 2 and 3, instructional personnel were divided by (a) professional service/continuing contract status (tenured) and (b) annual contract/probationary contract status (nontenured) to determine if there was a difference in how tenured and nontenured teachers were rated by administrative personnel and assigned to improvement plans as evidenced by instructional practice evaluations.

Levels of Feedback

Research Question 4 required a protocol and rubric to review the highest levels of feedback provided for within the instructional practice evaluations and improvement plans (if present) for each of the 758 subjects in this study. The following rubric and accompanying definitions, developed by Rafalski (2015) in a prior study and used with her permission, were used to gather data on feedback categories and levels.

Level 1 - No feedback: The observer provides no opinion in the comment section of the protocol.

Level 2 - Unrelated feedback or general statement: The observer gives some information in the comment section, but it is not relevant to the element or meaning cannot be interpreted.

Level 3 - Recount of observation events: This could include a narrative of what the teacher and students were doing during the observation, general statements of events, or notes the observer took to justify the rating given. In some instances, the observer included statements to support the effectiveness of a strategy.

Level 4 - General affirmation or praise statement: The observer either leaves a single word or phrase to indicate approval or adds a compliment to the end of a recount of observation of events.

Level 5 - Reflective feedback: The observer asks the teacher to think about the practice or a specific element in either a general or specific way.

Level 6 - Standardized feedback: The observer uses the cut and paste option in the protocol to leave systematized feedback.

Level 7 - Specific targeted feedback: The observer leaves differentiated and meaningful statements intended to improve the impact of an instructional strategy.

Data Collection

Prior to initiating data collection, the proposal for the study was examined by University of Central Florida Institutional Review Board (IRB), and it was determined that the study did not include human research (Appendix A). Thus, University of Central Florida IRB review and approval were not required as all data collected were de-identified, and secondary information was obtained from publically available databases. Value-added model scores were obtained from the Florida Department of Education (FDOE) through the Division of Accountability, Research, and Measurement via a public

record request made directly to the FDOE. Instructional practice scores, teacher improvement plans, and evaluation feedback were obtained through the district offices of each respective school district, by submitting formal public record requests (Appendix B) to each individual district office pursuant to Florida Statute §119.07.

Data Analysis

This study was guided by an interest in determining if grossly ineffective teachers, according to Florida VAM scores, had related instructional practice evaluations that were being used to determine placement on improvement plans. Additionally, this study served to determine if there was a difference between tenured and nontenured teacher instructional practice scores for teachers with the lowest 10% of VAM scores in 13 Central Florida school districts. Finally, the highest levels of feedback held within the local district performance evaluations and improvement plans were reviewed to search for evidence of administrative barriers within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults that exist and inhibit administrative evaluators from providing candid teacher performance evaluations relative to student achievement. The research questions and hypotheses governed the selection of statistical measures and analytical procedures to be used, as described in Table 2.

Table 2

Research Questions, Research Methods, Variables, Data Sources, and Data Analysis

	Research Questions	Research Methods	Variables	Data Sources	Analysis
1.	What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?	Quantitative	VAM scores and instructional practice scores	VAM scores and instructional practice scores	Pearson r Correlation
2.	What difference, if any, exists between PSC/CC and AC/PC teachers' instructional practice scores within Central Florida school districts?	Quantitative	Instructional practice scores and teacher contract statuses	Instructional practice scores	Independent samples t test, Cohen's d
3.	What percentage of PSC/CC and AC/PC teachers identified within the 10% of lowest VAM scores in Central Florida school districts have been placed on improvement plans?	Quantitative	Instructional practice evaluation content. Y/N contains improvement plan	Instructional practice evaluations	Descriptive statistics, frequencies and percentage of teachers assigned improvement plans by contract status and total population
4.	Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida school districts?	Qualitative	Feedback levels	Feedback content from instructional practice evaluations	Descriptive statistics, including a frequency distribution of feedback rubric ratings to illustrate findings

Quantitative data for instructional practice and student achievement (VAM scores) were collected, analyzed, and reported for a correlational analysis. The mean and standard deviation of the instructional practice scores for the subjects were compared using an independent samples *t*-test to determine whether there was a significant difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' scores within the lowest 10% of VAM scores in the represented school districts within the study. Frequencies and percentages were used to determine the number of teachers within the population of subjects selected who were placed on improvement plans by contract status (tenured or nontenured). Finally, descriptive statistics including a frequency distribution, range, mean, and standard deviation were used to provide a quantitative measure of the highest levels of feedback provided within instructional practice evaluations. This was accomplished using the method and instrument of data analysis developed by Rafalski (2015) with her explicit permission. A sample of Rafalski's rubric for feedback determination rating is included in Appendix C.

Summary

In this chapter, the purpose of this study and the research questions were restated. The selection of subjects, instrumentation, data collection procedures, and data analysis processes were also discussed. The selection of subjects yielded 758 teachers from 13 Central Florida counties to provide data for investigation of the research questions. Results of the data analysis are presented in Chapter 4.

CHAPTER 4 PRESENTATION AND ANALYSIS OF DATA

Introduction

The purpose of this study was to determine if grossly ineffective teachers, according to Florida VAM scores, had related instructional practice evaluations that were used to determine placement on improvement plans. Additionally, this study served to determine if there was a difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teacher instructional practice scores and assignment to improvement plans for teachers within the lowest 10% of VAM scores in 13 Central Florida school districts. Finally, the levels of feedback held within the local district performance evaluations and improvement plans were reviewed to search for evidence of administrative barriers within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults that exist and inhibit administrative evaluators from providing candid and actionable teacher performance evaluations relative to student achievement. This chapter presents the results of descriptive and inferential analyses for the four stated research questions.

The descriptive statistics for the population investigated were first reported, followed by the results of a Pearson r correlation employed to respond to Research Question 1, an independent samples t -test used to respond to Research Question 2, and descriptive statistics, frequencies, and percentages to respond to Research Questions 3 and 4. Finally, additional supporting analyses were reported for each school district investigated using descriptive statistics. Narrative descriptions, as well as

complementary tables and figures, were used in responding to each of the research questions. A summary of key findings concludes Chapter 4.

Overall Descriptive Statistics

Upon completion of data collection, the total population for this study decreased from 758 teachers in 13 school districts within Central Florida, to 528 teachers in 11 school districts within Central Florida, or 69.66% of the total population. Reasoning for the elimination of 230 teachers was due to a variety of conditions, such as: (a) incomplete personnel files held within school district central offices, yielding missing information for selected subjects; (b) elimination of charter school employees due to charter school personnel files not being held by school district central offices; (c) termination of employees prior to completion of the 2013-14 school year summative instructional practice evaluations; (d) retirement of employees prior to the completion of the 2013-14 school year summative instructional practice evaluations; and (e) duplicates of teacher names within the original value-added model (VAM) score data from the Florida Department of Education (FDOE) due to part-time employment at multiple schools. Both Marion County and Lake County teachers were disregarded due to lack of contract statuses shared by these school district central offices via public record request. The overall frequencies for the population of subjects who were included in this study as compared to the total population determined prior to data collection are illustrated in Table 3.

Table 3

Pre- and Post-Data Collection Frequencies of Study Teachers by County: 2013-14

Central Florida County	Frequency of Teachers in Lowest 10% of VAM Scores 2013-14	
	Pre-data Collection	Post-data Collection
Brevard	24	20
Highlands	11	7
Hillsborough	195	165
Lake	28	0
Manatee	20	8
Marion	34	0
Orange	99	80
Osceola	60	48
Pasco	54	46
Polk	121	63
Seminole	41	36
Sumter	17	17
Volusia	54	38
Total	758	528

Research Question 1

What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?

H₀₁: There is no observed relationship between the lowest 10% of teachers' VAM scores as correlated to local instructional practice scores within Central Florida school districts.

The Florida Department of Education provided VAM scores of all Florida teachers that were numerically ordered from least to greatest to determine the teachers in Central Florida school districts within the lowest 10% of VAM scores for the entire state

for the 2013-14 school year. Instructional practice scores were obtained through public records requests made to each Central Florida school district in accordance with Florida Statute §119.07. A total of 528 teachers within the total population of 758 yielded complete data sets for all investigations included in this study, or 69.66% of the total population.

Value-added model scores ranged between -2.28 and -0.63, which were correlated with instructional practice scores that ranged from one to four (1 = highly effective, 2 = effective, 3 = needs improvement/developing, and 4 = unsatisfactory). Value-added model scores displayed a mean of -0.90 and a standard deviation of 0.27. A histogram of VAM score frequencies within the population (N = 528) illustrated negative skewness as shown in Figure 2. Therefore, the greatest frequency of VAM scores were between -1.16 and -0.63, with some outliers as negative as -2.28.

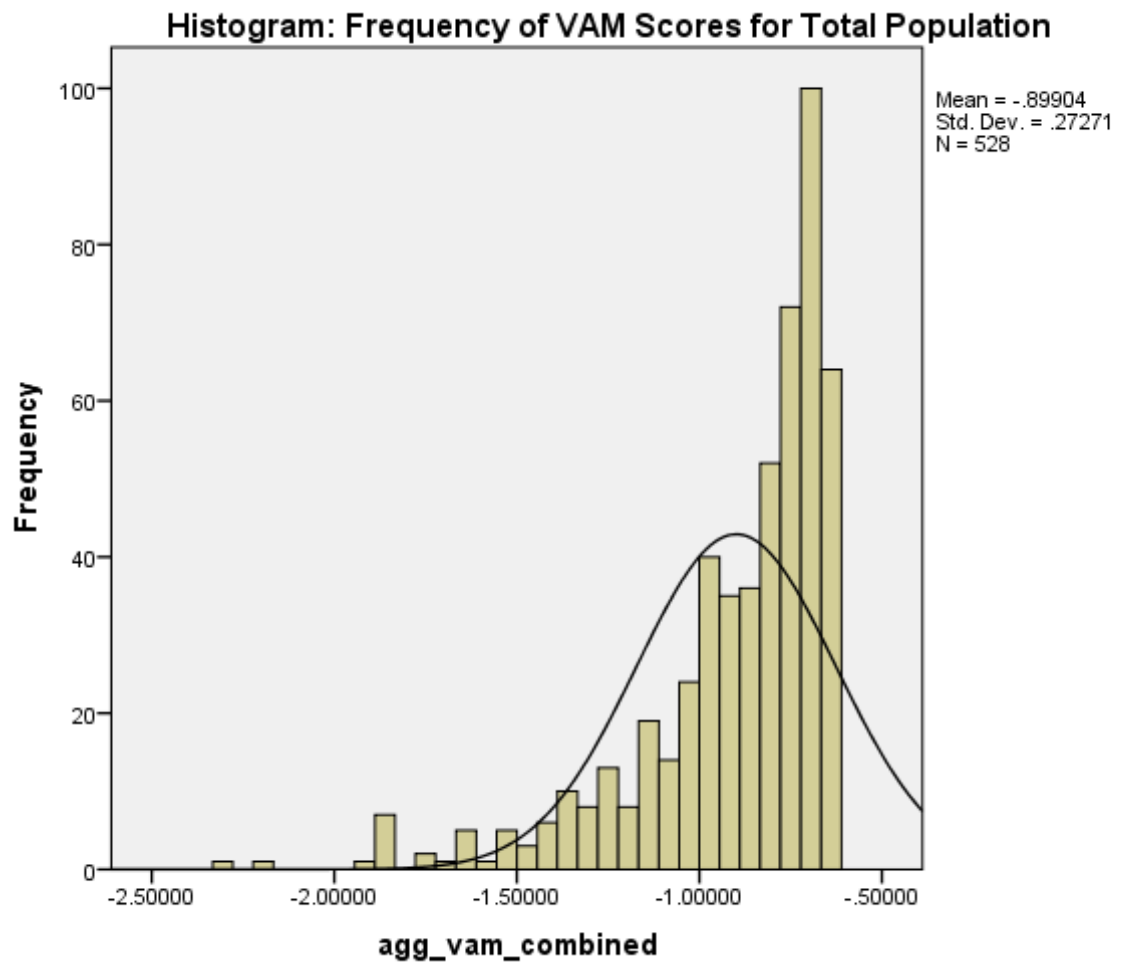


Figure 2. Histogram of Value-added Model (VAM) Score Frequencies Observed

Alternatively, instructional practice scores for the population displayed a mean of 1.65 (between highly effective and effective), and a standard deviation of 0.65 (N = 528). Descriptive statistics provided the frequency of instructional practice scores observed. Of the 528 summative instructional practice evaluations investigated, 43.2% of teachers were rated highly effective, 51.1% of teachers were rated effective, 3.6% of teachers were rated needs improvement/developing, and 2.1% of teachers were rated unsatisfactory. Table 4 presents the frequencies and percentage of teachers in each

instructional practice rating category. Additionally, a histogram of instructional practice score frequencies within the population illustrates positive skewness as shown in Figure 3. Hence, the greatest frequency of teachers were rated (1) highly effective and (2) effective, with few being rated (3) needs improvement/developing or (4) unsatisfactory.

Table 4

Frequencies and Percentage of Teachers by Instructional Practice Rating Category

Instructional Practice Rating	<i>f</i>	%
Highly Effective	228	43.2
Effective	270	51.1
Needs Improvement/Developing	19	3.6
Unsatisfactory	11	2.1
Total	528	100.0

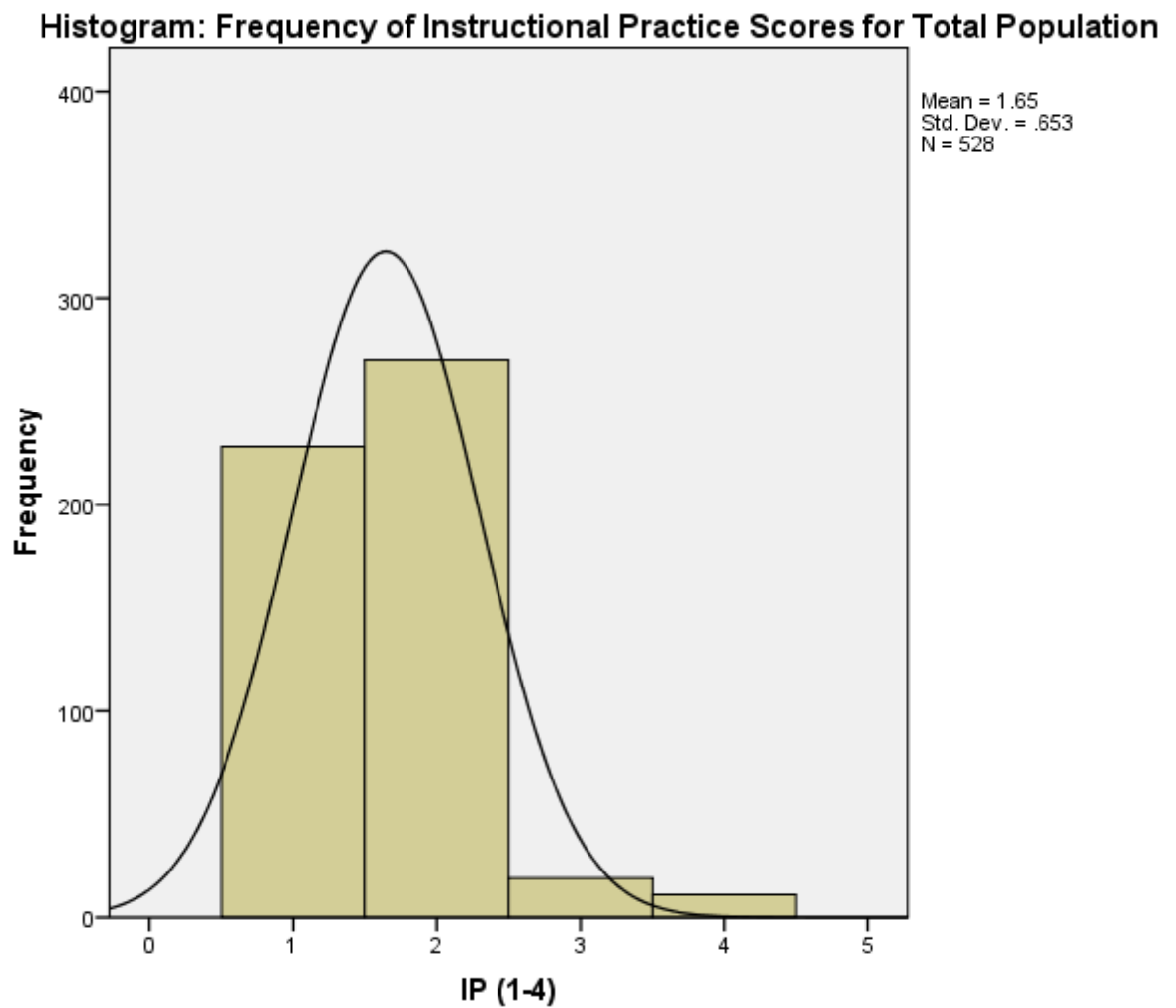


Figure 3. Histogram of Instructional Practice Score Frequencies Observed

A Pearson r correlation was used to determine if a relationship existed between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts. The result of this analysis produced a negative correlation: $r(526) = -.104, p = 0.016$. The null hypothesis was rejected at the $p < .05$ level; a very weak relationship existed between the lowest 10% of teachers' VAM scores and instructional practice scores in Central Florida school districts in the 2013-14 school year. However, the null hypothesis would be accepted at the $p < .01$ level, indicating no relationship existed between these two variables. A negative correlation was observed due to instructional practice scores being 1 = highly effective, 2 = effective, 3 = needs improvement/developing, and 4 = unsatisfactory. Therefore, a lower instructional practice numerical score designated a higher rating in terms of observed effectiveness by an administrative evaluator. Consequently, a negative correlation indicated instructional practice ratings were less favorable as correlated to increasingly negative VAM scores. However, this relationship was observed to be very weak at the $p < .05$ level (rejecting the null hypothesis), and no relationship was observed at the $p < .01$ level (accepting the null hypothesis). This correlational analysis is presented in Table 5.

Table 5

Pearson Correlation Coefficients: VAM Scores and Instructional Practice Scores

Descriptor	Instructional Practice Scores
Value-added (VAM) scores	-0.104*

Note. * $p = 0.016$

The significance of this correlation at the $p < .05$ level is likely due to a large sample size with the result of the correlation unlikely to have arisen by chance. Figure 4 displays a scatterplot to support the Pearson r correlational findings, illustrating plots oriented in a parallel fashion for each instructional practice score (1 = highly effective, 2 = effective, 3 = needs improvement/developing, and 4 = unsatisfactory), with a greater distribution of parallel plots located above the (1) highly effective and (2) effective categories on the x-axis, regardless of the VAM score rating spread. Consequently, regardless of the relative negativity of any particular VAM score, administrative evaluators rated teachers as (1) highly effective, and (2) effective most frequently on the instructional practice portion of the overall summative teacher evaluation.

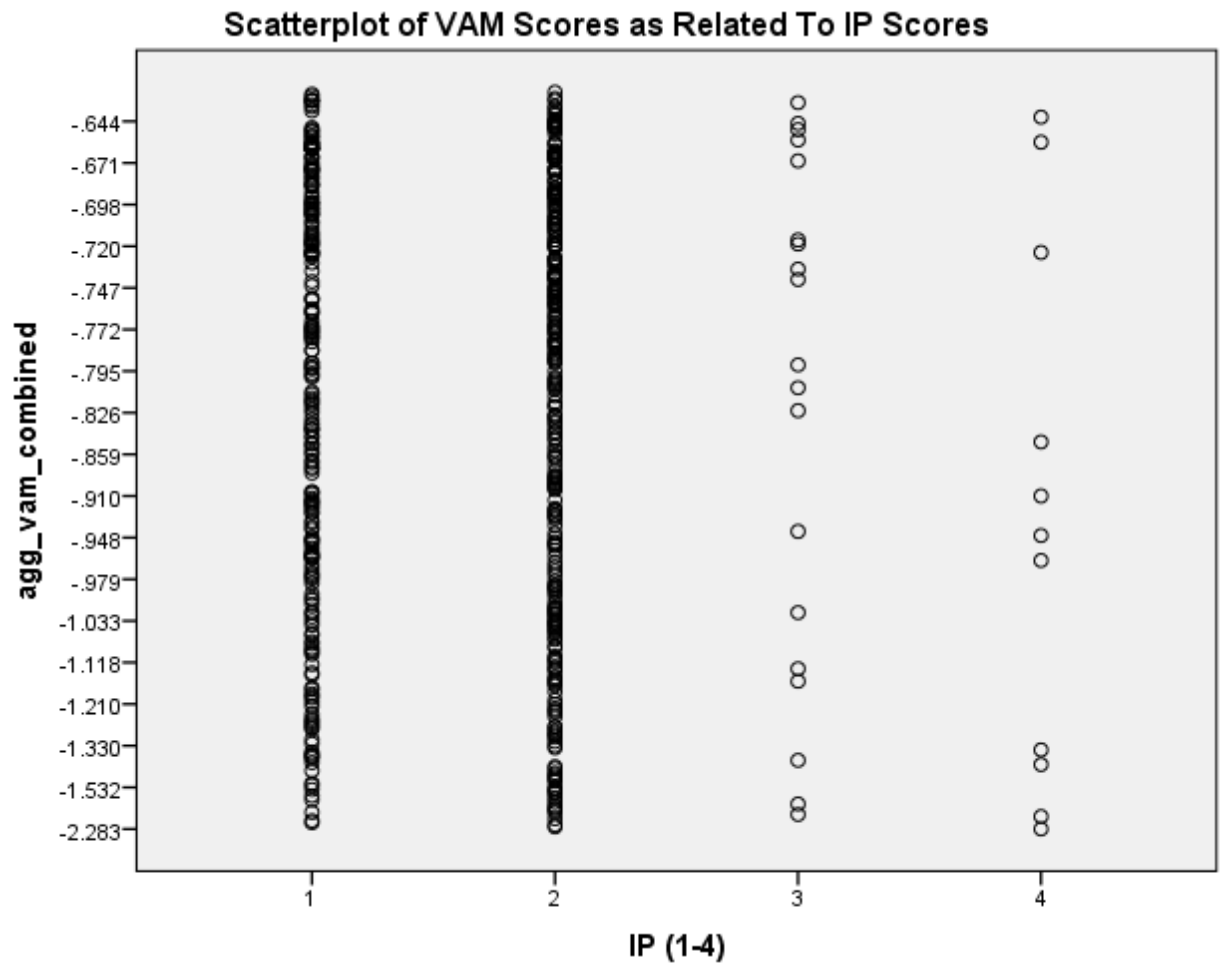


Figure 4. Scatterplot of Value-added (VAM) Scores as Related to Instructional Practice Scores

Research Question 2

What difference, if any, exists between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts?

H₀: There is no observed difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts.

A total of 11 Central Florida school districts provided summative evaluation data for the 2013-14 school year via public record requests in accordance with Florida Statute §119.07 for 528 teachers within the total population included in this study. Contract statuses of (a) professional service contract, (b) continuing contract, (c) annual contract, and (d) probationary contract were observed throughout the population. For the purpose of this study, professional service contract and continuing contract teachers were considered to be tenured, whereas annual contract and probationary contract teachers were considered to be nontenured. As shown in Table 6, of the total population of 528 teachers, 278 (52.7%) teacher evaluations examined were those of tenured teachers, and the remaining 250 (47.3%) evaluations were those of nontenured teachers.

Table 6

Frequency and Percentage of Teachers by Contract Status

Contract Status	<i>f</i>	%
Tenured: Professional Service Contract & Continuing Contract	278	52.7
Nontenured: Annual Contract & Probationary Contract	250	47.3
Total	528	100

To determine if there was a significant difference between the instructional practice scores for tenured and nontenured teachers, an independent samples *t*-test was performed. The results of the analysis are displayed in Table 7. The mean instructional practice score for the 278 tenured teachers within this study was 1.56 with a standard deviation of 0.578. The mean instructional practice score for the 250 nontenured teachers within this study was 1.74, with a standard deviation of 0.715. The overall mean for the total population was 1.65, with a standard deviation of 0.653.

Table 7

Mean and Standard Deviation for Instructional Practice Scores of Tenured versus Nontenured Teachers

Contract Status	<i>N</i>	Mean	Standard Deviation
Tenured	278	1.56	0.578
Nontenured	250	1.74	0.715
Total	528	1.65	0.653

The result of the independent samples *t*-test yielded a *t* value of -3.306, with a *p* value of 0.001 at 526 degrees of freedom; $t(526) = -3.306$, $p = 0.001$. Therefore, the null hypothesis was rejected; there was a significant observed difference between the instructional practice scores of tenured and nontenured teachers. Although there was a statistically significant difference between the mean instructional practice scores for tenured versus nontenured teachers, a Cohen's *d* effect size of -0.275 was calculated, and it was determined that this finding was of low practical importance. In brief, tenured

teachers were rated more favorably than nontenured teachers on the instructional practices portion of the summative teacher evaluation. Table 8 displays these data.

Table 8

Independent Samples Test Comparing Mean Instructional Practice Scores for Tenured vs. Nontenured Teachers

Descriptor	<i>t</i>	<i>df</i>	Significance (2-tailed)	Mean Difference	Standard Error Difference	95% Confidence Interval	
Equal Variances Assumed	-3.306*	526	.001	-.186	.056	Lower	Upper
						-.297	-.076

* $p < .01$

Research Question 3

What percentage of professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers identified with the lowest 10% of VAM scores in Central Florida school districts have been placed on improvement plans?

A total of 11 Central Florida school districts provided summative evaluation data for the 2013-14 school year via public record requests in accordance with Florida Statute §119.07 for the 528 teachers in the total population included in this study. Within the summative evaluation data collected, two teachers were prescribed improvement plans for the 2013-14 school year, representing 0.38% of the total population. One improvement plan was prescribed for a tenured teacher (0.40%, $N = 278$), and one improvement plan was prescribed for a nontenured teacher (0.36%, $N = 250$). Figure 5

consists of a pie chart which shows the percentage of teachers for whom an improvement plan was prescribed within the total population observed.

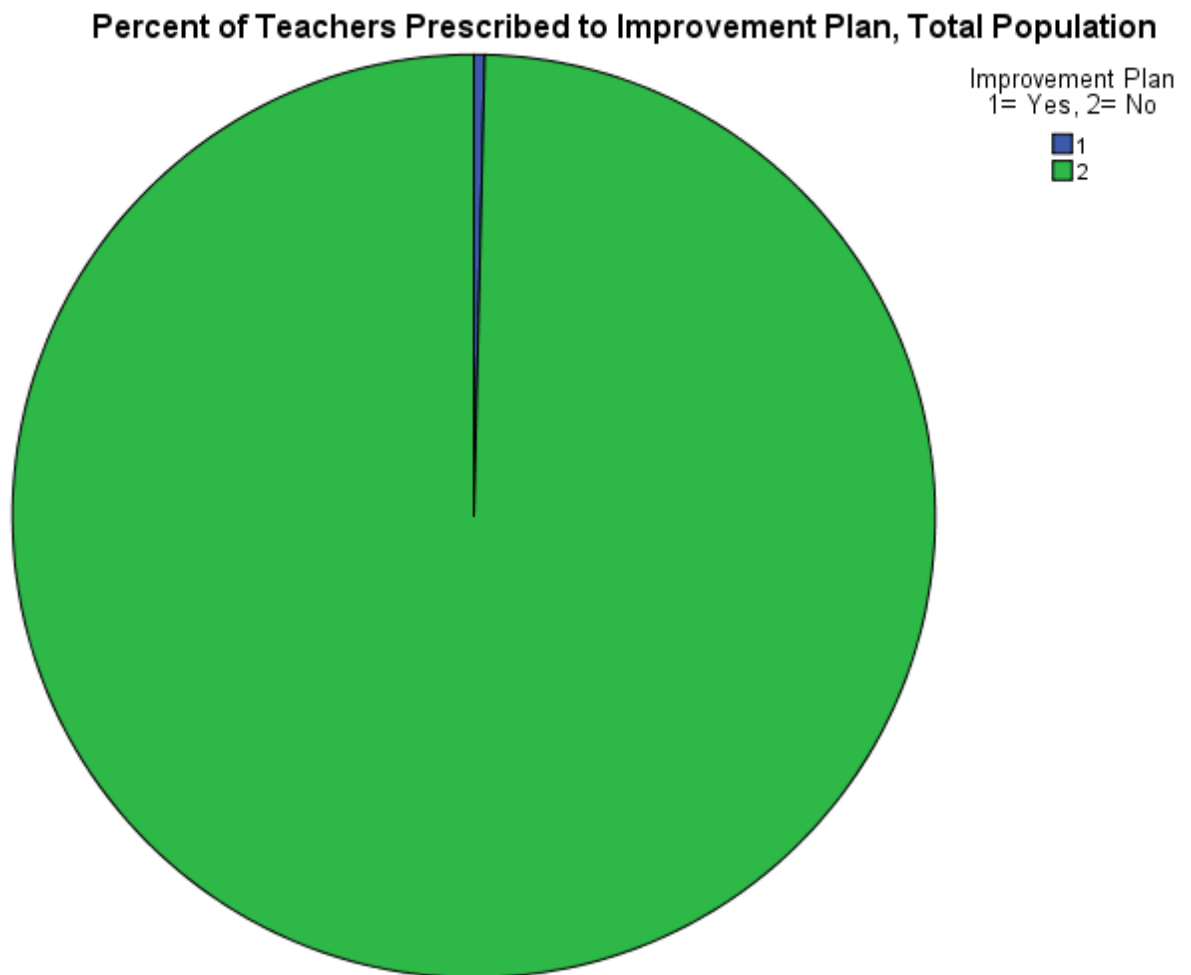


Figure 5. Pie Chart Representing Percentage of Teachers Prescribed an Improvement Plan

Research Question 4

Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida school districts?

A total of 11 Central Florida school districts provided summative evaluation data for the 2013-14 school year via public record requests in accordance with Florida Statute §119.07 for 528 teachers within the total population included in this study. Within the summative evaluation data were various feedback statements delivered to each teacher regarding instructional performance for multiple categories assumed to define teacher effectiveness. Each school district had a different evaluation system, and the highest levels of feedback for each individual's overall summative evaluation were examined for analysis. The following rubric and accompanying definitions were used to gather data on feedback categories and levels, developed and used with permission from a prior study conducted by Rafalski (2015):

Level 1 - No feedback: The observer provides no opinion in the comment section of the protocol.

Level 2 - Unrelated feedback or general statement: The observer gives some information in the comment section, but it is not relevant to the element or meaning cannot be interpreted.

Level 3 - Recount of observation events: This could include a narrative of what the teacher and students were doing during the observation, general statements of events, or notes the observer took to justify the rating given. In some instances, the observer included statements to support the effectiveness of a strategy.

Level 4 - General affirmation or praise statement: The observer either leaves a single word or phrase to indicate approval or adds a complement to the end of a recount of observation of events.

Level 5- Reflective feedback: The observer asks the teacher to think about the practice or a specific element in either a general or specific way.

Level 6- Standardized feedback: The observer uses the cut and paste option in the protocol to leave systematized feedback.

Level 7- specific targeted feedback: The observe leaves differentiated and meaningful statements intended to improve the impact of an instructional strategy.

Table 9 presents the highest levels of feedback observed for the total population of teachers observed for this investigation ($N = 528$). Of the 528 teacher evaluations analyzed, the highest levels of feedback percentages were as follows: 47.5% Level 1 (no feedback); 6.3% Level 2 (unrelated feedback or general statement); 18.4% Level 3 (recount of observation events); 12.5% Level 4 (general affirmation or praise statement); 2.1% Level 5 (reflective feedback); 1.7% Level 6 (standardized feedback); and 11.6% Level 7 (specific targeted feedback).

Table 9

Frequencies and Percentages of Highest Feedback Levels Observed

Feedback Level	<i>f</i>	%	Cumulative %
1	251	47.5	47.5
2	33	6.3	53.8
3	97	18.4	72.2
4	66	12.5	84.7
5	11	2.1	86.7
6	9	1.8	88.4
7	61	11.6	100.0
Total	528	100.0	100.0

Summary

In this chapter, data were analyzed to respond to four research questions associated with the analysis of low performing Central Florida teacher evaluation feedback and improvement plans as related to value-added model scores and instructional practice scores. Descriptive and inferential statistics were used in the analysis.

Research Question 1 addressed what relationship, if any, existed between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts for the 2013-14 school year. The actual population of teachers, for whom data were included in this study, was reduced to 528 from the initial total of 758 teachers due to a variety of limiting factors. To determine the magnitude of association between VAM scores and instructional practice scores, a Pearson r correlation coefficient was calculated, resulting in a very low (near zero) correlation: $r(526) = .104, p = 0.016$. Therefore, the null hypothesis was rejected at the $p < .05$ level.

There was a very weak relationship between VAM scores and instructional practice scores for teachers within the lowest 10% of VAM scores for the 2013-14 school year.

Research Question 2 focused on what difference, if any, existed between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts for the 2013-14 school year. From the total population of teachers with complete data sets ($N = 528$), 278 teachers were considered tenured, and 250 teachers were considered nontenured. To determine the difference between the mean instructional practice scores of tenured versus nontenured teachers, an independent samples t -test was used, resulting in the null hypothesis being rejected. There was a significant difference between the mean instructional practice scores of tenured versus nontenured teachers at 526 degrees of freedom with a t value of -3.306 ($p = 0.001$): $t(526) = -3.306, p < .01$. Although this finding was significant, the Cohen's d effect size was calculated to be -0.275, providing evidence of low practical and clinical importance.

Research Question 3 investigated the percentage of professional service contract/continuing contract (tenured) teachers and annual contract/probationary contract (nontenured) teachers within the population who were placed on improvement plans. Descriptive statistics were used to calculate frequencies and percentages of tenured and nontenured teachers who were prescribed improvement plans. Two improvement plans were found within the personnel files of the investigated population. One improvement plan was for a tenured teacher, and one improvement plan was for a nontenured teacher, resulting in 0.40% of tenured teachers ($N = 278$), and 0.36% ($N = 250$) nontenured

teachers within the lowest 10% of VAM scores in Central Florida for the 2013-14 school year being placed on improvement plans. The overall percentage of teachers being placed on improvement plans for the entire population was 0.38%, ($N = 528$).

Research Question 4 was used to demonstrate if the feedback provided in instructional practice evaluations and improvement plans targeted student achievement for teachers with the lowest 10% of VAM scores within Central Florida for the 2013-14 school year. Descriptive statistics, including frequencies and percentages of highest feedback levels as described by Rafalski (2015) were used to provide evidence of feedback types most commonly rendered to individual teachers in the population. Level 1 (no feedback) was most frequently rendered, resulting in 47.5% of the total feedback comments examined. Level 3 feedback comments (recount of classroom events) were ranked second most observed, resulting in 18.4% of the total feedback comments examined.

Table 10 presents an overall summary of the study, including research questions, variables, data sources, methods of analysis, and results. Chapter 5 is comprised of an elaborated summary, including discussion of the findings, implications for policy and practice, and recommendations for future research.

Table 10

Research Questions, Variables, Data Sources, Methods of Analysis & Results

#	Research Questions	Variables	Data Sources	Methods	
				of Analysis	Results
1	What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?	<p><u>Independent:</u> Instructional Practice Scores for the 2013-14 school year, ranged 1-4. 1 = highly effective 2 = effective 3 = needs improvement/developing 4 = unsatisfactory</p> <p><u>Dependent:</u> VAM scores for the 2013-14 school year, ranged from -2.28 to -0.63</p>	Value-added model (VAM) scores; Instructional practice scores	Pearson r Correlation	<p>$N = 528$</p> <p>$r(526) = -.104$ $p = 0.016$</p> <p>H_0 accepted, very weak (nearly no) relationship between VAM scores and IP scores.</p>
2	What difference, if any, exists between PSC/CC and AC/PC teachers' instructional practice scores within Central Florida school districts?	Mean instructional practice scores for tenured (PSC/CC) vs. nontenured teachers (AC/PC) with the lowest 10% of VAM scores for the 2013-14 school year.	Instructional practice scores and teacher contract statuses	Independent samples t test, Cohen's d	<p>$N = 528$</p> <p>Tenured $f = 278$</p> <p>Nontenured $f = 250$</p> <p>$t(526) = -3.306$, $p = 0.001$</p> <p>H_0 rejected, significant difference between IP scores for tenured vs. nontenured teachers was observed.</p> <p>Cohen's d effect size = -0.275; low practical importance.</p>

#	Research Questions	Variables	Data Sources	Methods	
				of Analysis	Results
3	What percentage of PSC/CC and AC/PC teachers identified within the 10% of lowest VAM scores in Central Florida school districts have been placed on improvement plans?	Instructional practice evaluation content, disaggregated by contract status. Yes/No contains improvement plan.	Instructional practice evaluations	Descriptive statistics, frequencies and percentage of teachers assigned improvement plans by contract status and total population	$N = 528$ Tenured $f = 278$ Nontenured $f = 250$ Two improvement plans discovered, one plan for tenured (0.40%), one plan for nontenured (0.36%). Total percentage of improvement plans for population = 0.38%
4	Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida school districts?	Feedback levels	Feedback content from instructional practice evaluations	Descriptive statistics, including a frequency distribution of feedback rubric ratings to illustrate findings	L1 = 47.5% L2 = 6.3% L3 = 18.4% L4 = 12.5% L5 = 2.1% L6 = 1.8% L7 = 11.6% Greatest frequency L1 (no feedback), followed by L3 (recount of classroom events).

CHAPTER 5 SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Introduction

This chapter extends upon chapters one through four by elaborating on the research study through an overall summary of the study, discussion of the findings, implications for policy and practice, and recommendations for future research. The summary of the study includes a restatement of the problem, purpose of the study, theoretical framework, research questions, and research methodology. The following section is dedicated to a discussion of the findings, grounded in the literature cited in chapter two. Finally, recommendations for further research are extended, followed by final conclusions for the entire research study.

Summary of the Study

To date, there has been limited research conducted on how Florida value-added model (VAM) student achievement data, instructional practice ratings, feedback, and improvement plans have been used to determine and remediate teacher performance. Additionally, there is little known regarding how barriers to candid and appropriate feedback, such as individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults contribute to administrative decisions to assign teachers with low VAM scores to improvement plans and subsequently pursue dismissal of such ineffective teachers.

The purpose of this study was to determine if grossly ineffective teachers, according to Florida VAM scores, had related instructional practice evaluations that were being used to determine placement on improvement plans. Additionally, this study was

conducted to determine if there was a difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teacher instructional practice scores and assignment to improvement plans for teachers within the lowest 10% of VAM scores in Central Florida school districts. Finally, the levels of feedback held within the local district performance evaluations and improvement plans were reviewed to search for evidence of administrative barriers within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults that existed and inhibited administrative evaluators from providing candid and actionable teacher performance evaluations relative to student achievement.

To understand the relationship between low student achievement measures and school leadership response to ineffective teaching, the study was grounded by the social systems theory of administrative behavior (Getzels & Guba, 1957). The basis of this theory was that there are multiple components that contribute to a behavior exhibited by an administrator, and balancing these components is necessary to achieve institutional goals (Getzels & Guba, 1957). The groundwork for the theoretical framework was the innate conflict and need for balance between the nomothetic and idiographic dimensions of a social system. It was theorized by Getzels and Guba that all administrative behavior was a function of the institutional role (nomothetic) and personality (idiographic) dimensions, where $B = f(R \times P)$. Per this equation, school administrators were theorized to be caught between individual and institutional conflicts of roles and expectations, (such as supervision, evaluation, and targeted feedback delivery toward the goal), versus

their own personalities and needs (harmony and belongingness). The administrator must achieve a balance of both role and personality to exhibit behavior conducive to achievement of the institutional goal (Getzels & Guba, 1957).

Following are the research questions and null hypotheses used to guide the study:

1. What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?

H₀₁: There is no observed relationship between the lowest 10% of teachers' VAM scores as correlated to local instructional practice scores within Central Florida school districts.

2. What difference, if any, exists between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts?

H₀₂: There is no observed difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts.

3. What percentage of professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers identified with the 10% of lowest VAM scores in Central Florida school districts have been placed on improvement plans?

4. Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida School districts?

Data for the study were collected to respond to the research questions. Value-added model scores from the Florida Department of Education (FDOE) were obtained through correspondence with the Division of Accountability, Research, and Measurement office. Instructional practice evaluations, including instructional practice scores and any feedback delivered by administrators, teacher contract statuses, and existing improvement plans were obtained via public record requests made to each of the 13 Central Florida school districts pursuant to Florida Statute §119.07. Of the 13 Central Florida school districts that held membership in the Central Florida Public School Boards Coalition, 11 school districts provided complete and viable data sets in response to the public record requests, yielding a total of 528 instructional practice teacher evaluations for data analysis as described by the research questions.

Discussion of the Findings

Research Question 1

What relationship, if any, exists between the lowest 10% of teachers' VAM scores and instructional practice evaluation scores in Central Florida school districts?

H₀₁: There is no observed relationship between the lowest 10% of teachers' VAM scores as correlated to local instructional practice scores within Central Florida school districts.

For the population of 528 teachers, value-added model (VAM) scores ranged between -2.28 and -0.63, which were correlated with instructional practice scores that

ranged from one to four (1 = highly effective, 2 = effective, 3 = needs improvement/developing, 4 = unsatisfactory). Value-added model scores displayed a mean of -0.90 and a standard deviation of 0.27. Alternatively, instructional practice scores displayed a mean of 1.65 (between highly effective and effective), and a standard deviation of 0.65. Of the 528 summative instructional practice evaluations investigated, 43.2% of teachers were rated highly effective, 51.1% were rated effective, 3.6% were rated needs improvement/developing, and 2.1% were rated as unsatisfactory. Furthermore, the Pearson r correlational analysis displayed a coefficient of -0.104, resulting in the null hypothesis being accepted ($p < .05$); there was a very weak observed relationship between the lowest 10% of teachers' VAM scores as correlated to instructional practice scores within Central Florida school districts. These data were in agreement with the findings of prior researchers as reported in the review of literature. A comparison of prior evidence as related to these findings is summarized in Table 11.

Table 11

Comparison of Prior Research as Related to Teacher Effectiveness Ratings

Researcher	Findings
Langlois & Colarusso (1988)	98.2% of teachers received a perfect evaluation score when districts used Pennsylvania State Education Department standard rating form.
Gallagher (2004)	Strong, positive, and statistically significant relationship between teacher evaluation scores and student achievement using Value-added model (VAM) teacher effects correlated with teacher evaluation scores in reading (correlation coefficient of .50, $p=.01$).
Milanowski (2004)	Combined VAM estimates and teacher evaluations scores for a sample of teachers in grades three through eight correlated at .43 for math, .32 for reading, and .27 for science.
Donaldson (2010)	Superintendent Hopkins (Maine, 2001) conducted a review of summative evaluations of all teachers upon becoming superintendent. Hopkins stated that the summative evaluations were like reading valentines; the evaluations were full of vague, meaningless praise, and overall, were absent of constructive criticism and actionable feedback.
Weisberg et al. (2009)	More than 99% of teachers received a satisfactory rating on a binary scale (satisfactory, unsatisfactory). 94% of teachers received one of the top two ratings on a multi-tiered scale.
Papay (2011)	Correlation coefficients ranged between .15 and .58 when correlating VAM scores with various standardized assessments in mathematics, reading, English-language arts, and the Stanford Achievement Test subtests for reading and mathematics.
Mela (2013)	99% of teachers in Brevard County, FL were rated effective or highly effective.
Winters & Cowen (2013)	Students assigned to teacher who would have been dismissed one or two years earlier according to a VAM-based policy of ineffective teacher dismissal (teachers at or below the fifth percentile) yielded an average 0.188 standard deviation decrease in achievement one year later, compared to students with teachers above the fifth percentile.
Pace (2015)	92.5% of teachers in Brevard County, FL were rated effective or highly effective.
Butler (2017)	94.3% of all teachers in Central Florida within the lowest 10% of VAM scores in the 2013-14 school year were rated effective or highly effective. -0.104 Pearson r correlation coefficient when investigating the relationship between VAM scores and instructional practice scores for the lowest 10% of teachers, according to VAM scores in the 2013-14 school year ($p < .05$).

The Lake Wobegon Effect is a term characterized by “a phenomenon in which most individuals or groups perform above average” (Wheeler & Haertel, 1993). In the case of summative teacher evaluations, numerous researchers have found that the Lake Wobegon Effect is prevalent. Teachers are almost always rated effective or highly effective, even in the presence of VAM-based evidence that suggests otherwise. Since Senate Bill 736 passed in 2011, the FDOE has restricted the standards for performance evaluations in an effort to ensure high-quality instruction for every student, and this has led to more rigorous standards for local instructional practice evaluations that must be approved by the FDOE. Even with VAM scores attached as a large percentage of the overall summative evaluation, a great majority of teachers have continued to be rated as effective and highly effective on the instructional practice portion of the evaluation. When nearly all teachers are rated effective and highly effective, the entire evaluation system becomes nullified. Teachers who are truly effective are not given the proper commendation, respect, and reverence for their dedicated work. In contrast, teachers who are truly ineffective are permitted to stay in the classroom with poor student growth and achievement results, and little, if any, consequence for poor teaching performance. The fact that ineffective teachers, according to Florida VAM measures, are not able to be identified by administrators, also fosters a system of evaluation that does not seek to invest in human capital by remediating teacher instructional practice and pedagogy for more favorable student growth outcomes.

Moreover, statutory changes since the passages of NCLB and RTTT have targeted the teacher evaluation process as a means to increase student achievement via a highly

qualified teacher in every classroom, to no avail (Hazi & Arrendondo Rucinski, 2009). Therefore, it is plausible that there is a complex and innate problem in administrative behavior patterns that inhibit accurate teacher performance evaluations, (e.g., the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults). Without further research to determine the root cause of the Lake Wobegon Effect, this problem encompassing teacher evaluation will be ongoing and pervasive, as stakeholders strive to increase student growth and achievement.

Research Question 2

What difference, if any, exists between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts?

H₀: There is no observed difference between professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers' instructional practice scores within Central Florida school districts.

Within the total population of 528 teachers and the lowest 10% of VAM scores, 278 (52.7%) of the teacher evaluations examined were those of tenured teachers. The remaining 250 teacher evaluations examined were those of nontenured teachers (47.3%). The findings in response to Research Question 2 indicated a statistically significant difference between the instructional practice scores of tenured and nontenured teachers in Central Florida school districts. Tenured teachers were more likely to receive a higher instructional practice score ($M = 1.56$; closer to 1, which coded for highly effective) than nontenured teachers ($M = 1.74$; closer to 2, which coded for effective). Therefore, the null hypothesis was rejected. This finding speaks to the social systems theory of administrative behavior and the hypothesized barriers to candid teacher evaluation

practices: individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults.

Because all teachers in the population were in the lowest 10% of VAM scores statewide, each of these respective teachers received negative VAM scores for the 2013-14 school year. The findings from Research Question 1 indicated that there was a low probability of any teacher within the population receiving an instructional practice score of needs improvement/developing or unsatisfactory. Therefore, these findings are significant not only statistically, but also offer significance in professional practice. Because the clear majority of teachers was rated either highly effective or effective, any difference between these two categories suggested that administrators were more likely to assign a lower instructional practice evaluation score to a nontenured teacher than to a tenured teacher, though not within the realm of the lowest two categories provided for by the state-approved evaluation systems. According to these analyses, the evaluation system in professional practice was more binary in nature, as the lowest ratings administrators were willing to assign to teachers was, on average, effective.

Although school administrators have often cited tenure as a major barrier to ineffective teacher dismissal, opponents of this view cite weakness in skill and ineffective management/supervision on behalf of school-based administrators as the root cause of this misconception (Davis, 2015; Zirkel, 2010). Such opponents have supported the philosophy that administrative behaviors and beliefs regarding tenured teacher dismissal are ill-contrived. If tenured teachers are more likely to score higher on an instructional practices portion of the Florida state-approved evaluation system due to administrator

rankings, despite being within the lowest 10% of VAM scores statewide, the problem would appear to be rooted in the actual implementation of the evaluation system by local school-based administrators, not the legal job protections prescribed by tenure status. Moreover, any dismissal of a tenured teacher under Florida Law would likely not be supported by instructional practice scores, as nearly every teacher is evaluated to be effective or highly effective on the instructional practices portion of the overall summative teacher evaluation.

Research Question 3

What percentage of professional service contract/continuing contract (tenured) and annual contract/probationary contract (nontenured) teachers identified with the lowest 10% of VAM scores in Central Florida school districts have been placed on improvement plans?

The results of further analysis related to Research Question 3, beyond the that already presented for Research Questions 2 and 3, further supported the argument that the instructional practices portion of the Florida state-approved teacher evaluation system was defunct in nature. Of the 528 teacher evaluations investigated, only two teachers within the lowest 10% of VAM scores were assigned to an improvement plan. One teacher was tenured; the other teacher was nontenured. Each of the improvement plans were from separate districts within Central Florida. The total number of teachers prescribed an improvement plan accounted for less than 1% of the total population.

Evidence of bureaucratic procedural interferences was present, specifically within one improvement plan investigated. This improvement plan contained a string of e-mail correspondence between school and district-based administrators. The teacher who

received this improvement plan had two consecutive years of needs improvement evaluations that initiated the professional improvement plan process (PIP) in this school district. The internal reporting form, submitted upon case completion, stated that the PIP was incomplete, nearly five months after its initiation. Reasoning for the incomplete PIP was due to the extensive nature of the improvement plan and administrative procedural faults associated with its implementation. The PIP required seven informal observations, and three formal observations to be conducted by the teacher's evaluator, aligned with two specific instructional practice goals: (a) noticing when students are not engaged and (b) maintaining a lively pace. Though all other informal and formal observations were completed by the administrator, the third formal observation was never conducted. Therefore, the PIP was considered incomplete, and the outcome of the entire process was nullified without resolution.

Numerous instances of bureaucratic procedural interferences and administrative procedural faults were evident throughout the e-mail correspondence. For instance, toward the PIP completion deadline, the school-based administrator sent an email to district-level personnel, stating that the third formal observation had not been conducted. When the district-level administrator asked if there was a reason for not meeting the formal requirement, the school-based administrator stated: "We did take notes--we just met and talked about the results of the most recent evaluation. We failed to do a third formal."

Furthermore, throughout the earlier e-mail correspondence, there was a lack of clarity for proper procedural implementation of the PIP by the school-based

administrator. In one instance, the draft of the PIP was due in December. Upon request by the district-based administrator to produce the draft, the school-based administrator stated that she was under the impression that the PIP was due in January. The district-based administrator corrected the school-based administrator, stating that the January implementation date was in fact the date to begin remediating the teacher, but the draft of the PIP was due in December to the district office for approval. The school-based administrator replied: “Then I messed up. Is there any way to rectify it or is the opportunity lost?” In fact, the very first e-mail correspondence from the district-level administrator, dated October 9, 2013, clearly charted the due date of December 20, 2013 for the PIP draft, with the implementation date beginning on January 15, 2014.

Thus, the qualitative evidence from one improvement plan in a Central Florida school district provided clear instances of administrative procedural faults and bureaucratic procedural interferences. Administrative procedural faults occur when an administrator lacks the skill necessary to reach the institutional goal. Bureaucratic procedural interferences occur when state and local district policies and personnel do not fully support school-based administrative actions and decisions. The competing demands encompassing the roles of school-based administrators, (i.e., time for completion of an improvement plan with numerous informal and formal observations with fidelity, and the ability to concentrate on the important process of increasing human capital by remediating instructional practice), make for a bureaucratic system confined by paperwork and deadlines governed by district-level personnel.

Research Question 4

Does the feedback reflected in the instructional practice evaluations and improvement plans target student achievement for teachers with the lowest 10% of VAM scores in Central Florida school districts?

Feedback provided within the instructional practices portion of the summative teacher evaluations for 528 teachers in Central Florida whose scores were in the lowest 10% of VAM scores for the 2013-14 school year were analyzed using a feedback rubric developed by Rafalski (2015). The feedback rubric was used to determine the frequency and quality of feedback delivered to teachers by administrative evaluators. Though 11 school districts provided full summative evaluation data for the instructional practices portion of the overall teacher evaluations, the volume of documentation for some school districts was substantially less than that of other school districts. For example, one Central Florida school district provided entire evaluations, including all observational data and final commentary feedback by evaluators. Other school districts provided one sheet of evaluative information for each teacher with numerical calculations for each evaluation component. When asked if the single sheet evaluations were the complete evaluation file, these school districts replied that the single sheets were the only data that were held by the central district offices; there were no other data to support a feedback level of greater than one (no feedback). When asked how feedback was delivered to teachers based on these simple evaluation documents, these school districts replied that evaluators were trained to hold conferences with the teachers to deliver feedback. Using the data provided, the researcher determined that the greatest frequency of feedback delivered to the population of teachers was Level 1 (no feedback) at 47.5%, followed by

Level 3 (recount of observation events) at 18.4%, and Level 4 (general affirmation or praise statement) at 12.5%. Of all the feedback observed, only 11.6% was specific targeted feedback (the highest level provided for by the rubric).

Based on these data and conversations regarding the volume and integrity of evaluation paperwork, it became clear that bureaucratic procedural interferences were a common factor within school districts. Supporting documentation for review of evaluations by district-level personnel were not available in some school districts. Should school board members, superintendent, union representative, teacher, or other stakeholders choose to examine the files (to supervise the content within teacher evaluations in these districts), they would find no data to support any employee performance decisions.

Furthermore, it was of much concern to find that the greatest feedback levels were distributed at the lower end of the feedback scale developed by Rafalski (2015), with 84.7% of the total feedback being between Levels 1 and 4. Although bureaucratic procedural interferences may account for much of the Level 1 feedback, the remaining results allude to the other two dimensions that are barriers to candid and actionable feedback delivery by evaluative personnel. Feedback Levels 1-4 include no feedback, unrelated feedback or general statements, recount of observational events, and general affirmations or praise statements. These feedback levels are quite benign in nature and are much simpler to deliver than feedback Levels 5-7 which include reflective feedback, standardized feedback, and specific targeted feedback. One could surmise, when observing the differences in these feedback levels developed by Rafalski, that there is a

clear dichotomy between the lower and upper levels of feedback in terms of how skillful and willful an evaluative administrator is in practice.

It is much less intrusive, and emotionally safer for administrators, to provide lower levels of feedback that do not include statements of instructional practice that may conflict with the beliefs teachers have that they are performing exceptionally well. Herein is the root of administrative conflict avoidance. Are administrators content to provide lower levels of feedback so as to not disturb teachers' sense of well-being and self-worth, or do they hold true to the belief that teachers are the single most important component in a student's educational growth and performance, and evaluate with deep integrity and seek to increase human capital by providing timely, candid, and actionable feedback on performance? Do administrators shy away from the higher levels of feedback because it may lead to a point of contention between teachers and administrative evaluators?

Administrative procedural faults may also account for the high distribution of lower-level feedback. Although some of the evaluations delivered by school districts were simple, single-paged documents, other school districts provided evaluation paperwork that, in some cases, exceeded 20 pages. It was common to observe evidence of administrators typing each spoken word by teachers and students, essentially script-taping the entire lesson. However, in these instances, administrators were providing low levels of feedback, merely sweeping over the lesson by recounting observational events (Level 3 feedback). Though appearing to be quite arduous and cumbersome, it seemed that evaluators were trained to provide this type of evidence for evaluations, without ever

providing any type of commentary following the recount of classroom events as to how a teacher may improve instructional practice. These administrative procedural faults are two-fold. The first fault is that of the district personnel and administrative professional development personnel who may train administrators to provide this type of feedback. The second administrative procedural fault is that of the actual evaluator who lacks the ability to recognize that this type of feedback does not allow for teachers to ascertain any true meaning from their script-taping, general praise, or unintelligible commentary.

Regardless of whether a teacher is evaluated as highly effective, effective, needs improvement/developing, or unsatisfactory, it should be considered an industry standard for teachers to receive valuable, candid, and actionable written feedback on instructional performance. All employees, regardless of how well they are teaching at any particular moment, should expect to be evaluated in an accurate and fair manner, with their unique teaching styles and characteristics highlighted. Their shortcomings should be respectfully acknowledged, and they should be advised regarding an appropriate plan for success. Teachers should expect to be delivered feedback that will further improve performance over time and develop a team mentality of administrator-teacher reciprocal accountability for increasing student growth and performance. It would be unimaginable in another profession to expect an evaluator to sit in the workplace and merely type each word that was said, delivering ratings based on such evidence. Though teaching in the public education arena is certainly not a business, e.g., where profit and shareholder value are tangible indicators of performance), student growth and performance are positive indicators of how well a teacher is educating students. If administrative evaluators are

not properly trained to provide specific, targeted feedback that respects the dignity and nobility of the teaching profession, the instructional practices portion of the overall teacher evaluation will never be the tool it was intended to be. Similarly, public education will never be the great equalizer it was envisioned to be.

Implications for Policy and Practice

There is little disagreement that teachers are the single most important force and factor in a student's educational progress. Schools require great teachers to be working with students each day to ensure that the future of each child is bright. However, it has become abundantly clear that evaluation systems, particularly in the state of Florida, do not appropriately distinguish who the truly great teachers are among the total population. Moreover, teachers who have low student growth measures, as indicated by VAM scores, are consistently being rated as effective or highly effective on the instructional practice portion of the summative evaluation with little to no feedback to render these ratings valid or reliable. The conclusion is, that nearly every teacher, regardless of student growth measures, is considered effective or highly effective. The findings support that the Florida system of teacher evaluation is not efficient and that it is not supportive of or rewarding for teachers. Regardless of whether teachers' VAM scores are high or low, the instructional practice portion of the evaluation inflates local evaluations, and this has led to far fewer needs improvement/developing or unsatisfactory summative ratings overall.

Three interrelated policy and practice implications have emerged from the results of this study. Stakeholders, such as Florida policymakers, executive-level school district

administrators, school-based administrators, teachers' unions, and educational researchers should consider the following three implications:

1. The results of this study suggested that there are clear concerns with relatedness of VAM scores to the instructional practices portion of teacher evaluations. District-based executives should be advised to have simple correlations run each year to determine if there is any relationship between VAM scores of teachers within their local school district and the current instructional practice ratings. Additionally, executives should sample evaluations within their school districts each year, and read the evaluations with a careful eye attuned to the types of ratings being earned, along with the feedback given upon issuance of a particular rating. Executives should determine which school-based evaluators are rating teachers appropriately when correlating instructional practice scores to VAM scores as well as those evaluators who are providing exemplary feedback targeted to the institutional goal of increasing student growth and achievement. Once these school-based evaluators are determined, these administrators may be asked to mentor struggling evaluators and train new administrators in evaluating teaching performance in an appropriate and valid manner.
2. It is recommended that school districts in Central Florida re-examine and re-design their improvement plan processes. The qualitative data evidence from one improvement plan and the subsequent email correspondence regarding the process provided insight as to how arduous an improvement plan may be for

school-based administrators who are encumbered with various other duties. Furthermore, the lack of improvement plans in general is cause for concern. A substantial contributor to the lack of improvement plans is that nearly every teacher is rated effective or highly effective. Therefore, there is no evidence to suggest that administrators can identify struggling teachers and develop plans to increase instructional performance. This creates a cascading effect of problems in how instructional practice evaluations are being implemented. School district executives should be advised to prepare annual reports on the total number of teachers with improvement plans so as to determine if there is a need to reduce the paperwork on behalf of the school-based administrator to ensure that the proper actions are being followed to remediate teaching performance and increase human capital.

3. Low levels of feedback (1-4) were far more frequent than higher levels of feedback (5-7) on teaching performance. One reason for the low levels of feedback identified in this study may have been due to school district offices not maintaining complete evaluation and observation data. Although maintenance of complete and accurate data files may be one segment of the overall lack of high-quality feedback, a second area of concern is related to administrators' preparation and training in delivering feedback in certain Central Florida school districts. For example, one school district's evaluations were largely script-taped recounts of observational events. It appeared that the evaluators in this district were trained to provide evidence to support a

particular rating, as opposed to feedback on how teachers could improve their instructional performance. District executives would be well-advised to speak with the professional development teams in their respective school districts to determine how school-based administrative evaluators are trained to execute the instructional practices portion of the evaluation system. Additionally, school districts should consider re-designing evaluator trainings with more emphasis on how to deliver objective, clear, candid, and actionable feedback, in addition to instructional performance evidence, to all teachers, regardless of summative performance ratings.

Recommendations for Future Research

The goal of this study was to determine if VAM scores were related to instructional practice evaluation scores and improvement plans. Data were collected and analyzed to test four research questions relative to this goal. Upon analysis and discussion of the findings, recommendations for future research as related to this study were deliberated. The findings, though meaningful for educational policy and practice in the state of Florida, were limited methodologically. The greatest limitation of this study was due to the nature of public record requests and the availability of pertinent data held at district-level offices. Of the 758 total teachers identified in the population, 230 teachers were removed from this study due to limitations previously cited (e.g., such as termination of employment prior to the conclusion of the evaluation process, retirement, and lack of supporting evaluation documentation from district central offices). Additionally, each school district had different ways of documenting their different

evaluation systems. Some included all observational and evaluation data; others included only a summative numerical score, absent of comments and feedback. Due to these differences and the likelihood of missing statistical evidence, the ability to generalize or form conclusions from this study was limited. Thus, suggestions are made for further research.

Further research into this subject should include a more in-depth study of comparison for VAM scores and instructional practice scores. To better correlate VAM scores to instructional practice scores, the population in future studies should include all teachers and not be delimited to a particular subsection of the Florida teacher populace. Furthermore, one section of the overall teacher evaluation that was not included in this study was the final one-third of the evaluation. This section may be compiled considering other indicators of performance, such as professional and job responsibilities recommended by the State Board of Education or identified by the district school board, peer reviews, objectively reliable survey information from students and parents based on teaching practices that are associated with higher student achievement, and other valid and reliable measures of instructional practice (Florida Statute §1012.34, 2015). Within the written evaluations, one teacher expressed her concern for this final one-third which included a team evaluation instrument where peers rated one another based on collaboration and mutual accountability:

The team evaluation instrument does not relate to what a teacher does in his/her classroom on a daily basis. It is truly a “performance” based evaluation dependent upon how effectively one entertains the assessor and displays the more concrete elements such as board configuration. In fact, the team evaluation is demeaning and punitive in nature. Innovative is in the eyes of the beholder-

teaching is heart deep. The academic and behavioral responses of my students (past and present) validate my teaching proficiency.

This qualitative evidence spawns more questions about the overall teacher evaluation in Florida. The statute has been written to be open to interpretation and, at the discretion of the local school boards and school districts, to decide how a teacher will be rated for the final one-third of his or her evaluation. A significant point of interest for future researchers would be to examine how different Florida school districts determine what will be used for this part of the evaluation and what effect each system has on the overall determination of an instructional practice score.

Another avenue of research could be to develop a study that includes more concrete data, displaying evidence of administrative barriers that exist and inhibit administrative evaluators from providing candid and actionable teacher performance evaluations relative to student achievement within the three dimensions of individual conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults. Though this study made inferences from various qualitative and quantitative data points on these barriers, more valid and reliable data could be determined by conducting interviews, surveys, and case studies with current administrative evaluators.

Many questions have been illuminated in this study regarding the Florida teacher evaluation system since the Student Success Act was enacted in July 2011. That nearly every teacher, despite VAM score outcomes, has been rated effective or highly effective on paper is a clear indication that the Lake Wobegon Effect is prevalent in Central Florida school districts. The question that requires further study is why this phenomenon is occurring. Researchers must investigate and determine the common factors behind this

phenomenon to provide a greater understanding of how educational policy-makers, district-level administrators, school-based administrators, and various other stakeholders may mitigate this problem of professional practice.

Summary

The findings of this study elaborated upon the work of numerous previous researchers on the topic of teacher evaluation, both in the state of Florida and in other locales. With the fairly new implementation of the value-added model (VAM) scores being used to calculate how individual teachers affect student academic growth, it has been determined that there is a lack of relationship between VAM scores as compared to instructional practice scores for individual teachers. Administrators are unable to discern which teachers in a population are truly effective or highly effective as defined by student growth and achievement. This finding was of even more interest with the discovery that tenured teachers were rated higher than nontenured teachers on the instructional practice portion of the overall summative teacher evaluation system used by Central Florida school districts. Additionally, improvement plans were not being implemented regularly to remediate teacher performance and increase human capital. Finally, nearly all feedback associated with the teacher evaluations examined was low-level and absent of feedback relative to student growth and achievement.

The literature reviewed indicated that school administrators have been viewed to be deeply conflicted between their institutional roles and expectations (supervision, evaluation, and feedback toward the goal) versus their own individual personality and needs (harmony and belongingness) (Getzels & Guba, 1957). Among the institutional

roles and expectations, paired with individual needs of an administrator, are the three dimensions of administrative conflict avoidance, bureaucratic procedural interferences, and administrative procedural faults, which have been inferred to contribute to the findings of this study. The literature and subsequent findings suggest that an evaluation system is only as impactful as the evaluators who are conducting the observations and determining the summative outcomes. To improve upon the current evaluation system, administrators must be aware of, and well-prepared for, the demands of evaluating, remediating, and providing feedback to teaching professionals relative to student achievement and growth in a manner that is simultaneously respectful, candid, fair, timely, and actionable.

APPENDIX A
UCF INSTITUTIONAL REVIEW BOARD EXEMPTION OF RESEARCH



University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

From : **UCF Institutional Review Board #1**
FWA00000351, IRB00001138
To : **Tara L. Butler**
Date : **May 09, 2016**

Dear Researcher:

On 05/09/2016 the IRB determined that the following proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56:

Type of Review: Not Human Research Determination
Project Title: AN INVESTIGATION OF LOW PERFORMING
CENTRAL FLORIDA TEACHER EVALUATION
FEEDBACK AND IMPROVEMENT PLANS AS
RELATED TO VALUE-ADDED MODEL SCORES
AND INSTRUCTIONAL PRACTICE SCORES
Investigator: Tara L. Butler
IRB ID: SBE-16-12271
Funding Agency:
Grant Title:
Research ID: N/A

University of Central Florida IRB review and approval is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are to be made and there are questions about whether these activities are research involving human subjects, please contact the IRB office to discuss the proposed changes.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

A handwritten signature in black ink that reads "Kanille Chay" followed by a horizontal line.

IRB Coordinator

APPENDIX B
SAMPLE LETTER FOR PUBLIC RECORDS REQUESTS

Date

To Whom It May Concern:

I am seeking the full summative instructional practice evaluations, including (1) total instructional practice determination (highly effective, effective, needs improvement/developing, unsatisfactory), (2) improvement plans (if applicable), (3) contract status (probationary contract, annual contract, or professional service contract), (4) instructional feedback/comments by evaluators held within the evaluations of the following teachers in accordance with Florida Statute §119.07. The format for the following list is as follows: SCHOOL NAME TEACHER NAME (Last, First, Middle Name or Initial, if applicable).

The records I am requesting are for the 2013-14 school year, and the school listed is the school in which the subsequent individual was assigned to in the 2013-14 school year.

Should you deny my request, or any part of the request, please stat in writing the basis for denial, including the exact statutory citation authorizing the denial as required by Florida Statute §119.07(1)(d).

Thank you,
Signature of Requestor
Contact Information (email or phone) of Requestor

APPENDIX C
SAMPLE RUBRIC FOR FEEDBACK DETERMINATION RATING

No Feedback Provided (Level 1)	Unrelated Feedback or General Statement Provided (Level 2)	Recount of Classroom Events (Level 3)- Justification for rating	General Affirmation Statement (Level 4)	Reflective Question (Level 5)	Standardized Feedback Provided (Level 6)	Specific Targeted Feedback Provided (Level 7)
	2-The message is Unintelligible	3-Recap has several different components (sometimes statement of percent of students being monitored or desired effect). Sometimes actually gives examples of what is wrong with no suggestion. You did this or that, teacher did this or	4-General praise. Good job, great job, excellent job, I liked, I loved, WOW!	5- Asks the teacher a question.	6- Examples: How might you adapt and create new strategies for chunking content into digestible bites that address unique student needs and situations? How might you expand your monitoring to involve more	7-Language like: 1. Reference to Resource Library or Reflective Teacher 2. Maybe try.... Or You might want to try.... 3. Consider.... 4. Recommendation.... Or I would recommend.... 5. Suggestion.... Or I suggest.... 6. It might be a good idea... 7. You should.... 8. This would have been good or great if...
		that, students did this or that, I observed this or that...			students? What are you learning about your students as you adapt and create new strategies? In addition to monitoring students by the use of choral responses, how else can you monitor students when chunking information?	9. To move to a higher level, do this_____. 10. Think about.... 11. I want you to.... 12. Next time.... 13. Always.... 14. Be sure to... or Make sure you.... 15. Doing this _____ would have been more effective. 16. You need to.... 17. Continue to.... 18. Remember this _____. 19. Do this _____. 20. This _____ is a good strategy. 21. I would like to see.... 22. Coaching idea...

Note. Rubric used in data analysis with permission of Rafalski (2015).

LIST OF REFERENCES

- Blacklock, K. (2002). Dealing with an incompetent teacher. *Principal*, 81(4), 26-28.
- Blase, J., & Blase, J. (2000). Effective instructional leadership: teachers' perspectives on how principals promote teaching and learning in schools. *Journal of Educational Administration*, 38(2), 130-141.
- Bridges, E.M. (1993). *The Incompetent teacher: Managerial responses. A revised and extended edition. The Stanford Series on Public Policy*. Bristol, PA: Falmer Press.
- Bridges, E.M. (1985). It's time to get tough with the turkeys. *Principal*, 64(3) 19-21.
- Brieschke, P.A. (1986). The administrative role in teacher competency. *Urban Review*, 18(4), 237-251.
- Cambron-McCabe, N.H., McCarthy M., and Thomas, S. (2004). *Public school law: teachers' and students' rights* (5th ed.). Boston, MA: Pearson
- Cooper, B.S., Ehrensall, P.A., Bromme, M. (2005). School-level politics and professional development: Traps in evaluating the quality of practicing teachers. *Educational Policy*, 19(1), 112-125.
- Daley, G., & Kim, L.. (2010). *A teacher evaluation system that works. Working paper*. Santa Monica, CA: National Institute for Excellence in Teaching.
- Darling-Hammond, L. (1996). What matters most: A competent teacher for every child. *Phi Delta Kappa*, 78(3), 193-200.
- Davis, L. (2015). Can Vergara fix our schools? *California Lawyer*, 35(8), 14-21.
- DeNisco, A. (2014). Removing weak links: New laws and policies empower administrators to dismiss ineffective teachers. *District Administration*, 2, 50-54.
- Donaldson, M.L. (2009). So long, Lake Wobegon? Using teacher evaluation to raise teacher quality. *Center for American Progress*. Retrieved from: https://www.americanprogress.org/wp-content/uploads/issues/2009/06/pdf/teacher_evaluation_execsumm.pdf
- Donaldson, M.L. (2010). No more valentines. *Educational Leadership*, 67(8), 54-58.
- Doran, H.C., & Izumi, L.T. (2004). *Putting evaluation to the test: A value-added model for California*. San Francisco, CA: Pacific Research Institute.

- Elmore, R. F. (2002). Bridging the gap between standards and achievement: The imperative for professional development in education. *Albert Shanker Institute*. Retrieved from: <http://www.shankerinstitute.org/resource/bridging-gap-between-standards-and-achievement>
- Florida Department of Education (2015). *Florida's value-added model (VAM) frequently asked questions*. Retrieved from: <http://www.fldoe.org/teaching/performance-evaluation>
- Florida State Senate Bill 736. (2011).
- Florida Statute §265.605. (2016).
- Florida Statute §267.17. (2016).
- Florida Statute §119.07. (2016).
- Florida Statute §119.071. (2016).
- Florida Statute §1012.01. (2015).
- Florida Statute §1012.33. (2016).
- Florida Statute §1012.335. (2015).
- Florida Statute §1012.34. (2015).
- Fuhr, D. L. (1993). Managing mediocrity in the classroom. *School Administrator*, 50(4), 26-29.
- Futernick, K. (2010). Incompetent teachers or dysfunctional systems? *Phi Delta Kappa*, 92(2), 59-64.
- Gallagher, H.A. (2004). Vaughn Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4), 79-107.
- Getzels, J.W., & Guba, E.G. (1957). Social behavior and the administrative process. *The School Review*, 65(4), 423-441.
- Hain, J.H., & Smith, G.J. (1966). Evaluation of teachers- the principal's dilemma. Retrieved from: <https://eric.ed.gov/?id=ED013490>

- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, 51(1), 73-112.
- Hattie, J., & Timperley, H. (2007). *The power of feedback*. Thousand Oaks, CA: SAGE.
- Hazi, H.M., & Arrendono Rucinski, D. (2009). Teacher evaluation as a policy target for improved student learning: A fifty state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, 17(5), 1-19.
- Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 1 101-136.
- Langlois, D.E., & Colarusso M.R. (1988). Don't let teacher evaluation become an empty ritual. *Executive Educator*, 10(5), 32-33.
- Mela, C.A. (2013). *An analysis of the School Board of Brevard County Instructional Performance Appraisal Instrument and the validity of its components*. Orlando, FL: University of Central Florida.
- Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33-53.
- Mock, S.G., & Melnick, S.A. (1991). The Marginal Teacher: A Survey of Principals' Attitudes. Retrieved from: <https://eric.ed.gov/?id=ED342739>
- National Commission on Teaching and America's Future (1996). *What matters most: Teaching for America's future*. Report of the National Commission on Teaching & America's Future. Woodbridge, VA: National Commission on Teaching and America's Future.
- National Council on Teacher Quality. (2010). Bumping HR: Giving principals more say over staffing. *National Council on Teacher Quality*, 1-13.
- Nettles, S.M., & Herrington, C. (2007). Revisiting the importance of direct effects of school leadership on student achievement: The implications for school improvement policy. *Peabody Journal of Education*, 82(4), 724-736.
- Nixon, A., Packard, A.L., & Dam, M. (2011). Teacher contract non-renewal: southeast and Midwest principals. *International Journal of Educational Leadership Preparation*, 6(4), 4-21.

- Ovando, M. N., & Ramirez, A., Jr. (2007). Principals' instructional leadership within a teacher performance appraisal system: Enhancing students' academic success. *Journal of Personnel Evaluation in Education*, 20(1-2), 85-110.
- Pace, D.P. (2015). *A study of the impact of Brevard County School's Peer Coaching Model on student achievement outcomes and teacher evaluation results*. (Unpublished doctoral dissertation). Orlando, FL: University of Central Florida.
- Papay, J.P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, 48(1), 163-193.
- Phillips, W., & Young, B. (1997). Just caring? Supervisors talk about working with incompetent teachers. *Journal of Educational Thought*, 31(2), 105-121.
- Platt, A.D., & Tripp, C.E. (2008). Communities that undermine learning: school leaders must distinguish between PLCs that genuinely serve greater student learning and groups that protect mediocre performance by both students and adults. *Leadership*, 1, 18-32.
- Pratt, F. (1996). The ethical principal and teacher nonretention. *Journal of Personnel Evaluation in Education*, 101(1), 29-36.
- Rafalski, S. (2015). *Policy implications of a teacher evaluation system: The relationship of classroom observations, levels of feedback, and student achievement outcomes*. (Unpublished doctoral dissertation). University of Central Florida, Orlando, FL.
- Randklev, B., & Lemon, D.K. (1990). When tenured teachers fail. *Principal*, 70(2), 44-45.
- Range, B. G., Duncan, H. E., Scherz, S. D., & Haines, C. A. (2012). School leaders' perceptions about incompetent teachers: Implications for supervision and evaluation. *NASSP Bulletin*, 96(4), 302-322.
- Robertson, M. (2015). Blaming teacher tenure is not the answer. *Journal of Law & Education*, 44(3), 463-471.
- Staples, S.R. (1991). *A study of tenured teacher dismissals in Virginia, 1987-1990*. Doctoral Dissertation, Virginia Polytechnic Institute and State University, 1990. *Dissertation Abstracts International*.
- Steinberg, M.P., & Donaldson, M.L. (2014). *The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era*. Policy brief. Retrieved from: <http://cepa.uconn.edu/home/research/>

- Stronge, J.H. (1995). Balancing individual and institutional goals in educational personnel evaluation: a conceptual framework. *Studies in Educational Evaluation*, 21, 131-151.
- Stronge, J.H., Ward, T.J., & Grant, L.W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, 4, 339.
- Torff, B., & Sessions, D. (2009). Principals' perceptions of the causes of teacher ineffectiveness in different secondary subjects. *Teacher Education Quarterly*, 36(3), 127-148.
- Tucker, P.D. (1997). Lake Wobegon: Where all teachers are competent (or, have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education*, 11, 103-126.
- Tucker, P. (2001). Helping struggling teachers. *Educational Leadership*, 5, 52-55.
- Vergara v. State of California, BC484642 (2014).
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., & New, T.P. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Executive summary* (2nd ed.) Brooklyn, NY: New Teacher Project.
- Wheeler, P., & Haertel, G.D. (1993). *Resource handbook on performance assessment and measurement: A tool for students, practitioners, and policymakers* Retrieved from: <https://login.ezproxy.net.ucf.edu/login?auth=shibb&url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED367686&site=eds-live&scope=site>
- Wiggins, G. (2012). 7 Keys to effective feedback: Advice, evaluation, grades- none of these provide the descriptive information that students need to reach their goals. What is true feedback- and how can it improve learning? *Educational Leadership*, 70(5), 11-16.
- Winters, M. & Cowen, J. (2013). Who would stay, who would be dismissed? An empirical consideration of value-added teacher retention policies. *Educational Researcher*, 42(6), 330-37.
- Yariv, E. & Coleman, M. (2005). Managing “challenging” teachers. *International Journal of Educational Management*, 19(4), 330-346.
- Zirkel, P.A. (2010). Teacher tenure is not the real problem. *The Phi Delta Kappan*, 1, 76-77.