
Electronic Theses and Dissertations, 2004-2019

2015

Central Florida Educational Leaders' Professional Perceptions of Race to the Top Components Concerning Teacher Evaluation and Compensation

Orin Smith
University of Central Florida



Part of the [Education 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, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Smith, Orin, "Central Florida Educational Leaders' Professional Perceptions of Race to the Top Components Concerning Teacher Evaluation and Compensation" (2015). *Electronic Theses and Dissertations, 2004-2019*. 1181.

<https://stars.library.ucf.edu/etd/1181>



Showcase of Text, Archives, Research & Scholarship

CENTRAL FLORIDA EDUCATIONAL LEADERS' PROFESSIONAL
PERCEPTIONS OF RACE TO THE TOP COMPONENTS CONCERNING TEACHER
EVALUATION AND COMPENSATION

by

ORIN SMITH

B.S. University of North Carolina, Wilmington, 2005

M.S. University of Central Florida, 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
2015

Major Professor: Rosemarye Taylor

© 2015 Orin Smith

ABSTRACT

This mixed-methods replication study was conducted to develop further understanding of the professional perceptions of educational leaders as to the fairness and impact of Race to the Top reforms concerning teacher evaluation and compensation on student achievement and growth. Graduate students in education and educational leadership from a target university were selected to complete an electronic survey to collect quantitative and qualitative data for analysis.

Quantitative results from the electronic survey revealed limited diversity in professional perceptions of the five identified components of RTTT based upon professional classification or percentage free and reduced lunch population at the school sites where assigned. Among the identified RTTT components, the component that provided for the use of school- or team-level VAM scores as part of the evaluation and compensation system was consistently viewed as the least fair and least impactful by respondents.

Analysis of the qualitative data revealed a number of themes that effected respondents' professional perceptions of the RTTT initiative. The use of a value-added model in RTTT reforms, the variables considered by the model, and communication and implementation problems associated with the reforms were the central areas of concern among survey respondents.

This study provided follow-up data to Windish's 2012 study and showed a negative general trajectory of the professional perceptions of educational leaders related to this high-profile, national educational reform effort.

More than anyone else, this work exists because of my mother.

I believe we are given many gifts in our lives; the most significant, and impactful of which, is the love and support of a dedicated and persistent mother who places the well-being of her children far in advance of her own. Considering that, I know I am blessed beyond measure.

Thank you for everything you have done for me, mom. I would not be who I am today without you.

ACKNOWLEDGEMENTS

Thank you to my Dissertation Chair, Dr. Rosemarye Taylor, for allowing me the opportunity to conduct this research. I attribute a great deal of my success in this endeavor to her unwavering support and exacting standards of excellence. To the rest of my committee, Dr. Barbara Murray, Dr. Walter Doherty, and Dr. Thomas Vitale, I am honored by the opportunity to work with such an accomplished and well-regarded group of educational leaders. Thank you for your guidance and insight that helped raise the quality of this study.

Thank you to the members of the 2012 Educational Leadership Cohort. From the first day of our program I was humbled not only by the breadth and depth of your experience, but also by how warmly you welcomed me, a relative novice in the field of education, into your community of professionals. I learned a great deal from each of you and am a vastly better leader as a result.

I would also like to thank Dr. Foard Jones, Dr. Dean Cleavenger, and Kay Wolf for their early belief in my ability to be successful in this pursuit. And finally, to Dr. Daniel Windish, I am grateful for the opportunity to follow-up on your original research. Thank you for creating such a high-quality framework upon which to develop this study.

TABLE OF CONTENTS

LIST OF TABLES	ix
CHAPTER 1 THE PROBLEM AND ITS CLARIFYING COMPONENTS.....	1
Introduction.....	1
Purpose of the Study	2
Statement of the Problem.....	3
Definition of Terms.....	4
Conceptual Framework.....	10
Reform Initiatives	10
Teacher Compensation and Evaluation	10
Learning and Poverty	12
Research Questions.....	14
Methodology.....	15
Population and Sample	18
Instrumentation	18
Data Collection	19
Data Analysis	19
Limitations	20
Significance of the Study.....	20
Summary.....	21
CHAPTER 2 REVIEW OF LITERATURE.....	22
Introduction.....	22
Improving Student Achievement	23
Learning and Poverty	24
Teacher Quality.....	29
Major Reforms to Public Education	33
No Child Left Behind.....	34
Race to the Top	36
Florida's Race to the Top.....	38
Teacher Evaluation and Compensation	41
Evolution of Teacher Evaluation	41
Purposes and Perceptions of Teacher Evaluation	42
Evolution of Teacher Compensation	44
Purposes and Perceptions of Teacher Compensation	46
Reforms to Teacher Evaluation and Compensation.....	48
Value-Added Models	49
Observations	56
Summary.....	62
CHAPTER 3 METHODOLOGY	64
Introduction.....	64
Population	65

Sample.....	66
Instrumentation	67
Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components of Race to the Top	67
Part A: Demographic Information	68
Part B: RTTT Background Information.....	69
Part C: RTTT and Student Achievement/Growth.....	71
Part D: Open-ended Responses.....	72
Data Collection Procedures.....	73
Collection of Quantitative and Qualitative Data.....	73
Data Analysis	74
Analysis of Quantitative Data.....	74
Research Question 1	75
Research Question 2	75
Research Question 3	75
Research Question 4	76
Analysis of Qualitative Data.....	76
Summary.....	77
 CHAPTER 4 RESULTS	 78
Introduction.....	78
Descriptive Statistics.....	78
Categorical Variables.....	79
Continuous Variables.....	79
Data Analysis for Research Question 1	85
Quantitative Analysis for Research Question 1	85
Qualitative Analysis for Research Question 1	86
Use of Value-Added Models	86
Implementation and Communication Issues.....	88
Data Analysis for Research Question 2	89
Quantitative Analysis for Research Question 2	89
Qualitative Analysis for Research Question 2	92
Data Analysis for Research Question 3	95
Quantitative Analysis for Research Question 3	95
Qualitative Analysis for Research Question 3	97
Effects of Poverty on Students.....	97
Effects of Student Poverty on Teachers.....	98
Effects of Socio-Economic Status on VAM Scores	99
Data Analysis for Research Question 4	100
Quantitative Analysis for Research Question 4.....	100
Qualitative Analysis for Research Question 4.....	101
Additional Analysis	101
Summary.....	109
 CHAPTER 5 SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS	 110

Introduction.....	110
Purpose of the Study	110
Population, Research Design, and Instrumentation	111
Summary and Discussion of the Findings	112
Research Question 1	112
Research Question 2	114
Research Question 3	116
Research Question 4	118
Implications for Policy and Practice	119
Recommendations for Future Research	121
Limitations of this study	122
Conclusion	122

APPENDIX A SURVEY OF THE FAIRNESS AND IMPACT OF TEACHER
EVALUATION AND COMPENSATION COMPONENTS OF RACE TO THE TOP 124

APPENDIX B WRITTEN CONSENT TO EDIT AND REPLICATE SURVEY AND
FOLLOW-UP PROTOCOL

APPENDIX C INSTITUTIONAL REVIEW BOARD APPROVAL

APPENDIX D INFORMED CONSENT.....

REFERENCES

LIST OF TABLES

Table 1 Research Questions, Variables, Data Sources, & Statistical Tests	17
Table 2 Summary of Literature Reviewed: Improving Student Achievement	32
Table 3 Summary of Literature Reviewed: Major Reforms to Public Education.....	40
Table 4 Summary of Literature Reviewed: Teacher Evaluation	59
Table 5 Summary of Literature Reviewed: Teacher Compensation.....	60
Table 6 Summary of Literature Reviewed: Reforms to Teacher Evaluation and Compensation	61
Table 7 Research Questions, Independent Variables, and Dependent Variables	80
Table 8 Descriptive Statistics for Selected RTTT Components: Impact on Student Achievement	82
Table 9 Descriptive Statistics: Change in Perceptions of RTTT from 2011 to the Date of this Study, by Professional Classification.....	84
Table 10 Perceived Impact of Components of RTTT on Student Achievement, by Professional Classification.....	91
Table 11 Perceived Impact of Components of RTTT on Student Achievement, by Respondents' FRL Group (0-24, 25-49, 50-74, 75-100)	96
Table 12 Change in Perceptions of the Impact of RTTT, by Professional Classification	100
Table 13 Perceived Impact of Components of RTTT on Student Achievement, by Doctoral Degree Program	103
Table 14 Independent Samples T-Test: Perceived Impact of RTTT Components, by Doctoral Degree Program	104
Table 15 Descriptive Statistics: 2012-2014 Comparison of RTTT Components' Mean Ratings of Impact on Student Achievement	106
Table 16 Descriptive Statistics: Self-reported Knowledge of RTTT for School District- based Administrators and Instructional Personnel.....	108
Table 17 Independent Samples T-Test: Self-reported Knowledge of RTTT, by Professional Classification (School District-based Administrators and Instructional Personnel)	108

CHAPTER 1 THE PROBLEM AND ITS CLARIFYING COMPONENTS

Introduction

The U.S. Department of Education describes the spirit of its Race to the Top (RTTT) initiative as follows:

Awards in Race to the Top will go to States that are leading the way with ambitious yet achievable plans for implementing coherent, compelling, and comprehensive education reform. Race to the Top winners will help trail-blaze effective reforms and provide examples for States and local school districts throughout the country to follow as they too are hard at work on reforms that can transform our schools for decades to come. (2014)

Since its implementation in 2011, RTTT funds have been distributed in 18 states and the District of Columbia (U.S. Department of Education, 2014). As of June 30, 2013, the state of Florida had received more than \$283 million in grant expenditures from the program (Florida: State-reported ARP: SY 2012-2013, 2014). States that were awarded RTTT funds were required to “design and implement new performance evaluation systems for teachers and to utilize the evaluations to determine compensation, promotion and retention of teachers” (Windish, 2012, p. 11).

As a RTTT funded state, Florida’s local education agencies (LEAs) began implementing revised teacher evaluation systems in 2011 to comply with RTTT requirements and newly adopted state statutes. The revised teacher evaluation system tied 50% of teacher evaluations to student growth, using a value-added model approved by the state commissioner of education in June of that year (Florida: State-reported ARP: SY 2010-2011, 2014). Continuing on through the 2012-2013 school year, Florida LEAs conducted extensive professional development in the areas of Deliberate Practice, Inter-

rater Reliability, and Monitoring and Implementation (Florida: State-reported ARP: SY 2012-2013) to build capacity for, and support adoption of, the new educator evaluation system. During this time, revised principal evaluation systems were also implemented. New principal evaluations linked student growth (as measured by student scores on standardized assessments) and leadership practice (as measured by an instrument such as Marzano's School Leadership Evaluation and Deliberate Practice (School leader evaluation model, 2012)) to a principal's annual level of performance (i.e.: Highly Effective, Effective, Needs Improvement, or Unsatisfactory) (School leader evaluation model, 2012).

In 2012, prior to the implementation of RTTT requirements by the state of Florida and Florida LEAs, Windish (2012) completed research on the professional opinions of educational leaders regarding RTTT teacher evaluation and compensation reforms. The Windish study sought to analyze the perceived potential impact of RTTT reforms on student growth, particularly for economically disadvantaged students (Windish, 2012). Since the publication of that study, RTTT requirements were implemented and long-standing systems, such as teacher tenure, reworked or overturned entirely. Considering the extent of RTTT reforms, it was prudent to investigate how, if at all, professional perceptions of RTTT requirements and their impact on student growth changed from prior to implementation in 2011, to the date of the current study.

Purpose of the Study

The purpose of this replication study was to explore the professional perceptions of educational leaders (administrative and instructional personnel) regarding selected

components of the RTTT initiative, including teacher evaluation and compensation, and the perceived effect the elements have on student achievement. Electronic surveys were distributed to central Florida educational leaders pursuing advanced degrees in education and educational leadership. Surveying educational leaders, who work in different types of school settings and educational organizations, helped to understand their perceptions of the different components of RTTT. Through this study, this researcher hoped to add to the body of knowledge concerning the types of policy reforms related to teacher evaluation and compensation that facilitate improved student learning in public schools, with particular emphasis on learning outcomes for economically disadvantaged student populations.

Statement of the Problem

At the time of this study, insufficient research had been conducted to understand the perceptions of educational leaders related to improving learning outcomes for economically disadvantaged students through requirements of RTTT. Since its passage, RTTT has created new and complex processes through which educational leaders evaluate and reward effectiveness in public school classrooms. Since the original study, (Windish, 2012), upon which this research is based, full implementation of RTTT components have been achieved in states across the country. With the added clarity that seeing a system in a live setting can provide, it was an important step in research to ask similar populations to the original study (educational leaders both administrative and instructional) how they perceived the reforms to be affecting the quality of instruction and student learning outcomes. Such research could provide much needed insight into

the perceived impact and limitations of policy priorities, such as the policies adopted through RTTT, in an attempt to inform educational policy in the future.

Definition of Terms

The following terms and/or phrases were defined for the purposes of this study as follows:

Administrative: Non-instructional positions. Both site-based and district-based positions are included (i.e., principal, assistant principal, executive director, superintendent).

Compensation: Salary and benefits paid by an employer to an employee in return for the completion of specified tasks, duties, and responsibilities related to the employee's role within an organization.

Common set of K-12 standards: "A set of content standards that define what students must know and be able to do and that are substantially identical across all States in a consortium. A State may supplement the common standards with additional standards, provided that the additional standards do not exceed 15 percent of the State's total standards for that content area" (Florida Department of Education, n.d.).

Economically disadvantaged students: "All students eligible for free or reduced lunch prior to testing are considered to be economically disadvantaged" (Florida Department of Education, as cited by Windish, 2012).

Evaluation: "An assessment of an individual's performance over a period of time based on evidence from multiple measures that reflect the performance level of the

individual's work on student learning, practice, and job responsibilities" (Florida Department of Education, 2012).

Effective principal: "A principal whose students, overall and for each subgroup, achieve acceptable rates (*e.g.*, at least one grade level in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple measures, provided that principal effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, high school graduation rates and college enrollment rates, as well as evidence of providing supportive teaching and learning conditions, strong instructional leadership, and positive family and community engagement" (Florida Department of Education, n.d.).

Effective teacher: "A teacher whose students achieve acceptable rates (*e.g.*, at least one grade level in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, multiple observation-based assessments of teacher performance" (Florida Department of Education, n.d.).

Formative assessment: "Assessment questions, tools, and processes that are embedded in instruction and are used by teachers and students to provide timely feedback for purposes of adjusting instruction to improve learning" (Florida Department of Education, n.d.).

Highly effective principal: "A principal whose students, overall and for each subgroup, achieve high rates (*e.g.*, one and one-half grade levels in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple

measures, provided that principal effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, high school graduation rates; college enrollment rates; evidence of providing supportive teaching and learning conditions, strong instructional leadership, and positive family and community engagement; or evidence of attracting, developing, and retaining high numbers of effective teachers” (Florida Department of Education, n.d.).

Highly effective teacher: “A teacher whose students achieve high rates (*e.g.*, one and one-half grade levels in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, multiple observation-based assessments of teacher performance or evidence of leadership roles (which may include mentoring or leading professional learning communities) that increase the effectiveness of other teachers in the school or LEA” (Florida Department of Education, n.d.).

High-need students: “Students at risk of educational failure or otherwise in need of special assistance and support, such as students who are living in poverty, who attend high-minority schools (as defined in this notice), who are far below grade level, who have left school before receiving a regular high school diploma, who are at risk of not graduating with a diploma on time, who are homeless, who are in foster care, who have been incarcerated, who have disabilities, or who are English language learners” (Florida Department of Education, n.d.).

High-poverty school: “Consistent with section 1111(h)(1)(C)(viii) of the ESEA, a school in the highest quartile of schools in the State with respect to poverty level, using a

measure of poverty determined by the State” (Florida Department of Education, n.d.). For the purposes of this study, a school will be considered high-poverty when 75% or more of their students qualify as economically disadvantaged.

High-quality assessment: “An assessment designed to measure a student’s knowledge, understanding of, and ability to apply, critical concepts through the use of a variety of item types and formats (*e.g.*, open-ended responses, performance-based tasks). Such assessments should enable measurement of student achievement (as defined in this notice) and student growth (as defined in this notice); be of high technical quality (*e.g.*, be valid, reliable, fair, and aligned to standards); incorporate technology where appropriate; include the assessment of students with disabilities and English language learners; and to the extent feasible, use universal design principles (as defined in section 3 of the Assistive Technology Act of 1998, as amended, 29 U.S.C. 3002) in development and administration” (Florida Department of Education, n.d.).

Instructional: Any teacher, academic coach, resource teacher, or other non-administrative position (Windish, 2012).

Instructional improvement systems: “Technology-based tools and other strategies that provide teachers, principals, and administrators with meaningful support and actionable data to systemically manage continuous instructional improvement, including such activities as: instructional planning; gathering information (*e.g.*, through formative assessments (as defined in this notice), interim assessments (as defined in this notice), summative assessments, and looking at student work and other student data); analyzing information with the support of rapid-time (as defined in this notice) reporting; using this information to inform decisions on appropriate next instructional steps; and evaluating

the effectiveness of the actions taken. Such systems promote collaborative problem-solving and action planning; they may also integrate instructional data with student-level data such as attendance, discipline, grades, credit accumulation, and student survey results to provide early warning indicators of a student's risk of educational failure" (Florida Department of Education, n.d.).

Performance level: "The summative ratings of performance over the evaluation period based on accumulated evidence of proficiency in each of the criteria of the evaluation system. There are four performance levels: highly effective; effective; needs improvement, or, for teachers in the first three years of employment, developing; and unsatisfactory" (Florida Department of Education, 2012).

Persistently lowest-achieving schools: "As determined by the State: (i) Any Title I school in improvement, corrective action, or restructuring that (a) Is among the lowest-achieving five percent of Title I schools in improvement, corrective action, or restructuring or the lowest-achieving five Title I schools in improvement, corrective action, or restructuring in the State, whichever number of schools is greater; or (b) Is a high school that has had a graduation rate as defined in 34 CFR 200.19(b) that is less than 60 percent over a number of years; and (ii) Any secondary school that is eligible for, but does not receive, Title I funds that (a) Is among the lowest-achieving five percent of secondary schools or the lowest-achieving five secondary schools in the State that are eligible for, but do not receive, Title I funds, whichever number of schools is greater; or (b) Is a high school that has had a graduation rate as defined in 34 CFR 200.19(b) that is less than 60 percent over a number of years. To identify the lowest-achieving schools, a State must take into account both (i) The academic achievement of the "all students"

group in a school in terms of proficiency on the State’s assessments under section 1111(b)(3) of the ESEA in reading/language arts and mathematics combined; and (ii) The school’s lack of progress on those assessments over a number of years in the “all students” group” (Florida Department of Education, n.d.).

Professional classification: Participants self-reported their employment position as either administrative or instructional (Windish, 2012).

Student achievement: (a) For tested grades and subjects: (1) a student’s score on the State’s assessments under the ESEA; and, as appropriate, (2) other measures of student learning, such as those described in paragraph (b) of this definition, provided they are rigorous and comparable across classrooms. (b) For non-tested grades and subjects: alternative measures of student learning and performance such as student scores on pre-tests and end-of-course tests; student performance on English language proficiency assessments; and other measures of student achievement that are rigorous and comparable across classrooms” (Florida Department of Education, n.d.).

Student growth: “The change in student achievement (as defined in this notice) for an individual student between two or more points in time. A State may also include other measures that are rigorous and comparable across classrooms” (Florida Department of Education, n.d.).

Student proficiency: Based on a score of 3 or higher on the Florida Comprehensive Assessment Test (FCAT) in the areas of reading or mathematics (Windish, 2012).

Conceptual Framework

Reform Initiatives

Over the past 50 years, education reform has not been lacking in bellwether moments. From the early space race dominance by the U.S.S.R. and Sputnik, which precipitated a national emphasis on math and science, to the 1983 publication of *A Nation at Risk: The Imperative for Educational Reform* (National Commission on Excellence in Education, 1983), which ignited a national conversation surrounding the quality of the U.S. public education system. In recent years, particularly since passage of the No Child Left Behind Act (NCLB) (2002), the reform movement has honed in on the practice of education reform through systemic accountability, rigorous standards, and market-based incentives for schools and educators (Jones, 2013; Ladd, 2012).

The latest iteration of national school reform, Race to the Top (RTTT), was passed as a part of the American Reinvestment and Recovery Act (ARRA) in 2009. This initiative prompted fundamental changes to state and local education systems in exchange for significant federal grant dollars to fund them. Among the suggested changes was the adoption of a performance-pay compensation plan for teachers that linked student growth to teacher “compensation, promotion and retention” (Windish, 2012, p. 11).

Teacher Compensation and Evaluation

What was historically based upon a single salary schedule, with moderators for level of education attained and years of experience (Jones, 2013), RTTT required a new, more complex tool for compensating educators, in the form of performance pay. To facilitate a performance pay system, state and local educational agencies were required to

design teacher evaluation tools that incorporated a mix of student achievement on standardized tests, administrative observations, and one additional metric of the school districts' choosing (Florida Department of Education, 2011).

Given the controversy over this requirement, opponents of the plan alleged that performance pay systems presented inconclusive evidence of their effectiveness at raising student achievement (Jones, 2013). However, other researchers contend that performance pay systems can be successful but only when utilized as part of a multi-pronged approach to evaluating teacher effectiveness. Burnett, Cushing, and Bivona (2012) summarize the schema as follows:

Because no single measure adequately captures the complexity of teaching, evaluation systems should include multiple measures of teacher effectiveness. Additionally, the mix of measures should align to the evaluation's purpose. A tight fit between measures and purposes can result in a more comprehensive and fair performance-based evaluation system that leads to greater buy-in among teachers, principals, and other stakeholders. (p. 15)

In 2007, the state of Florida became the first state in the nation to enact a state-wide performance pay plan. "The Merit Award Program (MAP) plan [required] at least 60% of a teacher's bonus be based on student performance, and the award must be distributed to individual teachers or teaching teams" (Jones, 2013, p. 150). According to Buddin et al. (2007), this offered researchers an opportunity to evaluate the effectiveness of such a plan in a real-world environment, looking at "teacher performance, recruitment, retention, and distribution across different types of schools and districts" (p. 22).

Research by Jones (2013) returned interesting findings as to the effect performance pay plans have on teachers. In the study, Jones (2013) examined teacher behavior under performance pay plans and provided specific data from Florida that showed performance

pay plans, with individual-level incentives (the type adopted by Florida), appeared to “increase teacher effort and turnover” (p. 149). This was contrary to his findings in other states where performance pay plans distributed bonuses at the school level (Jones, 2013). In states utilizing a school-level incentive structure, Jones (2013) found that teachers decreased the number of hours worked per week (a 12% decrease) while still increasing their time spent searching for new job opportunities outside of the teaching profession (Jones, 2013, p. 149). For performance pay plans to increase teacher effort at improving their instruction, plans should be designed to compensate individual-level performance, similar to what is done in the state of Florida.

Learning and Poverty

Much research has been completed on the devastating impact poverty has on the long-term development of children (Sirin, 2005; Jensen, 2009; Ladd, 2012). The most pressing impacts of childhood poverty can be categorized as emotional and social challenges (i.e., depression, emotional dysregulation, impatience and impulsivity, and inappropriate emotional responses), acute and chronic stressors (i.e., exposure to violence and abuse, separation or divorce, and material deprivation), cognitive lags (i.e., decreased language, working memory, rewards processing, and visual and special cognition), and health and safety issues (i.e., malnutrition, environmental hazards, and insufficient health care) (Jensen, 2009). However, negative side-effects of poverty are not permanent. As Jensen (2009) stated: “A brain that is susceptible to adverse environmental effects is equally susceptible to positive, enriching effects” (p. 45).

Evidence indicates that the disparity between high- and low-income students has almost doubled in the last 50 years (Ladd, 2012). That being considered, policy implications for economically disadvantaged students are of vital importance to researchers, policy makers, educators and educational leaders, and community members, as they will impact a greater and greater number of students every year.

Approximately 22% of all children in the U.S. live below the federal poverty line (FPL) of \$23,850 for a family of four (United States Census Bureau, 2014).

Unfortunately, this measure does not fully capture the population of children who are threatened by poverty. The National Center for Children in Poverty (NCCP) estimates that up to 45% of children qualify as low-income (living within 200% of the FPL) (2014). The challenges that low-SES children face are acute and have implications for their long-term health and contribution to society. NCCP describe the effects of poverty on children as,

Poverty can impede children's ability to learn and contribute to social, emotional, and behavioral problems. Poverty also can contribute to poor health and mental health. Risks are greatest for children who experience poverty when they are young and/or experience deep and persistent poverty. Research is clear that poverty is the single greatest threat to children's well-being. (National Center for Children in Poverty, 2014, para. 2)

While societal changes will need to be made to address the causes of childhood poverty, educational leaders must be particularly mindful of the implications their policies have on the academic achievement of this population of students.

Closing the achievement gap between high- and low-poverty students and schools is possible with the right mix of policy and effort. According to Darling-Hammond (2010), "The achievement gap would be much reduced if low-income minority students

were routinely assigned highly qualified teachers rather than those they most often encounter” (p. 44). Teacher quality matters when considering how to improve student achievement. When assessing research on teacher effectiveness, Windish (2012) concluded, “as teacher effectiveness increases, the lowest achieving students gain the most” (pp. 25-26). As such, the need to adopt high-quality systems to evaluate and provide meaningful and timely feedback, that assists educators in improving their practice, is critical to advancing the educational interests of public school students across the U.S.

Research Questions

The research questions used to guide this study are enumerated below. Research questions 1-3 closely aligned with those used by Windish (2012) and allowed for a direct comparison of results between his research and the current study. Research question four was utilized to assist this researcher in validating the analysis that compared the Windish (2012) findings to those of the current study. The research questions used in this research are as follows:

1. To what extent, if any, is there a relationship between administrative and instructional personnel’s self-reported knowledge of RTTT and the perceived fairness of RTTT requirements concerning teacher evaluation and compensation?
2. To what extent, if any, is there a difference between administrative and instructional personnel’s perceptions of the impact of RTTT teacher evaluation and compensation components on student achievement/growth?

3. To what extent, if any, is there a difference in the perceptions of administrative and instructional personnel who have different self-reported school poverty percentages about the impact of RTTT teacher evaluation and compensation components on student achievement/growth?
4. To what extent, if any, have administrative and instructional personnel changed in their perceptions of RTTT evaluation and compensation components, from the time RTTT was first implemented, to the date of this study?

Methodology

This mixed-methodology study utilized a survey to gather the perceptions of administrative and instructional personnel in central Florida regarding the effectiveness of RTTT at improving student achievement through teacher evaluation and compensation reforms. The survey, titled “Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components of Race to the Top” (Appendix A), contained items for the collection of quantitative and qualitative data. Johnson, Onwuegbuzie, and Turner (2007) describe the value and importance of mixed-method research in the following way, “It recognizes the importance of traditional quantitative and qualitative research but also offers a powerful third paradigm choice that will provide the most informative, complete, balanced, and useful research results” (p. 86).

The survey was replicated, with permission (Appendix B), from the survey and follow-up protocol used in Windish’s (2012) original study of RTTT components. Edits made to the instruments included clarifying components for the demographic information

questions in Part A, replacing Part D with open-ended response items similar to those from Windish's (2012) interview protocol, and general edits to the entire document to clarify language, based upon expert recommendation from educational leaders in the central Florida area. Table 1 identified the linkages between survey items and research questions, ensuring that questions and responses were aligned with the individual research questions that framed this study.

Table 1

Research Questions, Variables, Data Sources, & Statistical Tests

Research Questions	Variables to be Tested (Independent/Dependent)	Data Sources	Statistical Tests
1. To what extent, if any, is there a relationship between administrative and instructional personnel's self-reported knowledge of RTTT and the perceived fairness of RTTT requirements concerning teacher evaluation and compensation?	Self-reported knowledge score/Perceived fairness score	Survey items from Part B of Survey	Pearson Product-Moment Correlation
2. To what extent, if any, is there a difference between administrative and instructional personnel's perceptions of the impact of RTTT teacher evaluation and compensation components on student achievement/growth?	Self-reported professional classification/Perceived impact score	Survey items from Part C of Survey and professional classification from Part A	One-way analysis of variance (ANOVA)
3. To what extent, if any, is there a difference in the perceptions of administrative and instructional personnel who have different self-reported school poverty percentages about the impact of RTTT teacher evaluation and compensation components on student achievement/growth?	Self-reported school poverty percentage/Perceived impact score	Survey items from Part C of Survey and reported poverty percentage from Part A	One-way analysis of variance (ANOVA)
4. To what extent, if any, have administrative and instructional personnel changed in their perceptions of RTTT evaluation and compensation components, from the time RTTT was first implemented to the date of this study?	Self-reported professional classification/Self-reported prior perception score	Survey Items from Part C of Survey and Reported Prior Perception from Part A	One-way analysis of variance (ANOVA)

Population and Sample

The population for this study consisted of educational leaders (instructional and administrative) in the central Florida region. The population included classroom teachers from local elementary, middle, pre-kindergarten-8, and high school settings; school-based administrative deans, assistant principals, and principals; school district-based staff; college and university faculty and staff; as well as members of private businesses in the education industry. A convenience sample of educational leaders pursuing Master's of Education, Education Specialist, or Education Doctorate degrees from a target university were selected to represent the population.

Instrumentation

Since this study was a replication of research completed in 2012, this researcher utilized an updated version of the original survey. The follow-up interview protocol for the collection of qualitative data used in the original study was modified and incorporated as open-ended items in the current survey (Appendix A). Permission to use and modify the survey was provided by the original author, Dr. Daniel Windish (Appendix B). The original survey, entitled "Survey of the Potential Implementation and Impact of Teacher Evaluation and Compensation Elements from the Race to the Top Grant" was edited so that the instrument items reflected current research questions. The title of the new survey was "Electronic Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components from Race to the Top". Knowledgeable leadership professors at the target university, with significant experience using surveys, evaluated content-related evidence of validity. Self-reported demographic information, job

classification and job location data, as well as data on the percentage of students receiving free or reduced lunch at their current or last school was collected. Also of interest for this study was the self-reported data related to the participants' job classification and location prior to the full implementation of RTTT in 2011.

Data Collection

The required authorization for research involving human subjects was obtained from the target university's Institutional Review Board (Appendix D). Subsequent to obtaining informed consent (Appendix E), each participant completed the electronic survey. The link to the survey was included in an email sent to them from their graduate degree program coordinator.

Data Analysis

Data collected from completed surveys was analyzed using the program Statistical Package for Social Science (SPSS). Descriptive and other appropriate statistical tools were used to identify statistical significance within the data and produce the survey findings. Table 1 shows the alignment between research questions, sources of data, and the statistical tests used to analyze the data. Data collected from open-ended survey items was organized into a tabular display that allowed for the identification of themes and categories of responses for each research question. This data was added to the existing data from the quantitative sections of the survey to enhance the findings from the elements in question.

Limitations

Identifying the relationship between the perceptions of university students seeking advanced degrees in Education and Educational Leadership was the focus of this study. As such, the findings are limited in generalizability to urban school districts similar to those found in central Florida. Further, the perspectives of educational leaders seeking advanced degrees may not align with others in the field that choose not to pursue a graduate education. The survey addressed teacher evaluation and compensation components of RTTT, with particular emphasis on how the components affected student learning and academic achievement. The survey did not specifically address other areas of the initiative, about which educational leaders may espouse different perceptions of impact or fairness.

Significance of the Study

Through this study, this researcher intended to contribute to professional knowledge in the field of educational leadership on reform efforts involving teacher evaluation and compensation. The study focused on system-wide initiatives by the federal government, under RTTT, and their perceived impact on student achievement and growth. By the time this study was completed, RTTT was fully implemented in participating Florida districts, allowing educational leaders (who ultimately provided the data for this study) to have first-hand interaction with the RTTT components concerning teacher evaluation and compensation. By probing educational leaders as to their perceptions of RTTT efficacy at improving student achievement through evaluation and

compensation reforms, this researcher hoped to aid future policy makers in their consideration of student achievement-based policies.

Summary

This study was conducted to identify the perceptions of administrative and instructional personnel surrounding teacher evaluation and compensation components of RTTT. The findings from this study will assist policy-makers in their understanding of how perceptions of policies change over the course of their implementation.

The following components were introduced in this chapter: the problem and its clarifying components, the purpose of the study, a conceptual framework, population and sample information, as well as methodology and instrumentation constructs. Additionally, data collection, analysis, and limitations were also presented.

CHAPTER 2 REVIEW OF LITERATURE

Introduction

This review of literature illustrates the rationale for further research on administrative and instructional personnel's perceptions of evaluation and compensation reforms, specifically reforms implemented under the Race to the Top initiative. Much research has been conducted in the area of educational reform and its impact on academic achievement, across the spectrum of student demographic populations (Ladd, 2012; Lubienski & Crane, 2010; Sirin, 2005). The present study sought to build upon that body of knowledge by focusing on efficacy, as seen through the lenses of the educational leaders tasked with executing the reforms.

Researchers in this area have sought to understand the myriad factors that contribute to student achievement at the school-, teacher-, and student-level (Gawlik, Kearney, Addonizio, & LaPlante-Sosnowsky, 2010; Knoeppel, Logan, & Keiser, 2006; Konstantopoulos, 2009; Lubienski & Crane, 2010; Mangiante, 2011). To inform future education policy decisions, it is imperative to understand how factors such as socioeconomic status, teacher quality, and school supports influence students' learning outcomes.

To that end, this researcher explored relevant research in the areas of student achievement, learning and poverty, teacher quality, significant education reform efforts, and teacher compensation and evaluation models. Using Education Full Text, ERIC, Professional Development Collection Education, and PsychInfo databases provided by the university library, this researcher reviewed major journal publications, public and

private institutional reports, university dissertations, and other empirical sources germane to the current study. Similar to Windish (2012), the organization of the material was divided into three sections: (1) improving student achievement, (2) major reforms to public education, and (3) teacher evaluation and compensation.

Improving Student Achievement

To support the growth and maintenance of an effective, high-quality public education system in the United States, educational policy makers must maintain an intense focus on student achievement, particularly for the most disadvantaged student populations. A problem first raised by *A Nation at Risk: the Imperative for Education Reform* (National Commission on Excellence in Education, 1983), “the rising tide of mediocrity” (p. 1) threatens to undermine the long-term success of generations of U.S. students, as well as that of the nation as a whole. *A Nation at Risk* (1983) catalyzed a shift in public perceptions of public education that sought immediate remedy to systemic failures that hindered student achievement.

The level of achievement actualized by students is the conceptual assessment to which school teachers and administrators are held accountable. In recent education reforms, such as *Race to the Top* (2009), the ideas of accountability and assessment have become rigidly intertwined with one another. RTTT reforms installed accountability systems that used high-stakes assessments to measure student achievement and to make potentially adverse appraisals of teacher and administrator quality. The following subsections detail the research found related to learning, poverty, and teacher quality, and

establish an appropriate understanding of the myriad factors that impact student achievement as well as the effect educational reforms have on them.

Learning and Poverty

As evidenced by the research conducted by Sirin (2005), there is a direct link between poverty and learning outcomes. In Sirin's (2005) analysis of existing literature related to socioeconomic status (SES) and student achievement, the researcher concluded the following:

Our society may be failing in one of the greatest commitments of every modern society, that is, the responsibility to provide educational opportunities for each student regardless of social and economic background.... At present, one in five children in the United States lives in poverty, which puts many of these students at risk for poor school performance or failure. (pp. 445-446)

SES is a complex conceptualization, using multiple indicators, of an individual's position within society (Hattie, 2009). SES is one of the most frequently used contextual variables in educational research (Sirin, 2011, p. 417). Three common determinants of SES are parental income, parental education, and parental occupation (Hattie, 2009). However, researchers that study socioeconomic status in education often utilize only a single measure, Free and Reduced Lunch status, to identify poverty in their studies. Some researchers have raised concerns as to the validity of FRLS as an indicator of poverty (Lubienski & Crane, 2010; Sirin, 2005). "The use of participation in school lunch programs as a measure of SES, though common, is conceptually problematic" (Sirin, 2005, p. 444). Sirin found that FRLS was an imprecise measure that only weakly correlated to student achievement in higher grade levels (2005). Nevertheless, using FRLS as a measure of poverty provides a mechanism for researchers to estimate poverty

levels in a given population. Additionally, more detailed information may not be readily available and prohibitive for researchers to collect. That being said, researchers should apply caution when using such a measure within their studies (Sirin, 2005).

In acknowledgment of the limitations associated with using FRLS as a measure of SES, Lubienski and Crane (2010) sought to identify family background characteristics that were strong correlates of poverty. In their study, the researchers analyzed family demographic data from The Early Childhood Longitudinal Study—Kindergarten Class 1998 - 1999 (Lubienski & Crane, 2010). In their findings, they identified a number of specific behavioral differences between low- and high-SES families that correlate to student achievement. The number of books available and the presence of a computer in the home, age of the biological mother at her first birth, and access to pre-school care were found to be stronger indicators of reading and math achievement levels in kindergarten than traditional poverty measures (Lubienski & Crane, 2010). The researchers also found a strong relationship between participating in music lessons and achievement level. They stated that while "involvement in music lessons might not *cause* higher achievement [it] might instead serve as a proxy for other important aspects of parents' priorities, practices or resources" (Lubienski & Crane, 2010, p. 19). To wit, diversity of experience, like involvement in the arts, in early years can have a significant effect on achievement in subsequent years (Lubienski & Crane, 2010). The researchers argue that understanding the characteristics of poverty that correlate to academic achievement, such as the use of family resources, could have important policy and research implications for addressing poverty. As an example,

Given that the mean mother's age at first birth was less than 20 for the lowest-SES quintile of children, and given that mother's age at first birth was a significant predictor of children's achievement even after adjusting for other key demographic factors, this study suggests that programs aimed toward helping low-SES teens delay parenthood could hold promise for weakening generational poverty. (Lubienski & Crane, 2010, p. 23)

Generational poverty, one of the most difficult types of poverty to overcome, occurs where at least two generations of a family are born into poverty; families experiencing generational poverty do not have the faculties necessary to improve their situation and change their long-term life outcomes (Jensen, 2009).

The effects of poverty on children can be most acute in the early years of life (Hattie, 2009; Ladd, 2012). In supporting this fact, the researchers concluded,

Research documents a variety of symptoms of low SES that are relevant for children's subsequent educational outcomes. These include, for example, poor health, limited access to home environments with rich language and experiences, low birth weight, limited access to high-quality preschool opportunities, less participation in many activities in the summer and after school that middle-class families take for granted, and more movement in and out of schools because of the way the housing market operates for low-income families. (Ladd, 2012, p. 206)

Further, when comparing National Assessment of Education Progress (NAEP) data from years 1998-2009, Ladd (2012) identified within-state score variations in reading ($a = .10$) and mathematics ($a = .01$) for fourth and eighth graders that inversely correlated to changes in child poverty levels in the same state. "Consistent with the view that child poverty adversely affects student achievement, increases in child poverty rates during the last 10 years translated into reductions in average test scores" (Ladd, 2012, p. 208).

A family's position within society is also evidenced by the neighborhoods in which they live and the schools their children attend. Higher instances of violence, homelessness, illegal drug use, and exposure to environmental toxins, as well as

insufficient healthcare and nutritional food options are found in low-SES neighborhoods (Gallagher, Goodyear, Brewer, & Rueda, 2012; Hattie, 2009; Ladd, 2012; Sirin, 2005). Low-SES community attributes place low-SES students at an immediate disadvantage to their high-SES peers when considering school quality and the availability of resources. Nearly 50 percent of school funding is based on property values within a school district and the taxes they generate (Sirin, 2005). As a result, students from low-SES families often attend schools that are chronically under-funded, providing fewer resources to address their complex needs (Lagana-Riordan & Aguilar, 2009; Sirin, 2005). Beyond funding, researchers point to inexperienced teachers, low expectations, and higher teacher turnover and absenteeism as consequential impediments to achievement and upward economic mobility for low-SES students (Gallagher et al., 2012; Gawlik, Kearney, Addonizio, and LaPlante-Sosnowsky, 2010; Konstantopoulos, 2009). And as an unfortunate, if not predictable, outcome, “these events tend to rule out college as an option and perpetuate the cycle of poverty” (Jensen, 2009, p. 9).

Maslow’s hierarchy of needs theory establishes certain base needs, among them food, shelter, and security, that must be satisfied before higher-level achievement in life is possible (Robbins & Judge, 2010). Children living in poverty are less likely to experience the stability necessary to meet their basic needs. Higher rates of intra-district mobility and other instability in a student’s home life, impede long-term development (Popp, Grant, & Stronge, 2011). Accounting for variations in mobility rates is more troublesome given the current use of standardized testing and teacher accountability models. Gallagher et al. (2012) summarized the issue as follows.

A data-driven system in high mobility environments is less effective in part because no one takes responsibility for the learner who is moving among schools frequently, and it is more difficult to establish the value that any teacher or school is contributing to students who come and go throughout the year. (p. 160)

The challenges surrounding SES and academic achievement are not uniquely American phenomena. In nations around the world, high-SES students consistently outperform their low-SES peers (Ladd, 2012). Consistent with that finding, the 2009 Programme for International Student Achievement (PISA) test revealed that in reading, nations with higher percentages of low-SES students had lower test scores (Ladd, 2012). Given that nations such as Canada, Finland, and the Netherlands consistently outperform the U.S. on measures of student achievement like the PISA, it may not be a surprise that the U.S. has approximately twice as many students from low-SES families (with income less than 50% of the country's median household income) as higher performing nations (Ladd, 2012). While the statistic may explain some of the variation in performance levels between the nations, a comparison of scores for low-SES students within developed nations demonstrates that there is more to the disparity than what may be accounted for due to variations in population size alone,

U.S. students in families with [SES] below the median perform particularly badly relative to their low-[SES] peers in other countries, while U.S. students from more advantaged backgrounds perform reasonably well by international standards. That is, the largest shortfalls in performance among U.S. students are concentrated among those with relatively low [SES]. These shortfalls suggest there is room for the United States to do better by its disadvantaged students. (Ladd, 2012, p. 210)

As demonstrated, the challenges faced by low-SES students are myriad. Family characteristics associated with poverty limit the early development of children. Limited educational resources in the home result in early deficits in language and numeracy skills;

low-quality housing and neighborhood options expose low-SES children to environmental toxins, drugs and violence, poor nutrition, and decreased access to early childhood healthcare; the schools serving low-SES students are underfunded, reducing access to critical resources and supports, and the teachers working there have less experience and hold more negative views about the students they teach (Jensen, 2009). Students who live in poverty historically underperform compared to their high-SES peers. However, effective teachers providing the right supports have been shown to improve educational outcomes for low-SES students (Mangiante, 2011).

Given that, it is important to analyze the diverse teacher-level characteristics that improve student achievement as a way of developing an accurate conceptualization of quality teaching. The next section will discuss research on instructional practice and teacher attributes that affect teacher quality.

Teacher Quality

Teacher quality is regularly cited as the most significant school-level factor affecting academic achievement (Brown & Crumpler, 2013; Goldhaber, Goldschmidt, & Tseng, 2013; Goldhaber & Walch, 2012; Ladd, 2012; Mangiante, 2011). However, the specific attributes that constitute quality do not share a similar consensus in the literature. As a result, researchers continue to examine a variety of characteristics associated with teachers and the relationship the characteristics have to student achievement. There is evidence that teacher effects on student achievement are cumulative, to wit, having effective teachers in consecutive years significantly increases student learning (Konstantopoulos, 2009). Having a complete understanding of the qualities that

constitute teaching quality could have important long-term implications for the effectiveness of the public education system, particularly for the students most impacted by the achievement gap.

Historically, educational leaders have used a teacher's educational background, certification level, and years of teaching experience as a proxy for quality of instruction (Gawlik et al., 2010; Okpala, 2000). However, there is a body of research that challenges the use of these measures in isolation, as they do not account for the variety of teacher attributes that encompass quality instruction (Knoeppel, Logan, & Keiser, 2005; Mangiante, 2011; Okpala, 2000; Palardy & Rumberger, 2008).

State certification requirements and salary structures that address experience and degrees [held] are widely used as an effort by states to affect the quality of teachers. However, what teachers know about teaching and learning must be evidenced also by teaching practice. (Knoeppel et al., 2005, p. 8)

A study by Mangiante (2010) identified behaviors and characteristics of effective teachers in high poverty, minority schools. From a review of relevant qualitative studies, the researcher identified four main categories of behaviors that were consistently demonstrated by effective teachers. The behaviors were defined to include (a) personal beliefs and convictions that their students could succeed in rigorous coursework and that as teachers they were integral to their students' achievement; (b) instructional practices like setting clear learning goals, using well-scaffolded lessons, regularly monitoring student progress, and adjusting to individual needs; (c) interpersonal skills that enabled teachers to develop a trusting and committed relationship with their students wherein students were committed to their work in recognition of a commitment to their teacher; and (d) professional self-reflection on lesson outcomes, their causes, and ways to

improve, as well as regular collaboration among peers and the sharing of best practices. (Mangiante, 2010).

Other researchers found a similar correlation between teacher beliefs and student achievement (Gallagher et al., 2012; Lagana-Riordan & Aguilar, 2009; Popp et al., 2011). Effective teachers were more likely to develop caring relationships with their students, have solid classroom management strategies, and plan extensively for their instruction (Popp et al., 2011). Grossman, Loeb, Cohen, and Wyckoff (2013) found the use of Explicit Strategy Instruction, Guided Practice, and Intellectual Challenge strongly correlated to student growth and achievement in English Language Arts classrooms. Unfortunately for the students most in need of quality instruction, researchers noted that the effective strategies and practices were most commonly associated with teachers in high-SES schools (Gawlik et al., 2010; Konstantopoulos, 2009; Popp et al, 2011).

Table 2

Summary of Literature Reviewed: Improving Student Achievement

Subsection Summary of Findings	Authors
<p><i>Learning and Poverty:</i> SES is a major predictor of a child's educational outcomes. Low-SES students often lack access to resources necessary to support learning. However, high-quality instruction and teachers who are attuned to this issue can impact the achievement of economically disadvantaged students.</p>	<p>Gallagher, Goodyear, Brewer, & Rueda (2012); Gawlik, Kearney, Addonizio, & LaPlante-Sosnowsky (2010); Hattie (2009); Jensen (2009); Konstantopoulos (2009); Ladd (2012); Lagana-Riordan & Aguilar (2009); Lubienski & Crane (2010); Mangiante (2011); Okpala (2000); Popp, Grant & Stronge (2011); Robbins & Judge (2010); Sirin (2005)</p>
<p><i>Teacher Quality.</i> Teacher quality impacts student achievement. While the constituent components of teacher quality are not universally agreed upon, research shows that traditional measures of teacher quality are not strong correlates of student achievement.</p>	<p>Brown & Crumpler (2013); Gallagher et al. (2012); Gawlik et al. (2010); Goldhaber, Goldschmidt, & Tseng (2013); Goldhaber & Walch (2012); Grossman, Loeb, Cohen, & Wyckoff (2013); Knoeppel, Logan, & Keiser (2005); Konstantopoulos (2009); Ladd (2012); Lagana-Riordan & Aguilar (2009); Mangiante (2011); Okpala (2000); Palardy & Rumberger (2008); Popp et al. (2011)</p>

Major Reforms to Public Education

The education of our people should be a lifelong process by which we continue to feed new vigor into the lifestream of the Nation through intelligent, reasoned decisions. Let us think of education as the means of developing our greatest abilities, because in each of us there is a private hope and dream which, fulfilled, can be translated into benefit for everyone and greater strength for our Nation. (Kennedy, 1961)

A nation-wide system of free public schools, that provide all students the skills necessary to become productive citizens, is the keystone of social stability and economic prosperity (Mitra, 2011). Long a source of national pride, in the decades following the publication of *A Nation at Risk: The imperative for education reform* (National Commission, 1983), the U.S. public education system has been perceived as an ineffective institution, in need of reform.

As an acknowledgement of the importance of education to the public, candidates for elected office, at all levels of government, have championed proposals of myriad size and scope, to answer the call first sounded by *A Nation at Risk* (1983). Recent federal initiatives, like NCLB and RTTT, were developed and implemented to affect nation-wide educational outcomes, while many states instituted their own reforms that addressed local priorities, values, and standards. However, the empirical evidence provided to substantiate the quality of such reforms has varied (Levine & Levine, 2012), and subsequently, the results of the reforms are likewise, mixed. Hattie (2009) summarized one problem with previous reforms as follows:

So often money is added into the education system with little attention to the efficiency or effectiveness of education outcomes. It is not the amount of money spent that is important, but how it is spent. (Hattie, 2009, pp. 74-75)

In the following section, research on both federal and state educational reforms will be presented. Relevant literature on NCLB and RTTT were reviewed, as well as

studies that evaluated reforms to teacher evaluation and compensation systems by the states.

No Child Left Behind

The reauthorization of No Child Left Behind (NCLB) (NCLB, 2001), constituted a significant shift in national education policy “That dramatically expanded the historically limited scope and scale of federal involvement in K-12 schooling” (Dee & Jacob, 2011). NCLB originated developed in the state of Texas, under then-Governor George W. Bush. In the years following implementation of the state-level program in Texas, it was regarded as being highly successful at improving educational outcomes for all demographic subgroups, and was regularly touted by Bush while a presidential candidate (Dee & Jacob, 2011). Subsequent to Bush’s Presidential inauguration in 2001, NCLB, the federal version of the Texas education program, was passed by Congress and signed by the new President. The intent of the legislation was as follows:

[NCLB aims] to decrease the achievement gap and improve student performance so that 100 percent of U. S. students will meet predetermined standards in reading and math by the 2013-2014 school year. [NCLB has] three major requirements: that states (1) develop content standards to determine what students should know, (2) administer assessments to measure whether students are meeting those standards, and (3) institute accountability mechanisms to ensure that all students attain the proficiency standards. (Lagana-Riordan & Aguilar, 2009, p. 136)

NCLB required new standards for teachers meant to ensure only highly qualified teachers would be allowed in the classroom. To achieve this, the Act set minimum standards for teacher certification to include completion of a bachelor’s degree, a state-issued teaching certification, and demonstrated knowledge of the content being taught (Education Commission, 2007). In subsequent years, researchers have evaluated the efficacy of

certification as a predictor of teacher quality with generally negative results. Kane, Rockoff, and Staiger (2008) found that certification does not assure effectiveness.

There is not much difference between certified, uncertified, and alternatively certified teachers overall, but effectiveness varies substantially among each group of teachers. To put it simply, teachers vary considerably in the extent to which they promote student learning, but whether a teacher is certified or not is largely irrelevant to predicting their effectiveness. (Kane et al., 2008, p. 41)

Additional research analyzing the overall implications of NCLB also produced mixed results (Lagana-Riordan & Aguilar, 2009; Lee & Reeves, 2012; Palardy & Rumberger, 2008; Levine & Levine, 2012). Lagana-Riordan and Aguilar (2009) also identify a number of troubling issues to arise from the legislation. Particularly troubling, the researchers reported, were the impacts of the increased accountability measures on low-SES students and the school drop-out rates for impoverished populations.

Accountability mechanisms based on test scores can have a disparate impact on schools with larger populations of minority and low-income students. Small schools and those with highly concentrated at-risk and mobile populations, such as schools in urban and rural areas, are also more likely to fail to make AYP. Even more concerning is the evidence that accountability systems are exacerbating problems such as grade retention rates and dropout rates for minority and low-income students, even in states that claim that the achievement gap is closing. These phenomena have been linked to the intentional or unintentional retention of minority students in grades immediately preceding a 'testing grade' and the 'pushing out' of minority students who seem likely to negatively influence school test scores. This is especially concerning because students who have been retained are more likely to eventually drop out of school. (Lagana-Riordan & Aguilar, 2009, p. 137)

However, in the researchers' analysis of NCLB policies, they were also able to identify a number of positive outcomes (Lagana-Riordan and Aguilar, 2009). Specifically, as a direct result of the increase in the use of data, required by NCLB, schools and school districts developed more effective systems to collect and analyze that data; contributing to organization-wide efficiency improvements (2009). The researchers also point to

improvements in the quality of educational leaders as having a positive impact on student achievement; high-quality leaders tended to respond positively to high-stakes accountability models (Lagana-Riordan & Aguilar, 2009).

After reviewing NAEP data from 1990-2009, Lee and Reeves (2012) found similarly mixed if not negative progress toward closing the reading and math achievement gap. Specifically, they found that when considering pre-NCLB achievement gap trends for racial and SES subgroups, post-NCLB progress remained the same or slowed (Lee & Reeves, 2012, p. 216).

Race to the Top

The next iteration of federal education reform enacted after NCLB was known as Race to the Top (RTTT). RTTT was a federal education initiative that offered State Educational Agencies grant dollars in exchange for designing and implementing major changes to their state education programs that increased rigor in the curriculum and improved teacher and administrator effectiveness (Florida Department of Education, 2011). The program was passed as a component of the American Recovery and Reinvestment Act (ARRA) of 2009. This Act appropriated roughly 4.35 billion dollars to distribute to states that successfully completed a competitive grant application outlining their vision for reform (Smarick, 2011).

Among the issues to be addressed in a state's application, states needed to develop systems that improved the collection of student data "to measure student growth and to tie results to individual teachers" (Smarick, 2011, p. 61). Innovation in data collection by the states presents a major opportunity to provide teachers and administrators with a

quantity and quality of student data not previously available. The legislation asked states to then use that student data as a component of new teacher performance evaluations. Other areas weighing on a grant application were the states' proposals to improve low-performing schools, and the extent to which they were able to obtain stakeholder buy-in for their proposals (Smarick, 2011).

RTTT was a shift from the NCLB focus on teacher qualifications to an output-focused measure of teacher effectiveness. Effectiveness was defined by the initiative to mean,

A teacher whose students achieve acceptable rates (*e.g.*, at least one grade level in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, multiple observation-based assessments of teacher performance. (Florida Department of Education, n.d.)

States were tasked with developing evaluation systems that were able to accomplish the goal of identifying teacher effectiveness. In the fourth year of RTTT implementation, the development of evaluation systems was recognized as the most challenging component for states to actualize (McNeil, 2014). One researcher bluntly described it, and its impact on the broader initiative, as follows: "If [evaluation reform] doesn't work out, it will hurt the long-term legacy of RTTT -- it'll be another sign that the feds can get states and districts to do things but they can't make them do it well" (McNeil, 2014, para. 4). Many of the applications that earned grant dollars provided few binding details about specific components of their reforms. In 34 state applications, an explicit commitment was made to use evaluation systems to identify professional development needs, but only nine states were willing to tie the new evaluations to potentially adverse

employment decisions (Smarick, 2011). Among the states that proposed performance-pay plans (16), most (9) lacked sufficient detail or structure to estimate the likelihood that such plans would ever reach widespread implementation or be sufficiently effective at improving student outcomes (Smarick, 2011).

As one of the Obama Administration's largest domestic policy items, the success of RTTT is of critical importance to the Obama Presidential legacy (McNeil, 2014). An important factor impacting the long-term success of RTTT is the ability of states and school districts to maintain their reforms after federal grant dollars are exhausted. Performance-pay programs that are successful at raising teacher effectiveness will be more expensive to the states and school districts that adopt them. One school district superintendent posed the question: "If we don't have money for raises at all, how are we going to implement [performance-pay]?" (Boser, 2012, p. 30).

Florida's Race to the Top

At the time RTTT was passed into law, the state of Florida had already begun to implement its own reforms to the state's teacher evaluation and compensation systems (Jones, 2013). The 2007 Merit Award Plan (MAP) established the requirement that "at least 60% of a teacher's bonus must be based on student performance" (Jones, 2013, p. 150). The plan also stipulated that the bonuses were to be distributed on an individual level, instead of as a school-level bonus utilized elsewhere (Jones, 2013).

Within Florida's RTTT application was the requirement that all participating LEAs "make student growth the most significant component of compensation, ahead of

years of experience and academic degrees” (Smarick, 2011, p. 62). Additionally, participating Florida school districts were required to develop and administer standards-based assessments for all courses taught from kindergarten through 12th grade. The new testing requirement was a significant shift for the state, particularly in grades K-2 and in subject areas, like visual and performing arts and physical education, which had not previously been required to administer year-end assessments (Boser, 2012).

The requirement for new assessments throughout the public school system represented a boon for private contractors tasked with developing the new tests (Boser, 2012; Kolbe & Rice, 2012). While Florida planned to utilize only 20% of its RTTT funding to improve teacher and administrator effectiveness, their RTTT application anticipated using approximately 40% of award funds on outside contractors (Kolbe & Rice, 2012). The use of contractors was a consistent theme for the majority of RTTT applications, however Koble and Rice (2012) noted that only Hawaii had plans to spend as large of a percentage of their state-level award on private organization and contractors. Florida’s plan allocated approximately \$82 million to create new assessments in mathematics and reading, as well as interim assessments in the remaining core content areas (Kolbe & King, 2012). Additionally, approximately \$46 million was spent on an instructional tools database, and a textbook use study (Kolbe & King, 2012). The U.S. Department of Education tacitly endorsed the sizeable reforms proposed in Florida’s RTTT application when, in August 2010, the state was awarded 700 million in RTTT dollar to fund them; the largest award for a single state by the initiative (Boser, 2012; U.S. Department of Education, 2014).

Table 3

Summary of Literature Reviewed: Major Reforms to Public Education

Subsection Summary of Findings	Authors
<p><i>No Child Left Behind.</i> NCLB required states to raise standards for teachers. Established parameters to identify “highly qualified” teachers. Required schools to improve proficiency of all student subgroups with a goal of 100% proficiency by 2014.</p>	<p>Dee & Jacob (2011); Education Commission of the States (2007); Kane, Rockoff, & Staiger (2006); Lagana-Riordan & Aguilar (2009); Lee & Reeves (2012); Levine & Levine (2012); NCLB (2001); Palardy & Rumberger (2008)</p>
<p><i>Race to the Top (RTTT).</i> RTTT has four priorities: (1) More rigorous education standards; (2) Develop and integrate better data systems; (3) Measure teacher and principal effectiveness; (4) Turnaround low-performing schools. RTTT prioritizes “highly effective” over “highly qualified”. Largest federal education reform effort to push for adoption of performance-pay and high-stakes testing.</p>	<p>Boser (2012); Florida Department of Education (n.d.); Florida Department of Education (2011); McNeil (2014); Smarick (2011)</p>
<p><i>Florida’s RTTT.</i> Florida spending larger percentage of grant funds on independent contractors than other states. Requiring districts to develop standardized assessments for all K-12 subject areas. Districts concerned about the financial impact of new evaluation and compensation systems.</p>	<p>Boser (2012); Smarick (2011); Jones (2013); Kolbe & Rice (2012); U.S. Department of Education (2014)</p>

Teacher Evaluation and Compensation

Two of the main areas of focus under Race to the Top were teacher evaluation and compensation. To appropriately frame the current study, it is necessary to understand current evaluation and compensation systems as well as their evolution over time. Further, an analysis of the purposes and perceptions of compensation and evaluation systems, and their recent reforms, is necessary to contextualize the elements of focus in the current study.

Considering the impact reforms have on teachers and administrators, understanding their professional perceptions concerning the evaluation and compensation components of RTTT will provide useful insights to inform future research and policy priorities. In the following subsections, the purposes, perceptions, and evolution of teacher evaluation and compensation will be presented.

Evolution of Teacher Evaluation

The advent of Frederick Taylor's Scientific Management, which focused on measurement of productivity and efficiency, shaped teacher evaluation systems throughout the early part of the 20th century (Marzano, Frontier, & Livingston, 2011). As a result, feedback systems that utilized data and standardized tests were common in the education system, well into the 1940s (Marzano et al., 2011). At the conclusion of World War II, views of teacher evaluations began to shift toward a more individualized consideration of teacher development and the impact on student learning (Marzano et al., 2011). The focus on individual teachers coincided with a rise in supervisor observations as a way to better understand and improve the effectiveness of teacher practice.

The development of contemporary evaluation systems began in the 1980s with the publication of *A Nation at Risk* (National Commission, 1983) and *Teacher Evaluation: A study of effective practice* (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). These works propagated a narrative related to public education that focused on increased accountability, standardization of evaluation systems, specific training for evaluators, and an efficient allocation of educational resources (Marzano et al., 2011; Wise et al., 1983).

Later, the 21st century saw a shift from teacher behaviors to student achievement and the early proponents of linking student achievement data to teacher evaluation were honing their message (Marzano et al., 2011). As a result, NCLB (2001) was the first major effort to tie teacher-level outcomes to student subgroup proficiency in reading and math; RTTT (2009) sought to evaluate teachers on students' growth (Education Commission of the States, 2007; Smarick, 2011). Much of the evolution of evaluation systems has been informed by shifts in perceptions of their utility, by major educational stakeholders. The purposes and perceptions of evaluations are discussed in the following section.

Purposes and Perceptions of Teacher Evaluation

According to Delvaux et al. (2012), the most frequently cited purposes of teacher evaluation are accountability and professional development. The design of the evaluation, the process by which it is administered, and by whom it is administered, impact the perceptions of evaluation systems by the educators working under them (Delvaux et al., 2012).

The perceived fairness of the evaluation system is related to the acceptance of the evaluation system... [and the] procedural and distributive justice of the system.

Procedural justice concerns the different components of the procedure of evaluation, such as the attribution of an evaluator, the way in which evaluation criteria are established, and the existence of and the reactions toward an appeal procedure. Distributive justice concerns the understanding that teachers' received assessment is in line with their performance. (Delvaux et al., 2012, p. 3)

Wise et al., (1984) found that the majority of teachers surveyed believed their principals lacked the “resolve and competence” (p. 22) to complete accurate evaluations. Further, teachers believed their principals viewed the evaluation system as a burden which took time away from other areas of responsibility (Wise et al., 1984). School district administrators expressed concerns over the validity of the evaluations because of principals' desire to maintain positive relationships with their staff: “Principals' disinclination to be tough makes the early identification of problem teachers difficult and masks variations in teacher performance” (Wise et al., 1984, p. 22). That notwithstanding, the researchers consistently found that teachers identified “improved teacher-administrator communication and increased awareness of instructional goals and classroom practices” (Wise et al., 1984, p. 23) as positive effects of their evaluations.

Public perceptions of evaluation systems for teachers are consistently more negative. Aritomi et al. (2009) found that the average school district released approximately four teachers per year from employment in their schools. That information, coupled with weaker student performance on international student assessments, informed a public narrative that systems to evaluate teachers were ineffective, with teacher tenure specifically, being a contributing factor (Howell, Peterson, & West, 2011). In their study, approximately half of all respondents opposed teacher tenure while supporting increased teacher accountability for learning outcomes (Howell et al., 2011).

To improve both the real and perceived effectiveness of teacher evaluations, many researchers support the development of multiple measures of teacher performance to improve the accuracy and reliability of teacher evaluations (Brown & Crumpler, 2013; Burnett et al., 2012; Di Carlo, 2012; Grossman et al., 2013; Mangiante, 2011). Principal observations and peer assessments of instructional quality used in conjunction with student performance data and teacher-level goals or objectives (Goldhaber & Walch, 2012) all have proponents in the literature. Given that the complexity of teaching is hard to assess by a single instrument, multiple measures address concerns voiced by opponents of any one measure while holding teachers accountable for their performance.

Evolution of Teacher Compensation

Teacher compensation systems for the last 50 years of the 20th century were based primarily on a single salary schedule (Laine, Potemski, & Rowland, 2010), with some estimates at “nearly 100%” homogeneity (Podgursky & Springer, 2007, p. 552). By the early 21st century, shifts in national policy priorities facilitated a shift in compensation strategy as well. The traditional single salary schedules that differentiated individual salaries based on years of teaching experience and advanced degrees began to incorporate outcome-based compensation structures that considered student performance on standardized tests as a measure of teacher quality; high-quality teachers receiving bonus-based remuneration (Laine et al., 2010).

According to the 2011-2012 School and Staffing Survey (SASS), from the National Center of Education Statistics, 89.2 percent of school districts in the U.S. employed salary schedule for teachers; a decrease of 3.2 percent from the 2007-2008

school year (U.S. Department of Education, 2012). Among the school districts included in the survey, the average base salary for a teacher with a bachelor's degree and no teaching experience was \$35,500, while the average base salary for a novice teacher with a master's degree was \$38,700 (U.S. Department of Education, 2012). When considering years of teaching experience, average base salary for teachers with a bachelor's degree and 10 years of experience rose to \$44,900 and a teacher with the same experience and a master's degree received an average base salary of \$49,500 (U.S. Department of Education, 2012). In the same SASS year, 45% of districts offered teachers tuition reimbursement, approximately 25% offered incentive pay for National Board Certification, just over 11% for excellence in teaching, 6% for teaching in less desirable schools, and 13.5% of districts offered incentive pay to teachers in fields of shortage (U.S. Department of Education, 2012).

Recent educational reform efforts at the national level and their associated funding, have given "state and local education leaders an opportunity to push back on historical barriers that have prevented states and districts from defining and measuring teacher effectiveness, but also to align decisions about educator pay with evidence of effectiveness" (Laine et al., 2010). The increased funding that federal education initiatives offered to states to reform their teacher compensation systems is short-lived, and may do little in the long-run to sustain the programs. Performance-pay systems largely rely upon bonuses to recognize high-performing teachers; a design which may preclude states from embracing the programs where budgets are limited (NGA, 2011).

One researcher proposed recommendations for future teacher compensation designs that focus on the following elements (Shields, 2012): (1) A base salary

competitive with other professions that attract top academic candidates, raises commensurate to performance and contribution level, with differentiation sufficient to attract individuals with the knowledge and skills desired in the profession (p. 7), (2) Compensation incentives for teachers in the lowest-performing schools that are of a sufficient magnitude to adequately convey the difficulty associated with those types of teaching positions (p. 12), (3) Provide for role differentiation that involves increased responsibilities and involvement with more difficult student populations. Role differentiation within school structures would help retain the highest performing teachers at the school-level and compensate them in a way that considers the economic value of role (p. 16), (4) Pay incentives based upon student performance that are incorporated into base pay instead of provided as a one-off award (p. 21), and (5) Provide a financially sustainable system that uses limited district resources in the most effective way, and that is prioritized such that the potential need for reallocation of resources from other areas or even raising new revenue is accepted (p. 24).

Purposes and Perceptions of Teacher Compensation

While the evolution of teacher compensation systems in the U.S. is complex, so too are the purposes of the systems and the perceptions of them, held by the public, educators and educational leaders. Compensation systems are created to allow organizations to attract and retain a high-quality workforce (Shields, 2012). The traditional salary schedules used in education that differentiate based upon years of teaching experience and level of education are not without their benefits. Goldhaber and Walch (2012) described their advantages, “Everyone is rewarded equally based on

objective criteria, it is predictable and easy to understand, and it is favored by teachers over alternative systems” (p. 1068).

Traditional compensation systems were designed to support a seasoned educator workforce at a time when years of teaching experience was thought to strongly correlate to student achievement; a theory challenged by contemporary research (Gawlik et al., 2010; Laine et al., 2010; Okpala, 2000; Shields, 2012). Advocates for alternative compensation systems have increased public support for performance-based pay to greater than 70 percent in 2009 (Laine, 2010). Regardless of the type of compensation system utilized, the American public believes teacher salaries are too low (Dolton & Mercenaro-Gutierrez, 2013). However, a 2013 study found that teachers’ salaries were “considerably higher than the respondents thought they were and higher than the respondents identified to be a fair wage” (Dolton & Mercenaro-Gutierrez, 2013, p. 49). The same survey found that 80% of respondents in the U.S. supported some form of performance-based compensation for teachers (Dolton & Mercenaro-Gutierrez, 2013).

Teachers are less supportive of performance-based compensation systems, particularly when including a value-added model to estimate student growth. Teachers believe that the model is unfair and difficult to understand (Boser, 2012). Possibly as a result, teachers indicate an increased willingness to leave the profession as a result of performance-pay implementation in their district (Jones, 2013). The findings noted above are in contrast to that of school administrators. School administrators tend to favor performance-pay programs to improve teacher quality (Lagana-Riordan & Aguilar, 2009)

Reforms to Teacher Evaluation and Compensation

The traditional system of teacher evaluation and compensation has its roots in the early 20th century (Consortium, 2012). Since that time, the most commonly used attributes to estimate teacher quality and in turn, teacher evaluation and compensation have been: level of education, years of teaching experience, and type of certification (Gawlik et al., 2010). Beginning in the 1990s however, local, state and federal education agencies sought to bring about substantive changes to the way teachers are evaluated and compensated for their work (Consortium, 2012). Evaluation and compensation reforms have used myriad systems and structures, incentives and enforcements to reach a similar goal: improving teacher quality, the rigor in curriculum, and overall student outcomes. Contemporary compensation systems identified four main teacher-level areas to receive salary supplements: (1) National Board Certification; (2) Excellence in teaching; (3) recruitment and retention of teachers in the lowest performing schools; and (4) recruitment and retention of teachers in high-need subject areas (Aritomi et al., 2009). However, educational researchers have questioned the practice due to the fact that the factors only weakly correlate to student achievement and learning (Gawlik, 2010; Knoepfel et al., 2005; Palardy & Rumberger, 2008).

Teaching is a complex, multifaceted process that is difficult to evaluate with a single tool. Researchers posit, “because no single measure can identify all strengths and weaknesses of teacher practice, performance-based compensation systems should include multiple measures of performance to accurately identify areas of needed support” (Burnett et al., 2012, p. 3). Trends in evaluation and compensation reform have incorporated multiple measures, typically value-added models and principal observations,

to differentiate the quality of instruction and compensate teachers accordingly. Even considering the multiple measures incorporated in many new evaluation systems, Master (2014) summarized their limitations.

While these measures address central aspects of teachers' work, they may do a poor job of accounting for teacher impacts on valued student outcomes other than actual tested achievement, such as motivation, character development, or achievement outside the scope of standardized tests. They may also miss valuable teacher contributions that occur outside of regular classroom practices, such as organizational leadership, relationships with students and families, or collaboration with peers. (pp. 207-208)

The following sub-sections review the major evaluation and compensation elements present in the literature. The pertinent research and empirical evidence offered by advocates and opponents of each element's use is discussed.

Value-Added Models

The development of the value-added model (VAM) sought to reinterpret traditional differentiated compensation models. Originally, VAMs used student test data and socio-economic demographics in a complex statistical tool meant to provide a more precise estimate of the quality of teacher-level inputs (Sanders & Rivers, 1996). The Tennessee value-added assessment system (TVAAS), designed by researchers at the University of Tennessee, was one of the first mechanisms developed to utilize such an approach for evaluating and compensating teachers (Schafer et al., 2012). TVAAS incorporated multiple years of performance data from standardized tests in several subject matter areas (Schafer et al., 2012) to estimate teacher contribution to student growth above what might be expected from a student, given prior levels of performance (2012). In subsequent years, value-added models have attracted much attention by

education reform advocates. Despite VAMs popularity among education reform advocates and its inclusion in education reforms like RTTT (2009), research into the effectiveness of VAM has found mixed results.

Critics of the model have challenged VAM's use of student-level characteristics as though they have a linear impact on a teacher's effectiveness in the classroom. Konstantopoulos (2009) believes that there is a fundamental flaw in the assumption that the factors used in developing a value-added model account for all of the student-level factors effecting teacher quality. As one researcher stated, "such [student-level] factors are not just main effects easily controlled, [rather] they interact with the teacher's ability to be effective all year long and they interact with other student factors, as well" (Schafer et al., 2012, p. 2).

Researchers also note that the use of VAM to evaluate teachers is complicated where large numbers of teachers teach courses not previously subject to test-based accountability (Schafer et al, 2012; Mangiante, 2011). One study found that as of 2009, approximately 70% of teachers in the state of Florida taught subjects or classes with no year-end testing (Schafer et al., 2012). As a result, school districts are forced to create year-end tests for all subjects and rely upon school-level VAM scores to evaluate teachers where sufficient historical student data is not yet available (Boser, 2012). In Florida specifically, its own policy precedent prohibiting school-level awards contradicts the utility of the design. Florida's mandated individual-level performance-pay incentives were shown to have a distinct impact on teacher behavior, when compared to states and districts that utilize school-level awards (Jones, 2013). Under performance-pay, teachers in Florida reported a 25% increase in hours worked per week when compared to teachers

in school-level incentive system (Jones, 2013). However, Jones (2013) also found that under either structural design, teachers were more likely to seek employment opportunities outside of teaching profession as a result of performance-pay.

Even with new subject-area exams for previously non-tested subjects, researchers question the measurement's validity and reliability. Mangiante (2011) stated, "the usefulness of any statistical analyses is contingent on the quality of the data that are provided" (p. 48). Researchers have shown that the reliability of value-added models is not limited to situations where the availability of quality data inputs are limited (Goldhaber et al., 2013; Herlihy et al., 2014; Schafer et al., 2012). An analysis by Schafer et al. (2012) of the current VAM literature found instability in scores at both the school- and teacher-levels from one year to the next. The researchers reported a correlation as low as 0.34 for school-level effects and as low as 0.2 at the teacher-level, across consecutive years, raising questions about the fairness of such a measure to inform high-stakes employment decisions (Schafer et al., 2012). The same researchers identified stability concerns across grade, course, test form, and model variations (2012).

Goldhaber et al. (2013) found that variations among models were sufficient to move a teacher's VAM score from one quintile to the next by changing the specifications of the model being used. The researchers also considered between-course variation, and the logic of such a construct, to wit, "a biology course is not necessarily a good proxy for previous achievement in a chemistry course, even if it is the science course that was taken in a prior school year" (Goldhaber et al., 2013). However, even some opponents of VAM concede its potential of becoming a powerful diagnostic tool, greatly benefiting the formative assessment process but are still quick to caution against its use in high-stakes

decision-making (Grossman et al., 2013; Mangiante, 2011; Podgursky & Springer, 2007). According to Schafer et al. (2012), educational leaders are simply not ready to implement such a system.

At this point we do not appear to have models that are so accurate that they can ignore or compensate for the context of the instruction. Indeed, it may be doubtful that effective teaching is a simple construct that is independent of the characteristics of the students or the context of the classroom. (Schafer et al., 2012, p. 4)

Under performance-pay plans, teachers are likely to increase their reward-seeking behavior; not all of which are constructive or positive (Jones, 2013). Researchers have raised the specter of teachers or administrators engaging in potentially negative behaviors to unfairly augment or otherwise manipulate new accountability systems to avoid reprisal or earn reward (Goldhaber & Walch, 2011; Levine & Levine, 2012; Podgursky & Springer, 2007). “Incentive schemes that tie teacher pay to achievement gains by students, either for the individual teacher or the ‘team,’ provide incentives for cheating or other opportunistic behaviors” (Podgursky & Springer, 2007, p. 569). Specifically, an over-classification of students as ESE or ELL, removing students from the testing environment through out-of-school suspensions on test days, grade retention policy manipulation or generous use of test exemptions were all cited as conceivable ways teachers and administrators could try to deceptively offset potentially negative test results (2007).

Proponents of value-added models contend that in order to be used effectively, certain thresholds of data quantity and quality must be met (Knoeppel et al., 2005; Mangiante, 2011). Di Carlo (2012) proffered recommendations to address data concerns while implementing a value-added model, including (a) avoiding mandated high weights

of VAM scores on teacher evaluations: in the early years of implementing VAM, weighting VAM scores as only 10-20% of an evaluation would allow districts to evaluate the accuracy of their model over time without having significantly undue effects on teachers; (b) paying attention to all components of the evaluation system, and effectively designing additional measures of sufficient strength as to offset the actual impact of a teacher's VAM score. E.g.: if scores from principal observations are the same across a school, 100% of the ultimate differentiation between teachers will be due to VAM scores alone; (c) addressing the error by requiring multiple years of data before VAM scores can negatively impact teachers, or for a similar effect, adjusting the weight of VAM scores by sample size to reduce the impact of error in the measurement; and (4) regularly monitor outputs of the evaluation system and analyze them for accuracy and stability (Di Carlo, 2012).

In practice, some researchers have found positive effects of using student performance data in teacher evaluation and compensation system. Goldhaber &Walch (2012) analyzed performance outcomes under Denver Public School's ProComp teacher pay initiative. They found that while existing teachers were given the ability to opt-in to the new plan, there were statistically significant differences in level of performance between teachers who chose to do so and teachers who did not (Goldhaber &Walch, 2012). In their study, students in classrooms with teachers who voluntarily chose to participate in the program had achievement levels 0.4 to 0.7 standard deviations higher than their peers who were taught by teachers outside of the ProComp system (Goldhaber & Walch, 2012). The researchers posit that the act of voluntarily participating in such a program would correlate to a teacher's success under it, possibly indicating the presence

of a selection bias. Selection bias could have long-term implications for the success of performance-pay plans. To wit, “workforce composition effects are potentially quite important. . . . Much of the gains associated with pay for performance in the private sector results from more productive workers sorting into a performance-based system” (Goldhaber & Walch, 2012, p. 1071).

Individuals who are more motivated by high-stakes accountability systems are more likely to perform well under them (Goldhaber & Walch, 2012). The findings supported earlier research by Podgursky and Springer (2007), which considered teacher turnover under performance-pay systems. Their research indicated that teacher turn-over under performance-pay systems would result in less productive teachers leaving the field with more productive teachers staying and being rewarded for their performance (Podgursky & Springer, 2007). The researchers suggested that, when considering the effects of performance pay systems on employee behavior, the systems “will tend to attract and retain individuals who are particularly good at the activity being incentivized and repel those who are not” (Podgursky & Springer, 2007, p. 557).

In addition to a predisposition toward programs that include high-stakes accountability, age and years of teaching experience also correlate to teacher perceptions of new accountability systems. Jones (2013) found that new teachers responded less negatively to performance-pay than did their more experienced peers. Therefore, the results of high-stakes evaluation and compensation systems may improve as older teachers retire and younger teachers take their place (Jones, 2013, p.163).

Even with such mixed empirical evidence, education reform has proceeded with high-stakes systems of evaluation and compensation. Concurrent to NCLB, a 2006

federal program, the Teacher Incentive Fund (TIF), was created to encourage school districts to reform their compensation systems to specifically include performance-pay models (Podgurski & Springer, 2007). The program, created under the Bush administration, was continued and expanded by the Obama administration in 2009 as the Teacher and Leader Innovation Fund (TLIF). The Obama administration planned to leverage the program's resources to "more fairly distribute high-quality teachers among differently resourced schools, improve educator-preparation programs, develop additional professional opportunities for effective teachers, strengthen evaluation systems, remove ineffective teachers, improve professional development, and support school turnaround efforts" (Smarick, 2011, p. 60). As a precursor to RTTT, TLIF grants offered districts a variety of reform options that would be supported by generous federal grant dollars; a system similar to what they would experience under RTTT (Smarick, 2011). TIF and TLIF programs laid the groundwork for subsequent reform efforts, like RTTT, that encouraged school districts to implement value-added models in lieu of traditional evaluation systems. The shift toward VAM is controversial, and its criticism is summarized by Schafer et al (2012).

Besides geography, variables such as socioeconomic status, individual aptitude, home environment, and per-pupil expenditure, while associated with each other, nevertheless may all be needed to represent institutional challenge adequately. Constructing models to incorporate variables such as these and comparing outcomes with programs that have common environments may prove to have more value than VAM. (Schafer et al., 2012, pp. 18-19)

The value-added model adopted by the state of Florida uses two years of prior test data, accounts for multiple student-level characteristics, and holds teacher- and school-level characteristics as random effects (American Institutes of Research, n.d.). The

specific measured characteristics for students used in the Florida model include the following: (a) the number of subject relevant courses in which a student is enrolled (those linked to an FCAT course), (b) two prior years of test scores from the course, (c) disability status, as indicated by a student's receipt of special supports for a disability, (d) status as an English language learner, (e) Gifted status, (f) number of days present in the class over the course of the year, (g) a measure of mobility based on the number of school transitions made during a year, (h) retention, (i) class size, and (j) a measure of homogeneity within the students' test scores from the previous year (AIR, n.d.). It is important to note that even though VAMs were designed with the variable (Sanders & Rivers, 1996) and empirical data plainly demonstrates its impact on student achievement (Ladd, 2012; Lubinski & Crane, 2010; Sirin, 2005), Florida does not include a direct measure of socio-economic status as a predictor variable in the state's value-added model for teacher evaluation (AIR, n.d.).

Observations

Effective instruction that drives student growth involves a variety of elements, many of which are difficult to capture by standardized tests and value-added models alone (Grossman et al., 2013). Many states, particularly those awarded RTTT funds, have implemented multiple measures to evaluate teacher effectiveness including direct observation of teaching practice (Grossman et al., 2013). While VAM opponents often cite the use of peer and supervisor observations as an alternative to value-added models (Di Carlo, 2012), RTTT participants were encouraged to utilize both tools in their proposed evaluation reforms.

Supervisor observations of teachers have been a consistent component of many teacher evaluation systems since the middle of the 20th century (Marzano et al., 2011). As a formative assessment, observations can provide teachers with specific feedback to impact the quality of their instruction (Brown & Crumpler, 2013). However, as is the case with student performance data, supervisor observations have not previously been used to inform high-stakes employment decisions. As a result, researchers have studied supervisor observation protocols to better understand their utility in the evaluation of educators (Brown & Crumpler, 2013; Burnett et al., 2012; Grossman et al., 2013; Podgursky et al., 2007; Praetorius et al., 2013).

Critics of observation practices argue that supervisor observations are challenged by a lack of training for administrators conducting the observations. Researchers have demonstrated that while principals were able to identify their most and least effective teachers through observations, they were less likely to be able to differentiate between teachers in the middle of the performance spectrum (Burnett et al., 2012; Podgursky et al., 2007). Observations that are unable to differentiate levels of teacher quality undermine administrators' ability not only to ensure professional development opportunities are provided in the most effective ways, but further, "one cannot use teacher performance to inform personnel decisions if the performance evaluation systems in place fail to distinguish teachers from one another" (pp. 230-231). Burnett et al. (2012) also found that principal observation scores identified more teachers as highly effective than did outside observers, limiting the effectiveness of the tool at identifying teachers in need of additional supports to improve their practice.

The utility of principal observations of teaching practice are also limited in situations where the observer lacks content-area knowledge being taught, such as in foreign language courses. In situations when an observer's content-area knowledge is lacking, Brown and Crumpler (2013) recommend the use of a peer evaluator, with specific content expertise, as a more effective alternative to traditional supervisor evaluations. The researchers presented the benefits of such a system as being two-fold: "(1) it would help create more accurate and professionally informed assessment of a teacher's performance, and 2) it would serve formative purposes, helping a foreign language teacher improve his or her teaching" (p. 149).

Observations in any form rely upon the proposition that teaching quality is a fixed constant and that the limited amount of time spent on an observation is sufficient to quantify its value (Praetorius et al., 2013). To wit, "the question of whether the quality of the observed lessons is sufficiently indicative of the lessons the teachers generally conduct is crucial" (Praetorius, 2013, p. 2). In the study, the researchers found that the number of observations required to reliably estimate teacher effectiveness across different dimensions of teaching varied considerably depending on the dimension under consideration. Specifically, the researchers found that while a teacher's proficiency with classroom management and personal learning support could be estimated with a reliability coefficient of 0.7 after one visit, estimates of students' cognitive activation during lessons required at least nine visits to achieve the same level of reliability (Praetorius et al., 2013).

Table 4

Summary of Literature Reviewed: Teacher Evaluation

Subsection Summary of Findings	Authors
<p><i>Evolution of Teacher Evaluation.</i> Early evaluation systems included some measure of educational outputs but more recent models place greater emphasis on outputs for high-stakes employment decisions.</p>	<p>Education Commission of the States (2007); Marzano, Frontier, & Livingston (2011); Smarick (2011); Wise, Darling-Hammond, McLaughlin, & Bernstein (1984)</p>
<p><i>Purposes and Perceptions of Teacher Evaluation.</i> Professional development is a shared purpose, but teachers are more critical of new models that are difficult to understand. Younger teachers more receptive to reforms. Multiple measures should be used for an accurate assessment of teacher quality.</p>	<p>Aritomi et al. (2009); Brown & Crumpler (2013); Burnett, Cushing, & Bovina (2012); Devaux et al. (2012); Di Carlo (2012); Gawlik et al. (2010); Goldhaber & Walch (2012); Grossman et al. (2013); Howell, Peterson, & West (2011); Jones (2013); Mangiante (2011); Wise, Darling-Hammond, McLaughlin, & Bernstein (1984)</p>

Table 5

Summary of Literature Reviewed: Teacher Compensation

Subsection Summary of Findings	Authors
<p><i>Evolution of Teacher Compensation.</i> Vast majority of school districts use a salary schedule to compensate teachers. Recent reforms have pushed to change this toward performance-pay systems.</p>	<p>Laine, Potemski, & Rowland (2010); NCES (2012); NGA (2011); Podgursky & Springer (2007); Shields (2012)</p>
<p><i>Purposes and Perceptions of Teacher Compensation.</i> The public remains supportive of teachers and believes they are compensated too little. More support for performance-pay systems to improve accountability and student performance. Administrators are generally more supportive of new models than teachers.</p>	<p>Boser (2012); Dolton & Mercenaro-Gutierrez (2013); Gawlik et al. (2010); Goldhaber & Walch (2012); Jones (2013); Lagana-Riordan & Aguilar (2009); Laine et al. (2010); Okpala (2000); Shields (2012)</p>

Table 6

Summary of Literature Reviewed: Reforms to Teacher Evaluation and Compensation

Subsection Summary of Findings	Authors
<p><i>Reforms to Teacher Evaluation and Compensation.</i> Reforms found to have limited impact on student achievement. Policy decisions do not tend to align with empirical evidence that exists in the field.</p>	<p>Aritomi et al. (2009); Boser (2012); Brown & Crumpler (2013); Burnett et al. (2012); Consortium for Policy Research in Education (2012); Di Carlo (2012); Gallagher et al. (2012); Gawlik et al. (2010); Goldhaber et al. (2013); Goldhaber & Walch (2012); Grossman et al (2013); Herlihy (2014); Jones (2013); Knoeppel et al. (2005); Konstantopoulos (2009); Lagana-Riordan & Aguilar (2009); Levine & Levine (2012); Mangiante (2011); Marzano et al. (2011); Master (2014); Palardy & Rumberger (2008); Praetorius et al. (2013); Podgursky & Springer (2007); Schafer et al. (2012); Smarick (2011)</p>

Summary

The literature reviewed in the preceding chapter established a contextual and research frame for the present study. Over the course of the five years since RTTT was passed, as a component of the ARRA, advocates and opponents of the evaluation and compensation systems proposed therein have sought to validate their views with a mix of evidence and conjecture. To date, insufficient research has been conducted to evaluate the implications of RTTT reforms for the educational leaders at the school and school district levels, who implemented the reforms at the most granular level. “A better understanding of [educational leaders’] perceptions could make a difference in creating policies that will truly impact student achievement in a positive way” (Windish, 2012, p. 72).

The literature reviewed presented evidence of the significant difficulties children growing up in poverty experience in their attempts to obtain an education. Many of the factors associated with low-SES, like limited resources in the home, poor nutrition, exposure to toxins, drug abuse and violence in their neighborhoods, and limited parental engagement in education are outside of the control of teachers and administrators in the schools but are highly predictive of a student’s long-term educational and life outcomes. However, the factors that most negatively impact educational outcomes are not addressed by major policy priorities enacted by state and federal governments.

The last decade of educational reforms, from NCLB to RTTT, has focused on increased accountability for teachers and administrators by way of rigid achievement mandates and high-stakes testing of student learning. Reform strategies have produced mixed results, but researchers have identified improvement to data and evaluations

systems that provide improved access to student performance information and timely feedback as having a positive impact across subjects and schools.

In the chapters that follow, the methodology for the current mixed-methods study will be detailed, the results from the electronic survey presented, and finally the results and implications from the data discussed. The current study bookends one completed by Windish (2012) and presents a longitudinal evaluation of educational policy in practice, as perceived by administrative and instructional personnel in large and small, urban and suburban, school districts in central Florida.

CHAPTER 3 METHODOLOGY

Introduction

The purpose of this study was to enhance the understanding of professional perceptions of administrative and instructional personnel regarding teacher evaluation and compensation components of Race to the Top. Professional perceptions of administrative and instructional personnel enrolled in graduate programs at the target university were sought to facilitate understanding of the effects the teacher evaluation and compensation reforms had on student achievement and growth.

To answer the research questions from the current study, an electronic survey (Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components from RTTT) was implemented (Appendix A). The survey was modified, with permission (Appendix B), from the original study being replicated (Windish, 2012). Windish's study (2012) utilized an electronic survey to collect quantitative data and a follow-up interview to collect qualitative data. Much of Windish's (2012) survey was incorporated into the electronic survey used in the current study with some adjustments, as will be noted in the subsequent sections of this chapter.

Research questions one and two sought to understand the extent to which administrative and instructional personnel perceived the RTTT initiative as being fair or impactful at the end of the survey implementation period (June 30, 2014). The third research question sought to understand how school poverty affected the perceived impact of RTTT by administrative and instructional personnel. The final research question examined any change in perception of the reforms that might exist for administrative and

instructional personnel from prior to implementation of RTTT in 2011, to the date of this study (May-June 2014).

Through this study, the researcher hoped to contribute to the body of knowledge in educational leadership related to teacher evaluation and compensation reforms that lead to improved student achievement. Additionally, the perception of the impact of this reform effort, funded at the federal level, with public tax dollars, may impact public perceptions of future reform efforts and ultimately their ability to be passed into law by state or federal legislative bodies. The methodological processes used in the current study are detailed in four subsections within this chapter: selection of participants (population and sample), survey design, data collection, and analysis.

Population

Administrative and instructional personnel in the field of education in central Florida, who have obtained or are pursuing advanced degrees in education (i.e., Doctorate degree, Education Specialist degree, or Master's degree), comprised the population for this study. Access to the population was gained through the target university's graduate degree program coordinators in both education and educational leadership. The aforementioned educational leaders were selected as the population to be studied because of their efforts to advance their knowledge of education byway of their pursuit of an advanced degree in the field. Leaders in the graduate programs utilized are described as:

Focusing on analyzing and evaluating program effectiveness, reviewing current research, and leading change that is evidenced-based. Professionals who wish to advance their leadership opportunities [that] were currently working as administrators or teachers in elementary and secondary schools as well as other organizations or agencies. (Windish, 2012, pp. 87-88)

The educational leaders are representative of myriad professional classifications within education and from a variety of central Florida school districts (Windish, 2012). The university catalog for the programs described the students enrolled as professionals seeking to expand their leadership opportunities. Enrolled students were working in the field of public education (serving as administrators or teachers), higher education, or in outside business or industry.

Sample

Students enrolled in the target university's graduate-level programs in education and educational leadership were used to represent a convenience sample of the broader population of central Florida administrative and instructional personnel with advanced degrees in the field of education. The targeted graduate programs included the university's M.Ed. in Educational Leadership, Ed.S. in Educational Leadership, Ed.D. in Education, Executive Ed.D. in Educational Leadership, and Modified Core in Educational Leadership. Individuals who recently completed one of the two doctoral programs were also included in the sample. The administrators for each program provided an initial estimate of the number of students enrolled in each program. The initial estimate indicated approximately 392 total students were enrolled in the graduate programs selected (165 in M.Ed. in Educational Leadership, 12 in Ed.S. in Educational Leadership, 110 in Ed.D. in Education, 95 in Executive Ed.D. in Educational Leadership, and 10 in Modified Core in Educational Leadership). The 392 graduate students comprised the sample for the study.

Instrumentation

Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components of Race to the Top

This researcher sought to better understand the current perceptions, as well as any long-term shifts in perceptions, of administrative and instructional personnel as it related to the teacher evaluation and compensation components from RTTT. At the time of this study, the RTTT initiative was concluding and educational leaders at all levels of public education had been interacting with the federal legislation for multiple years. The timing presented a unique opportunity to replicate the original study conducted by Windish (2012) and evaluate current perceptions as well as long-term shifts in perceptions related to the impact and fairness of the legislation.

The current study implemented an electronic survey using Qualtrics® Survey Software. The survey was a replication the one used by Windish (2012) to collect data related to pre-implementation RTTT professional opinions. The electronic survey was modified with permission (Appendix B) from the original in ways that helped to answer the research questions for the current study. The revisions made by this researcher were consistent with research findings by Dillman, Smyth, and Christian (2010) and served primarily to make the instrument more intuitive to respondents, collect more in-depth demographic data that expanded the ways in which the data could be analyzed, and incorporated new survey items to assist in answering the current research questions. The changes made to the survey will be detailed in the following subsections. The original study (Windish, 2012) used a follow-up interview to collect qualitative data. The response rate for the follow-up interview was quite low ($n = 4$), so for this study, the

researcher chose to incorporate open-ended questions into a single, electronic survey in an attempt to improve the response rate for the qualitative items.

Working with experienced professors of Education and Educational Leadership at the University, the current survey was divided into four parts: Part A, Demographic Information; Part B, RTTT Background Information; Part C, RTTT and Student Achievement/Growth; and Part D, Open-Ended Responses. Each element is detailed in the following subsections. For each of the sections enumerated below, content-related validity was obtained through consultation with experienced university professors with subject area expertise in educational research and survey methodology.

Part A: Demographic Information

Part A of the survey included respondent demographic information, including gender and graduate degree program. In addition, the following demographic information was solicited in reference to their current status (at the time of the study), as well as their status prior to implementation of RTTT in 2011: professional classification, school level, employment location's percentage of free and reduced lunch (FRLP), and school district where employed. The data allowed the researcher to analyze any changes in perceptions over time. The original survey by Windish (2012) used two categories to identify professional classification (administrative or instructional). For the current study, the researcher sought to obtain additional information related to professional classification. To that end, categories of professional classification that distinguished between school and school district-based administrative personnel as well as between school and school district-based instructional coaches were included.

Current professional classification was identified using a partially closed question design in order to reduce “the number of items respondents have to consider at once and still collect data for the key items of interest” (Dillman, Smyth, & Christian, 2009, p. 75). A partially closed question design, that allowed a text entry box for response options not provided, was also selected for School Level and School District Where Employed items to similarly ensure all relevant data was collected without overwhelming respondents with potential response options. FRLP data was collected using equally distributed and mutually exclusive categories (Dillman et al., 2009) as follows: 0-24, 25-49, 50-74, and 75-100. Response categories for I Don’t Have Enough Information and Not Applicable were added in the event respondents did not know the FRLP at their school, or were not employed by a school or school district.

Part B: RTTT Background Information

The second part of the electronic survey collected data pertaining to level of knowledge of RTTT, the source of that information, and the perceived fairness of RTTT from the respondents’ knowledge of the teacher evaluation and compensation components. This section also queried the extent to which respondents’ perceptions of RTTT changed from prior to implementation in 2011 to the date of this study, and the extent to which the two components concerning teacher evaluation and compensation improved the quality of public education in the state of Florida.

The first item asked respondents “Where have you received your information on RTTT” (Appendix A), using a multiple-answer, check-all-that-apply format. Sources included: school districts, professional organizations, colleagues, graduate classes, guest

speakers, webinars, state conferences, collective bargaining units, media/news, educational journals/publications, FLDOE, U.S. DOE, email communication from RTTT, email blasts, and Other (Please Specify and a text box were added following the Other category to allow respondents the ability to include an additional source of information beyond those enumerated).

Participants were asked to rate their knowledge of RTTT using the following scale: expert knowledge, great knowledge, moderate knowledge, little knowledge, or no knowledge. To add clarification, two statements were added: under expert knowledge, the clarifier “Can facilitate a seminar on RTTT” was added, and under the no knowledge option, the clarifier “Have not heard of RTTT” was placed.

The next item asked respondents to rate the fairness, based on their knowledge, of RTTT components concerning teacher evaluation and teacher compensation. For each component, respondents used the following scale to record their responses: extremely fair, fair, unfair, extremely unfair, I don’t have enough information, and not applicable. The responses for I Don’t Have Enough Information and Not Applicable were separated from the other responses to align the conceptual and visual midpoints of the scale (Dillman et al., 2009) and ensure that responses are not negatively skewed based on a misalignment in the perceived middle of the scale.

Following, respondents were asked to rate the extent to which their perceptions of RTTT changed from prior to implementation of RTTT in 2011, to the date of the survey, using the following scale: much more favorable, somewhat more favorable, no change, somewhat less favorable, much less favorable, I don’t have enough information, or not

applicable. I Don't Have Enough Information and Not Applicable response options were also separated from the scaled responses, as explained above.

The final item in the section asks respondents to rate the extent to which the new teacher evaluation and compensation models improved the quality of education in the state of Florida. The item used the following 4-point unipolar scale: greatly improved, improved, somewhat improved, and not at all improved. An additional item, I Don't Have Enough Information, was added in the event respondents did not feel sufficiently knowledgeable to provide a response and was also separated from the scale for clarity.

Part C: RTTT and Student Achievement/Growth

Part C of the electronic survey asked respondents to rate the impact of five specific provisions of the RTTT components concerning teacher evaluation and compensation. The five provisions to be rated were: (1) The first 50% of Teacher Evaluation is based on student performance on a Statewide Assessment (VAM). (2) The first 50% of Teacher Evaluation of those who teach a subject or level in which students are not tested is based on school-wide or team performance (VAM). (3) The second 50% of Teacher Evaluation is based on administrator observations of core effective practices and at least one additional metric. (4) Teachers may be able to optionally participate in a separate performance pay scale (Performance Pay). And (5) Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements. Respondents were asked to rate each of the items on the following 5-point bipolar scale: strong positive impact, positive impact, no impact, negative impact, and strong negative

impact. The response option I Don't Have Enough Information was also provided as a separate, off-scale response.

Part D: Open-ended Responses

Part D of the electronic survey contains six questions that solicit information about respondents' interaction with RTTT components concerning teacher evaluation and compensation. The first item, have you been assessed under RTTT's requirements for professional evaluation and compensation? If an affirmative response was given to the above item, respondents were then asked "Do you believe your evaluation was fair?" Both survey items are closed-ended, binary questions that allowed the researcher to compare perceptions of RTTT and personal experience with those components by individual respondents. The final four items are open-ended to allow respondents to provide "thick, rich, descriptive information" (Dillman, Smyth, & Christian, 2009, p.115). The first two open-ended questions are as follows: (1) "How has your professional perception of RTTT's fairness changed from 2011 to today?" (2) "How has your professional perception of RTTT's impact on student achievement/growth changed from 2011 to today?" For the third question, respondents were first presented a qualifying, closed-ended item: "Has your professional classification changed since 2011?" If respondents provide an affirmative response to the item, they were then presented with the first of two open-ended follow-ups: "How has the change in your professional classification impacted your perception of RTTT?" Respondents then proceeded to the second open-ended follow-up: "How does your perception of RTTT compare with other professionals with whom you have had related discussions?" If

respondents provided a negative response to the closed-ended question “Has your professional classification changed since 2011?” they skipped the first follow-up and were only presented the second follow-up. The final open-ended question in the section was as follows: In your experience, how does school poverty relate to teachers’ and administrators’ evaluations under the new performance evaluation system? Consistent with research findings that show the size of response areas effect response length and thematic content (Dillman et al., 2009), under each of the aforementioned open-ended items, respondents were afforded a large response area allowing multiple lines of text for responses.

Data Collection Procedures

Data for this study were collected from May 22-July 30, 2014. The following sections describe the data collection procedures used to collect the qualitative and quantitative data from the electronic survey.

Collection of Quantitative and Qualitative Data

In May of 2014, the program coordinators for the five previously identified graduate degree programs assisted with the data collection by distributing emails containing the informed consent document and a link to the Qualtrics® survey to approximately 392 enrolled students. The program coordinators sent a follow-up email on June 10. The program coordinators used internal distribution lists to reach potential respondents for the survey.

The electronic survey was closed to potential respondents on July 31, 2014. At that time, a total of 142 participants (36.22%) completed the electronic survey. Due to limitations of the email system at the target university, the exact number of students who received the email invitation to participate in the survey is unknown and therefore the calculated response rate of 36.22% may actually be higher. Of the 142 respondents, 49 (34.5%) respondents identified as being enrolled in the Ed.D. in Education program, 46 (32.4%) from the Executive Ed.D. in Educational Leadership program, 4 (2.8%) from the Ed.S. in Educational Leadership program, 34 (23.9%) from the M.Ed. in Education program, and 6 (4.2%) from the Modified Core in Educational Leadership program, and 3 (2.1%) from the Other category.

Data Analysis

A mixed-method approach was taken to analyze the data from this study. This approach included both quantitative and qualitative methodologies, as described in the following sections.

Analysis of Quantitative Data

Responses to the electronic survey were entered into SPSS version 22 using numerical values assigned by the survey program for each response. This researcher chose to recode select values to align numerical values on an intuitive scale, based upon the response options for a given question (i.e., on a unipolar scale, the intuitive bottom of the scale was recoded to have the smallest numerical value while the intuitive top of the scale was given the largest number value for the scale). Values were recoded to facilitate

interpretation of the data by this researcher and did not affect respondents' interactions with the survey. The analysis used to answer each research question is detailed in the following subsections.

Research Question 1

To determine the extent to which, if at all, there was a relationship between administrative and instructional personnel's self-reported knowledge of RTTT and the perceived fairness of RTTT components concerning teacher evaluation and compensation, a Pearson product-moment correlation coefficient was calculated. This analysis was performed separately to evaluate the perceived fairness of the RTTT teacher evaluation process and the perceived fairness of the RTTT compensation requirements.

Research Question 2

To determine the extent to which there was a difference in the perceived impact of RTTT components concerning teacher evaluation and compensation on student achievement and growth, a one-way analysis of variance (ANOVA) was conducted. The five individual RTTT elements from Part C of the electronic survey acted as the dependent variable in the test, while the professional classification data from Part A acted as the independent variable.

Research Question 3

A one-way ANOVA was performed to determine the extent to which there was a difference in the perceptions of administrative and instructional personnel, with different

self-reported school poverty percentages, on the impact of RTTT components on student achievement and growth. To calculate the ANOVA, this researcher used the school poverty percentage categories (i.e., 0-24, 25-49, 50-74, and 75-100) as the independent variable and the perceived impact scores from Part C of the electronic survey as the dependent variable.

Research Question 4

To determine if administrative and instructional personnel's professional perceptions of RTTT components concerning teacher evaluation and compensation have changed over time, this researcher conducted a one-way ANOVA using respondents' professional classifications as the independent variable and their self-reported change in perception of RTTT, as the dependent variable.

Analysis of Qualitative Data

Quantitative data for this study was collected using open-ended items in the electronic survey. "One apparently common purpose for combining qualitative and quantitative methods is to use the results from one method to elaborate, enhance, or illustrate the results from the other" (Greene, Carracelli, & Graham, 1989, p. 266). Each open-ended survey item aligned with the research questions that guided this study and provided additional depth of understanding to the analysis of the quantitative data.

Responses from the open-ended items were organized to allow the researcher to analyze the data and identify common concepts or themes. Survey responses to the open-ended items were compiled and common concepts or themes identified by highlighting

similar or repeating phrases within the responses. The data, in the form of categories or themes, were incorporated into the quantitative data from the closed-ended survey questions to further support the overall research findings.

Summary

In this chapter, the methods and procedures used for the current study were enumerated. The population was described as well as the procedures used to identify the sample. The quantitative and qualitative measures used in answering the four research questions were described and the elements of the electronic survey were also explained. The statistical tests used to analyze the data collected from the electronic survey were described. The results of this analysis are detailed in Chapter 4 of this study and include the narrative responses from open-ended survey questions.

CHAPTER 4 RESULTS

Introduction

The purpose of this study was to enhance the understanding of professional perceptions of administrative and instructional personnel related to RTTT's teacher evaluation and compensation reforms. Results from the electronic survey used to complete this study (SFITECC-RTTT), including qualitative and quantitative data, are reported in the following chapter.

Descriptive statistics for the variables used to analyze the research questions were compiled and the previously identified statistical tests were employed. Analysis of the quantitative data was supplemented by qualitative data collected from open-ended survey items. The qualitative data served to develop a more in-depth understanding of the relationships between the variables being considered.

Descriptive Statistics

The design of the SFITECC-RTTT collected data on numerous variables. The only variables discussed in this section are those that were used in analysis of the research questions for the current study. The categorical variables for this study were professional classification and self-reported free and reduced lunch percentage at the school where respondents worked or interned. The continuous variables related to respondents' perceptions of RTTT and included (a) knowledge of RTTT, (b) perceived fairness of RTTT teacher evaluation reforms, (c) perceived fairness of RTTT compensation reforms, (d) perceived impact of five selected RTTT evaluation and

compensation reform elements on student achievement, and (e) the change in perception of the RTTT reforms from 2011 to the date of this study.

Categorical Variables

The frequencies for professional classification and percentage of free and reduced lunch at respondents' schools were calculated from the survey data. Of the 138 respondents who provided information about their professional classification, 12 (8.6%) indicated their current professional classification as school district-based administrator, 15 (10.9%) as school-based administrator, 62 (44.9%) as instructional, 11 (7.9%) as school district-based instructional coaches, 6 (4.3%) as school-based instructional coaches, and 32 (23.2%) as other. Among the respondents who selected other for their professional classification, 15 identified themselves as college or university faculty and staff, 9 as employed outside the field of education, 4 as other school staff, 2 as graduate students, and 2 as other school-district staff.

For the 138 respondents who provided a response to the question requesting the percentage of free and reduced lunches at the school where employed, 52 (37.7%) indicated either not applicable or they did not have enough information. Of the remaining 86 respondents, 10 (11.6%) indicated 0-24%, 20 (23.3%) indicated 25-49%, 25 (29%) reported 50-74%, and 31 (36%) reported 75-100%.

Continuous Variables

Continuous variables used to analyze the research questions in the current study included self-report values and those assigned by the Qualtrics ® software to represent

specific item values. The continuous (dependent) variables are displayed in Table 7 with the corresponding categorical (independent) variables used to analyze each research question.

Table 7

Research Questions, Independent Variables, and Dependent Variables

Research Question	Independent Variable	Dependent Variables
1	Self-reported knowledge of RTTT	Perceived fairness of RTTT – teacher evaluations; perceived fairness of RTTT – teacher compensation
2	Self-reported professional classification	Perceived impact of RTTT in the five selected areas of teacher evaluation and compensation reforms
3	Self-reported free and reduced lunch population percentage	Perceived impact of RTTT in the five selected areas of teacher evaluation and compensation reforms
4	Self-reported professional classification	Self-reported change in perception of RTTT from 2011 to the date of this study

Of the 132 respondents who rated their level of knowledge of RTTT using the SFITECC-RTTT scale, 9 (6.81%) indicated no knowledge of RTTT, 36 (27.2%) indicated little knowledge, 62 (44.9%) indicated moderate knowledge, 22 (15.9%) reported having great knowledge, and 3 (2.1%) indicated expert knowledge of RTTT. The group mean was 2.80, between little and moderate knowledge.

Based on their knowledge, 104 respondents provided a rating of the perceived fairness of RTTT's teacher evaluation reforms. Among them, 16 (15.3%) rated the teacher evaluation reforms as extremely unfair, 48 (46.2%) as unfair, 38 (36.5%) as fair, and 2 (1.9%) rated the teacher evaluation reforms as extremely fair. The average rating for teacher evaluation reforms was 2.25, slightly above unfair. An additional 30 survey respondents indicated they did not have enough information to rate the fairness of the teacher evaluation reforms. Using the same scale as above, 100 respondents rated the fairness of teacher compensation reforms under RTTT. Of the 100 respondents who rated the compensation reforms under RTTT, 25% ($n = 25$) rated the compensation reforms as extremely unfair, approximately half (52%) as unfair, and 23% rating the reforms as either fair ($n = 21$) or extremely fair ($n = 2$). Similar to that of teacher evaluation reform, the average rating for the perceptions of teacher compensation reform was unfair (2.0).

Table 8 presents the relevant descriptive statistics for the five identified components of RTTT's teacher evaluation and compensation reforms. The component with the highest mean score (3.63) was the component offering teachers at the lowest performing schools recruitment and retention salary enhancements. The component with the lowest mean score (2.28) was that which required 50% of the evaluations for those teaching non-tested subjects, be based upon a school-wide or team VAM scores.

Table 8

Descriptive Statistics for Selected RTTT Components: Impact on Student Achievement

Selected RTTT Component	N	Min	Max	Mean Score	Standard Deviation
1st 50% of Teacher Evaluation is based on student performance (VAM)	104	1	5	2.60	1.075
1st 50% of Teacher Evaluation for non-tested subject or level is based on school-wide or team performance (VAM)	103	1	4	2.28	1.023
2nd 50% of Teacher Evaluation is based on administrator observations of core effective practices and at least one additional metric	104	1	5	3.49	.955
Teachers may be able to optionally participate in a separate performance pay scale	93	1	5	2.89	1.016
Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements	92	1	5	3.63	.969

Survey respondents were asked to rate the extent to which their perceptions of RTTT have changed, from prior to implementation in 2011, to the date of this study. Summary statistics, by Professional Classification, are presented in Table 9. The average change reported for all respondents (2.56) indicated a slightly less favorable view of RTTT at the time of this study, compared to 2011. School-district staff (administrators and instructional coaches) had a mean of 2.73, close to No Change, while the mean rating for Instructional staff was 2.36; the lowest average rating among the professional classifications. School-based Instructional Coaches reported the highest mean score (3.33) for change in perception of RTTT, but due to the small number of respondents in the category (N = 6) caution should be taken when interpreting that result.

Table 9

Descriptive Statistics: Change in Perceptions of RTTT from 2011 to the Date of this Study, by Professional Classification

Professional Classification	N	Mean	Std. Deviation
School District-based Administrator	11	2.73	.905
School-based Administrator (principal, assistant principal, dean)	13	2.54	.967
Instructional (classroom teacher, counselor, dean, specialist)	47	2.36	1.241
School District-based Instructional Coach	11	2.73	1.191
School-based Instructional Coach	6	3.33	.816
Other (Please Specify)	14	2.64	.745
Total	102	2.56	1.095

Data Analysis for Research Question 1

Research Question 1: To what extent, if any, is there a relationship between administrative and instructional personnel's self-reported knowledge of RTTT and the perceived fairness of RTTT requirements concerning teacher evaluation and compensation?

Quantitative Analysis for Research Question 1

Analysis of Research Question one was completed using two Pearson product-moment correlation coefficients. While not intended to identify causality, a correlation coefficient will show the strength, and direction, of a relationship between variables. Only one statistically significant relationship was identified between the two tests.

The first test to identify a correlation considered the relationship between self-reported knowledge of RTTT and perceived fairness of teacher evaluation reforms under RTTT. The Pearson correlation was not found to be statistically significant.

$$r(103) = .091, p > .05$$

This indicated that respondents' self-reported knowledge of RTTT was not related to their perceived fairness of the initiative's teacher evaluation reforms.

The second test to identify a correlation considered the relationship between self-report knowledge of RTTT and perceived fairness of teacher compensation reforms under RTTT. For the second test, the Pearson correlation coefficient was found to be statistically significant.

$$r(99) = .240, p < .05$$

This result indicated that the two variables were positively related, to wit, as respondents' self-reported knowledge of RTTT increased, so too did their rating of the fairness of teacher compensation reforms.

Qualitative Analysis for Research Question 1

Qualitative data were collected to enhance understanding of the quantitative data used in answering research question one. Respondents were asked the following question, “How has your professional perception of RTTT's FAIRNESS changed from 2011 to today?” Responses ($n = 94$, 68%) were grouped by content or phrasing. The two most common themes or categories to emerge from the open-ended responses, as factors impacting perceptions of fairness, were related to the use of value-added models as a component of evaluation and compensation systems ($n = 57$, 60.6%) and problems associated with the implementation and dissemination of program requirements and specifications ($n = 19$, 20.2%).

Use of Value-Added Models

The first major theme centered on the inclusion of a value-added model in the teacher evaluation and compensation reforms. Respondents expressed concerns related to multiple components of VAM, the most frequent of which are discussed in the following section.

First, the validity and reliability of the standardized tests which inform student growth used in the models were of concern to respondents (10.5%), with one stating, “Basing teacher salaries on tests that have not been tested and proven reliable and valid is inherently unfair”. Along the same line, another respondent provided the following insight,

I know teachers that have had a very high instructional practice score and a very low VAM score. I know teachers that have very low instructional [practice] scores and very high VAM [scores]. In the beginning, I believed that good

teachers would have VAM scores that reinforced their efforts in the classrooms, but my perception of fairness has changed based on my experience.

Second, the use of school-wide VAM scores was a frequent concern for respondents (17.5%) where they are used to evaluate individuals who teach content areas lacking historical student performance data, or individuals who oversee multiple school sites (school district-based personnel). One respondent stated their position, “I am exposed to more situations where teachers are being evaluated based on students they have never taught; in addition, some are being evaluated by the standardized test results which are produced from content areas other than what they teach”. Another respondent provided a similar insight, “I left the classroom, where I had direct impact on my VAM score, to [go to] a district position, where the VAM scores are based on multiple schools and many students who I have little to no impact on”.

The final sub-theme involved the design of the value-added model and the student-level demographic factors that it accounts for, specifically socio-economic status. Multiple respondents (8.8%) indicated that the lack of a direct measure of socioeconomic status impacted their perception of the fairness of the value-added model used in their district. One respondent stated, “There are too many variables that can impact a student and their performance (socioeconomic status, domestic factors, coming into class behind to begin with, etc.)”. The same idea was conveyed by another respondent, stating that the current model used in their school district “...doesn't take into account... different demographics of students (home life, socioeconomic status, etc.) from teacher to teacher”.

Implementation and Communication Issues

The second major theme identified from the qualitative data involved the implementation of RTTT reforms by state and local educational agencies and the communication of key components of the reforms to stakeholders. In this area, respondents expressed specific concerns related to reform implementation (26.3%) as well as a lack of understanding of local and state reform efforts (73.7%).

One respondent, mentioning implementation concerns as negatively impacting perceptions of RTTT's fairness, stated "The lack of systemic support from FLDOE has negatively impacted my perception of RTTT's fairness [and] because of the variations in application of the key concepts (teacher evaluation and performance pay)". Another respondent expressed a similar sentiment by stating, "It is my opinion that state DOE and local districts have mishandled the creation of course assessments for non-core classes". Only one respondent provided a positive comment related to the implementation of RTTT reforms stating,

The district continues to prepare end of course exams for every course code in use in the district. These tests will serve as the instrument to judge teacher effectiveness. This should be an improvement over the current model where teachers of courses other than reading and math have school average scores factored into their VAM score.

The component of RTTT implementation regarding new testing requirements was only acknowledged by two respondents, and aligns with the second sub-theme identified by this researcher; a lack of understanding of reform specifications by stakeholders most impacted by the reforms.

A number of respondents to the survey question ($n = 14$, 11.76%) indicated a lack of knowledge surrounding RTTT reform efforts. One respondent provided the following

comment, “I don't know much about RTTT. Marzano dominates discussion regarding student and teacher success/pay”, while another respondent confessed, “I really don't know enough about what is district policy and what is RTTT initiative to know the difference”. After three years of implementation and professional development, the lack of understanding of RTTT reforms by survey respondents (45% of whom were classroom teachers) should be of concern to school district personnel.

Data Analysis for Research Question 2

Research Question 2: To what extent, if any, is there a difference between administrative and instructional personnel’s perceptions of the impact of RTTT teacher evaluation and compensation components on student achievement/growth?

Quantitative Analysis for Research Question 2

To answer research question two, a one-way ANOVA was performed to analyze the mean difference between respondents’ professional classification and their perceptions of the impact of RTTT reforms on student achievement/growth. The results of this analysis are presented in Table 10.

Respondents did not vary in their perceptions of the impact of RTTT reforms based upon professional classification, inasmuch as no statistically significant relationships ($p < .05$) were identified between the professional classifications and the perceived impact scores on the selected RTTT components. Therefore, no further statistical analysis was necessary between the means of the professional classification groups. That being said, the component closest to being identified as statistically significant ($p = .125$) from this test was the component that read, “1st 50% of Teacher Evaluation for non-tested subjects or levels is based on school-wide or team performance

(VAM)”. This may indicate a modest divergence in perceptions of the efficacy of this type of reform at improving student learning among the different professional classification groups. However, this component reported the lowest mean impact score ($\mu = 2.28$) of the five RTTT components, indicating that while there is greater variation in perceptions of this component, the general consensus as to its impact is still quite low.

Table 10

Perceived Impact of Components of RTTT on Student Achievement, by Professional

Classification

Selected RTTT Components	Sum of Squares	df	Mean Square	F	Sig.
1st 50% of Teacher Evaluation is based on student performance (VAM)					
Between Groups	6.360	5	1.272	1.106	.362
Within Groups	112.678	98	1.150		
Total	119.038	103			
1st 50% of Teacher Evaluation for non-tested subjects or levels is based on school-wide or team scores (VAM)					
Between Groups	8.966	5	1.793	1.777	.125
Within Groups	97.869	97	1.009		
Total	106.835	102			
2nd 50% of Teacher Evaluation is based on administrator observations of core effective practices and at least one additional metric.					
Between Groups	3.685	5	.737	.800	.552
Within Groups	90.306	98	.921		
Total	93.990	103			
Teachers may be able to optionally participate in a separate performance pay scale.					
Between Groups	2.695	5	.539	.508	.769
Within Groups	92.230	87	1.060		
Total	94.925	92			
Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements					
Between Groups	2.686	5	.537	.558	.732
Within Groups	82.748	86	.962		
Total	85.435	91			

Qualitative Analysis for Research Question 2

To expand the utility of the quantitative data collected to answer research question two, survey respondents were asked the following open-ended question, “How has your professional perception of RTTT's IMPACT on student achievement/growth changed from 2011 to today?”. Of the responses to the survey item ($N = 84$), only seven respondents (8%) indicated a more positive perception of RTTT’s impact at the time of the current study, than in 2011. The majority of responses reported a more negative perception ($n = 46$, 55%) of RTTT’s impact on student achievement/growth than in 2011. Eighteen respondents (21%) indicated there was no change in their perceptions, while 13 (15%) indicated that they did not know.

Among the positive responses ($n = 7$), 71.4% identified their current professional classification as either an administrator (school district-based and school-based) or a school district-based instructional coach, while 28.6% identified their professional classification as instructional. For the negative responses ($n = 46$), a plurality of respondents (45.6%) identified their professional classification as instructional, 23.8% as administrative (school- and school district-based), 13% as school district-based instructional coaches, 6.5% as school-based instructional coaches, and 8.7% as other.

Respondents indicating a more positive perception of the impact of RTTT mentioned clarity of expectations ($n = 3$), improved decision-making based upon student data ($n = 3$), and greater accountability for students and teachers ($n = 4$) as factors positively affecting their perceptions of RTTT. According to one respondent,

The evaluation system used in RTTT has impacted teachers in making them look at and understand the data of their students. It has switched the paradigm of how

teachers [teach] from ‘covering’ material to data based decision-making based on the needs of the students.

Another respondent echoed a similar sentiment, “My perception has shifted from intuition and teacher professional discretion to the importance of using data to drive instruction and differentiate [instruction] based upon multiple factors”.

The majority of responses (55%) indicated having a more negative perception of RTTT’s impact, than in 2011. The main themes found in these responses included teachers teaching toward standardized test content rather than for mastery of the curriculum ($n = 20$), increased student and teacher stress/anxiety ($n = 9$), and VAM as an unreliable measure of quality teaching ($n = 4$).

Respondents expressed concern over the practice of teachers spending significant instructional time preparing students to take standardized assessments. “Our students are learning for the test only, rather than experiencing learning through authentic project-based activities. They are ‘learning’ for the short term rather than the long term” stated one respondent. Along the same line of thought, another respondent stated,

I see that RTTT is making our teaching less focused on critical thinking, as many levels of higher level thought cannot be assessed with a multiple-choice test and therefore teachers focus more on what will be assessed by FCAT than what’s best for kids.

A discussion of the added stress or anxiety that standardized tests elicited for teachers and students was a second sub-theme among the negative responses. One respondent stated, simply, “It has only added more stress and anxiety to teachers, students, and administrators.” Another respondent provided a similar response by stating, “I see the stress that these assessments place on teachers and take away from teaching and planning time. I think that the process is negatively affecting student growth”.

The third sub-theme from this survey item involved negative perceptions of the use of a value-added model. One respondent provided the following assessment of VAM's impact,

In the beginning, I believed the movement had great promise to improve student achievement because of the focus on growth in performance and the VAM formula's 'leveling the playing field.' However, because there is such a delay in getting the VAM scores, the emphasis on using the scores to improve teaching and learning has been lost. It has become a summative measure, post-mortem, and I don't think it's changing instructional practice, which is the first step in improving student achievement.

Further, respondents expressed concerns surrounding specific elements of value-added models. One respondent questioned the wisdom of school-wide VAM scores being used in lieu of individual teacher scores where historical data was not available, "How does that identify ineffective teachers? How are they evaluated on students they don't even teach? There is no direct, correlational relationship between those teachers and impact on student achievement." Expanding further, the same respondent added, "In theory, I can see how legislators would believe teacher evaluations tied to student achievement should have a positive correlation but there are just too many variables that decrease the validity of the VAM scores."

Data Analysis for Research Question 3

Research Question 3: To what extent, if any, is there a difference in the perceptions of administrative and instructional personnel who have different self-reported school poverty percentages about the impact of RTTT teacher evaluation and compensation components on student achievement/growth?

Quantitative Analysis for Research Question 3

For research question three, a one-way ANOVA was performed to analyze the mean difference between respondents' self-reported free and reduced lunch percentage (FRL) and their perceptions of the impact of RTTT teacher evaluation and compensation components. The results of this analysis are presented in Table 11.

Respondents did not differ in their perceptions of the impact of RTTT teacher evaluation and compensation reforms by their schools' free and reduced lunch percentage. From the ANOVA, no statistically significant ($\alpha < .05$) relationships were found between the selected RTTT components and the schools' poverty measure (FRL), therefore no additional statistical analysis is necessary. However, similar to the quantitative findings from research question two, the second RTTT component ("1st 50% of Teacher Evaluation for non-tested subjects or levels is based on school-wide or team performance (VAM)") was the closest to being identified as having a statistically significant ($p = .391$) difference between FRL categories. The lack of statistical significance, as well as the low mean rating (2.28) by survey respondents indicates a general negative perception of this type of evaluation system, regardless of the socio-demographics of the students served; echoing the findings enumerated in the quantitative findings for research question two.

Table 11

Perceived Impact of Components of RTTT on Student Achievement, by Respondents'

FRL Group (0-24, 25-49, 50-74, 75-100)

	Sum of Squares	df	Mean Square	F	Sig.
1st 50% of Teacher Evaluation is based on student performance (VAM).					
Between Groups	4.784	5	.957	.821	.538
Within Groups	114.255	98	1.166		
Total	119.038	103			
1st 50% of Teacher Evaluation for non-tested subjects or levels is based on school-wide or team performance (VAM).					
Between Groups	5.500	5	1.100	1.053	.391
Within Groups	101.335	97	1.045		
Total	106.835	102			
2nd 50% of Teacher Evaluation is based on administrator observations of core effective practice and at least one additional metric.					
Between Groups	1.920	5	.384	.409	.842
Within Groups	92.071	98	.939		
Total	93.990	103			
Teachers may be able to optionally participate in a separate performance pay scale.					
Between Groups	3.976	5	.795	.761	.581
Within Groups	90.949	87	1.045		
Total	94.925	92			
Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements.					
Between Groups	.653	5	.131	.133	.984
Within Groups	84.781	86	.986		
Total	85.435	91			

Qualitative Analysis for Research Question 3

Qualitative data to assist with answering research question three was collected from survey respondents who answered the question “In your experience, how does school poverty relate to teachers' and administrators' evaluations under the new performance evaluation system”. From the responses (N = 74, 52.1%), three major themes emerged: 1) Poverty has a negative effect on students' ability to learn ($n = 35$), 2) teachers in low-SES schools face greater challenges than do their peers in high-SES schools ($n = 20$), and 3) the value-added model does not effectively account for the effects of poverty ($n = 16$).

Effects of Poverty on Students

The most common theme found in the responses to the open-ended question involved the negative developmental effects of poverty on students. Within that theme, two sub-themes were also identified: limited access to resources in the home (14.4%) and a lack of prior knowledge and lower proficiency that inhibits growth (20.6%).

Limited access to resources, including technology, and supports from home was frequently mentioned by respondents.

The research tells us these kids are often not read to and enter school with severe vocabulary deficiencies or come to school hungry. Children in poverty often times do not receive the support they need at home or have the responsibilities of taking care of siblings [or other] situations that can lead to excessive absences.

This is a major difference between low- and high-SES students, as one respondent stated, “Poor children do not have the support at home to improve like a middle class or upper class student”. Other respondents spoke to an “emphasis on survival”, “excessive absences”, and a failure to meet students' “basic needs” as factors that impair low-SES

students' ability to achieve in school; a factor that has a compounding effect year-over-year, placing them farther and farther behind their peers as they age.

Effects of Student Poverty on Teachers

The second theme identified in the responses involved the effects poverty has on classroom teachers of low-SES students. One respondent provided the following description of the issue,

At first I did not think that it would be a significant impact until I was moved to a high poverty school from an affluent school and saw more challenges and priorities fight for my attention versus what I had to focus on in the more affluent school.

Respondents ($n = 4$) also referenced the idea that lower-SES students do not perform as well as their higher-SES peers on standardized assessments, which in turn impacts a teacher's value-added scores. One respondent stated, "Although the evaluation system is based on student growth, students from high poverty schools face a slew of different challenges that make achieving those gains much harder". This results in a system where "teachers and administrators in [low-SES schools] are held to the same standards and suffer as a result of their students' difficulty in performing".

The selection effect was also mentioned by respondents ($n = 3$) as having the impact of, "encourage[ing] teachers to leave schools of high poverty because of the difficulty those schools have in accelerating student growth on a consistent basis."

Effects of Socio-Economic Status on VAM Scores

The final theme from the responses involved the effect socio-economic status has on VAM scores. One respondent connected the issues of poverty and VAM scores by stating,

Poverty definitely has an effect, but the current system does not account for SES. Low SES students usually have lower parental involvement, reduced access to technology, reduced access to print, and in some cases higher instances of absences and behavior problems. All of this affects student achievement.

In the State of Florida, the adopted value-added model does not utilize a direct measure of socio-economic status. This fact that was not missed by educational leaders, “Zip codes can predict, usually, how well students will perform on high-stakes tests.... It is unfair that socioeconomic status is not a factor in VAM”.

Another respondent approached the question from a holistic perspective, “VAM measuring growth in essence is a good thing. However it lacks total understanding of the whole human. Poverty directly impacts a whole child and therefore adds to the myriad of factors that play with educating ANY child”.

Data Analysis for Research Question 4

Research Question 4: To what extent, if any, have administrative and instructional personnel changed in their perceptions of RTTT evaluation and compensation components, from the time RTTT was first implemented to the date of this study?

Quantitative Analysis for Research Question 4

To answer research question four, a one-way ANOVA was performed to analyze the mean difference between respondents' professional classification and their self-reported change in perception of RTTT from 2011 to the date of this study. The ANOVA results for this analysis are presented in table 12.

The mean difference between the professional classification groups and their self-reported change in perception from 2012 to 2014 was not found to be statistically significant ($\alpha < .05$), to wit, the respondents in each of the professional classification groups reported similar changes, on average, in their perceptions of RTTT from 2012 to 2014.

Table 12

Change in Perceptions of the Impact of RTTT, by Professional Classification

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.154	5	1.231	1.028	.406
Within Groups	114.993	96	1.198		
Total	121.147	101			

Qualitative Analysis for Research Question 4

To consider research question four with additional depth, survey respondents who indicated their professional classification had changed since 2011 were presented with the question “How has your change in professional classification impacted your perception of RTTT?” For the survey item, qualitative data was collected from 20 respondents.

Half of all responses to the survey item ($n = 10$) indicated their change in professional classification did not impact their perceptions of RTTT. Twenty-five percent of respondents ($n = 5$) indicated that the change in their professional classification negatively affected their perceptions of RTTT. And 15% of respondents ($n = 3$) expressed having a broader view of RTTT’s impact after their change in professional classification, where one respondent stated,

My change in professional classification has allowed me to see the impact of RTTT from a more global perspective and see the impact on many more teachers, students and schools. My change in professional classification has also allowed me to gain greater understanding of the reach of my impact as well as how limited my reach is on many students, whose achievement still impacts my evaluation. This has led me to have less faith in the RTTT as it [is] currently.

Additionally, one respondent (5%) indicated a mixed effect on their professional perceptions of RTTT. Of the 20 respondents to this item, only one (5%) indicated that their change in professional classification positively impacted their perceptions of RTTT, simply stating, “I have learned more about high quality teaching”.

Additional Analysis

As additional analysis, this researcher compared perceptions of RTTT’s impact between the Doctor of Education and the Executive Doctor of Education, in Educational

Leadership programs. A similar analysis was performed by Windish (2012), and replicating the analysis allowed for a pre-, post- analysis to evaluate if a statistically significant relationship existed between members of the Education Doctoral (non-administrative leaders) and Educational Leadership Doctoral (executive-level leaders) programs. Descriptive statistics from this analysis are presented in Table 13, and results from the independent-samples T-Test are presented in Table 14.

The results of the independent samples t-tests found a statistically significant relationship ($p = .041$) between the means of the two groups on only one RTTT component; the use of school-wide VAM scores in subjects or levels where standardized tests were not historically performed. This finding showed that respondents in the two doctoral degree programs had similar perceptions of the impact of the selected RTTT reforms, with the exception of the school-wide VAM provision, wherein Education Doctoral students' mean rating indicated a significantly lower perceived impact on that element than was reported by Educational Leadership Doctoral students. This result was significantly different than the results found by Windish (2012) where a statistically significant ($\alpha < .05$) relationship was identified between all of the selected RTTT components. Windish's findings indicated that prior to implementation of RTTT, administrative (Educational Leadership Doctoral students) and non-administrative (Education Doctoral students) personnel professed significantly different perceptions of the potential impact of the teacher evaluation and compensation reforms; with non-administrative personnel rating their perceptions of the potential impact much lower than respondents in administrative positions.

Table 13

Perceived Impact of Components of RTTT on Student Achievement, by Doctoral Degree

Program

	Doctoral Degree Program	N	Mean	Std. Deviation	Std. Error Mean
IMPACT -1st 50% of Teacher Eval based on student performance (VAM).	Ed.D. in Education	33	2.61	1.116	.194
	Executive Ed.D. in Educational Leadership	38	2.87	1.018	.165
IMPACT -1st 50% of Teacher Eval (non-tested areas) based on school-wide or team performance (VAM).	Ed.D. in Education	32	2.09	.893	.158
	Executive Ed.D. in Educational Leadership	38	2.58	1.030	.167
IMPACT -2nd 50% of Teacher Eval based on administrator observations and at least one additional metric.	Ed.D. in Education	32	3.34	1.066	.188
	Executive Ed.D. in Educational Leadership	39	3.59	.818	.131
IMPACT -Performance pay scale (Performance Pay).	Ed.D. in Education	26	3.69	1.192	.234
	Executive Ed.D. in Educational Leadership	36	3.03	.845	.141
IMPACT -Recruitment and retention salary enhancements at low-performing schools.	Ed.D. in Education	26	3.46	1.029	.202
	Executive Ed.D. in Educational Leadership	36	3.75	.874	.146

Table 14

Independent Samples T-Test: Perceived Impact of RTTT Components, by Doctoral

Degree Program

		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
IMPACT -1st 50% of Teacher Eval based on student performance (VAM).	Equal variances assumed	-1.036	69	.304	-.262	.253
IMPACT -1st 50% of Teacher Eval (non-tested areas) based on school-wide or team performance (VAM).	Equal variances assumed	-2.085	68	.041	-.485	.233
IMPACT -2nd 50% of Teacher Eval based on administrator observations and at least one additional metric.	Equal variances assumed	-1.100	69	.275	-.246	.224
IMPACT - Performance pay scale (Performance Pay).	Equal variances assumed	-1.298	60	.199	-.335	.258
IMPACT - Recruitment and retention salary enhancements at low-performing schools.	Equal variances assumed	-1.190	60	.239	-.288	.242

Following this analysis, Table 15 compares the aggregate mean scores of RTTT components reported by Windish (2012) to those compiled in the current study. The mean scores shown were compiled from survey items asking respondents to rate the

impact of selected RTTT elements using a 5-point bipolar scale, where a score of 1 corresponded to strong negative impact and a score of 5 corresponded to strong positive impact. As shown, the difference in overall mean scores decreased from 2012 to 2014 on four of the five RTTT components. The component providing for recruitment and retention salary enhancements for teachers at the lowest performing schools was the only component to receive a higher mean rating of its perceived impact on student achievement in 2014 than in 2012. Overall, respondents perceived the majority of selected RTTT components as being less impactful at the time of the current study than in 2011.

In totum, the data in Table 13 and Table 15 indicate that not only did administrative and instructional personnel have more similar perceptions of the impact of RTTT at the time of this study, but those perceptions were also less favorable on four out of five RTTT components than they were prior to implementation (Windish, 2012).

Table 15

Descriptive Statistics: 2012-2014 Comparison of RTTT Components' Mean Ratings of Impact on Student Achievement

Selected RTTT Component	2012 N*	2012 Mean Score*	2014 N	2014 Mean Score	Difference in Means
1st 50% of Teacher Evaluation is based on student performance (VAM)	52	3.02	104	2.60	-0.42
1st 50% of Teacher Evaluation for non-tested subject or level is based on school-wide or team performance (VAM)	52	2.88	103	2.28	-0.6
2nd 50% of Teacher Evaluation is based on administrator observations of core effective practices and at least one additional metric	48	3.79	104	3.49	-0.3
Teachers may be able to optionally participate in a separate performance pay scale	46	3.15	93	2.89	-0.26
Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements	43	2.93	92	3.63	0.7

*(Windish, 2012, p. 107)

The final tables display the descriptive statistics (Table 16) and Independent Samples T-Test results (Table 17) comparing self-reported knowledge of RTTT to the school district-based administrator and instructional professional classifications. Using a 5-point scale, where the number 1 corresponded to no knowledge, 2 to little knowledge, 3 to moderate knowledge, 4 to great knowledge, and 5 to expert knowledge, the mean rating for RTTT knowledge for school district-based administrators was 3.67, close to

Great Knowledge, while the mean rating for the instructional personnel was 2.72, close to Moderate Knowledge. The mean difference between the two professional classification groups was shown to be statistically significant ($p = .000$), that is to say, school district-based administrators' level of knowledge of RTTT was not shared by instructional personnel at the school level.

Table 16

Descriptive Statistics: Self-reported Knowledge of RTTT for School District-based Administrators and Instructional Personnel

Current Professional Classification	N	Mean	Std. Deviation	Std. Error Mean
School District-based Administrator	12	3.67	.778	.225
Instructional (classroom teacher, counselor, dean, specialist)	60	2.72	.783	.101

Table 17

Independent Samples T-Test: Self-reported Knowledge of RTTT, by Professional Classification (School District-based Administrators and Instructional Personnel)

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Rate your knowledge of Race to the Top using the following scale:	Equal variances assumed	.001	.973	3.840	70	.000	.950	.247	.457	1.443
	Equal variances not assumed			3.855	15.78	.001	.950	.246	.427	1.473

Summary

In this chapter, the procedures used to collect the quantitative and qualitative data for this mixed-methods study were reviewed. Next, the descriptive statistics for categorical and continuous variables used in the analysis of the research questions were presented. Finally, for each research question, analyses of the quantitative and qualitative data were presented. Additional analysis of data collected from the electronic surveys was discussed to highlight interesting relationships not considered in the research questions.

CHAPTER 5 SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

Introduction

Within this chapter is a restatement of the purpose of this study, as well as an overview of the study's population, research design, and instrumentation. The remainder of this chapter presents a summary of the research findings from the electronic survey and a discussion organized by research question, the implications for policy and practice, as well as recommendations for future research.

Purpose of the Study

The purpose of this replication study was to analyze the professional perceptions of educational leaders (instructional and administrative) regarding components of the RTTT initiative involving teacher evaluation and compensation, and the perceived effect the elements had on the academic performance of students. To collect the data for this study, electronic surveys were distributed to central Florida educational leaders pursuing advanced degrees in education and educational leadership. Surveying educational leaders, who work in different types of school settings and educational organizations, may help to understand their perceptions of the fairness and impact of different components of RTTT. Through this study, this researcher hoped to add to the body of knowledge that exists concerning the types of policy reforms that facilitate improved student learning in public schools.

Population, Research Design, and Instrumentation

A convenience sample of educational leaders pursuing Master's of Education, Education Specialist, or Education Doctorate degrees from a target university were selected to represent the population for this study. Using this sample of the broader population, a mixed-methods study was conducted to analyze the professional perceptions of educational leaders surrounding specific RTTT components.

This researcher modified, with permission (Appendix B), the survey used by Windish (2012) in the study this research replicated. The final survey (Appendix A) used for this research (SFITECC-RTTT) was distributed to graduate and doctoral level students at a target university during the 2014 summer semester. Of the 138 respondents that provided their professional classifications, 12 (8.7%) identified their roles as school district-based administrator, 15 (10.7%) as school-based administrator, 62 (45.0%) as instructional personnel, 11 (8.0%) as school district-based instructional coaches, and 6 (4.3%) as school-based instructional coaches; 32 (23.2%) respondents indicated their professional classification as other. The breakdown of professional classifications identified in this study aligned closely with those found by Windish (2012); respondents to the original study were 50% ($n = 27$) administrative and 50% ($n = 27$) instructional. Respondents to the current study provided both qualitative and quantitative data by completing the electronic survey and the data was compiled and analyzed by this researcher to present and discuss in the following chapter.

Summary and Discussion of the Findings

In this section the findings from the electronic survey, related to each research question, are discussed. Quantitative and qualitative data from the survey are presented and the extent to which those findings were, or were not, in alignment with other researchers, including the Windish (2012) study upon which this research is based, are discussed.

Research Question 1

To what extent, if any, is there a relationship between administrative and instructional personnel's self-reported knowledge of RTTT and the perceived fairness of RTTT requirements concerning teacher evaluation and compensation?

The findings from the quantitative analysis conducted using a pair of Pearson product-moment correlation coefficients did not indicate a statistically significant relationship existed between respondents' self-reported knowledge of RTTT components concerning teacher evaluation and compensation and the perceived fairness of the components. The quantitative findings were supported by the qualitative data that showed largely homogenous perceptions of the fairness of the RTTT reforms, regardless of knowledge level of the RTTT components. In addition to the aforementioned analysis, this researcher conducted an Independent Samples T-Test to determine if a statistically significant relationship existed between two specific professional classifications (school district-based administrators and instructional personnel) and level of knowledge of RTTT. The results of this analysis are presented in Table 16 and Table 17 and show a strong, statistically significant difference ($p = .000$) between school district-based administrators and instructional personnel as to their knowledge of RTTT reforms.

Administrators ($\mu = 3.67$) reported having a much greater level of knowledge of RTTT than did instructional personnel ($\mu = 2.72$). This finding indicated a breakdown in communication existed between the school district- and school-levels, where pertinent information related to RTTT implementation would otherwise have been shared. Qualitative findings supported this proposition and presented two main themes.

The first theme identified from the qualitative data involved the use of a value-added model to determine evaluation scores and compensation increases. VAM represented the most common concern related to fairness espoused by respondents. The first VAM-related concern surrounded the validity and reliability of the testing instruments used to calculate VAM scores. Without extensively testing the assessments used, the validity and reliability of the data the instruments collect cannot be assured; a concern that is repeatedly acknowledged in the literature (Goldhaber et al., 2013; Herlihy et al., 2014; Schafer et al., 2012). Additionally, respondents repeatedly mentioned the use of school-wide VAM scores as negatively impacting their perceptions of RTTT's fairness. Research has shown the use of school-wide VAM scores to have a negative impact on teacher's time investment in their work, while increasing turnover (Jones, 2013). This study's findings of a negative perception of school-wide VAM scores are therefore not surprising.

Finally, the inclusion, or omission, of the right student-level moderating variables (specifically SES) was also an issue for respondents concerned with the value-added model. This concern is validated by the literature in this area. Research by Goldhaber et al. (2013) found that the construction of a particular value-added model, and the variables

it considers, could have significant effects on a teacher's final VAM score and potentially a negative impact on compensation or retention with a school district.

The second major theme surrounded implementation and communication issues related to RTTT reforms. Similar to a finding by Windish (2012), initiative communication has been, and continues to be, an issue for educational leaders in the field. Many respondents expressed uncertainty or a lack of clarity related to RTTT-specific reforms that were being implemented in their districts. This indicated that state and federal education policy makers were not being effective in disseminating critical information to educators at the school-level who are most immediately affected by the reforms.

Research Question 2

To what extent, if any, is there a difference between administrative and instructional personnel's perceptions of the impact of RTTT teacher evaluation and compensation components on student achievement/growth?

The quantitative findings from the one-way ANOVA failed to identify any statically significant differences between the professional classifications and perceptions of the impact of the RTTT components concerning teacher evaluation and compensation. Results from the qualitative analysis found three major themes.

A statistically significant relationship was not found between professional classifications and their perception of the impact on student achievement/growth for any of the five specified RTTT components. The findings from this study were different from those in Windish's (2012) study, where a statistically significant difference between administrative and instructional personnel was found between three of the selected RTTT

components. For the components requiring 50% of a teacher's evaluation to be based on student performance on standardized assessments, 50% of a teacher's evaluation in non-tested subjects be based on school-wide or team scores, and a separate pay scale for teachers at the lowest performing schools, the mean difference between administrative and instructional personnel was sufficiently large to be identified as statistically significant (Windish, 2012). However, in the current study the mean difference between respondents, based on professional classification, for each of the same three RTTT components was not statistically significant. This finding indicated that educational leaders, at all levels, had developed more similar perceptions of the impact of RTTT since 2011. Considering that, when comparing the mean scores for each of the selected RTTT components from the current study, to those found in the Windish study (Table 15), all but one of the selected RTTT components in the current study received a lower mean score. This finding indicated that not only were educational leaders more unified in their perceptions but those perceptions were more negative at the time of this study, than prior to implementation in 2011. To buttress this conclusion, data presented in Table 9 shows this survey's respondents' overall self-reported perceptions of RTTT to be slightly less favorable than they were in 2011.

It is worth noting that in the current study, the RTTT component that was closest to being identified as statistically significant ($p = .125$) was the second component, which read, "1st 50% of Teacher Evaluation for non-tested subjects or levels is based on school-wide or team performance (VAM)". The mean scores for this RTTT component, by professional classification, showed the greatest mean difference between school district-based administrators ($\mu = 2.90$, close to "No Impact") and instructional personnel ($\mu =$

2.08, close to “Negative Impact”). While not statistically significant, the difference between the two groups’ perceptions of the second component of the teacher evaluation and compensation reforms from RTTT possibly indicates a lingering divide between school district-based administrators and instructional personnel as to the impact of school-wide VAM scores on student achievement and growth.

As with research question one, the instructional personnel’s negative perceptions of school-wide VAM scores are supported by research from Jones (2013) who found that teachers working under school-wide VAM designs were sufficiently displeased with the plan that they decreased the number of hours worked per week, and increased the amount of time spent looking for work outside of the field.

Research Question 3

To what extent, if any, is there a difference in the perceptions of administrative and instructional personnel who have different self-reported school poverty percentages about the impact of RTTT teacher evaluation and compensation components on student achievement/growth?

The quantitative analysis from a one-way ANOVA produced no statistically significant differences between the perceptions of survey respondents with different self-reported school poverty percentages as to the impact of RTTT reforms. Similar to findings in research questions one and two, the findings indicate rather homogenous perceptions of RTTT reforms. Analysis of the qualitative data found three main themes.

No statistically significant relationships were found between the perception scores of the selected RTTT components and self-reported school poverty level. Similar to findings in research question one and two, this indicated that respondents espoused similar perceptions of the impact of RTTT on student achievement and growth in the

final year of the initiative's funding, regardless of the level of poverty present in the school at which they work.

The three themes to emerge from analysis of the qualitative data centered on poverty's effect on, 1) student performance/learning, 2) teacher's ability to achieve desired outcomes, and 3) value-added scores. The themes were similar to those identified by Windish (2012). In his study, he found evidence of teacher anxiety related to the reforms, teachers viewing these reforms as punitive for teachers in high poverty schools, teachers in high poverty schools having a harder time reaching students due to factors outside of the new evaluation system, with teachers at low poverty schools being less likely to be negatively impacted by the reforms (Windish, 2012, p. 144).

In both the present study and Windish's (2012) study, there was no statistically significant difference in perceptions of the impact of RTTT between respondents, based on self-reported school poverty levels. The consistency of perceptions in this area, over time, indicated that educators at all levels believed there to be a strong correlation between poverty and academic achievement, and that the correlation was not adequately addressed by the RTTT reforms. The qualitative data collected discussing poverty's impact on student academic achievement is well supported in the literature (Gallagher et al., 2012; Gawlik et al., 2010; Hattie, 2009; Jensen, 2009; Konstantopoulos, 2009; Ladd, 2012; Lagana-Riordan & Aguilar, 2009; Lubienski & Crane, 2010; Mangiante, 2011; Okpala, 2000; Popp, Grant & Stronge, 2011; Robbins & Judge, 2010; Sirin, 2005). Educators are right to perceive student poverty as a significant hurdle in the learning process. Further, an evaluation system that seeks to compensate for student-level

variables, yet fails to adequately account for SES, may not be effective at identifying quality learning environments or improving learning outcomes.

Research Question 4

To what extent, if any, have administrative and instructional personnel changed in their perceptions of RTTT evaluation and compensation components, from the time RTTT was first implemented to the date of this study?

The results from a one-way ANOVA found no difference between the professional classifications and their self-reported change in perception of RTTT reforms from 2011 to the date of this study. The quantitative analysis indicated that the extent to which respondents changed in their professional perceptions of the reforms was rather homogenous, regardless of professional classification.

Consistent with findings described under each of the three previous research questions, the qualitative data showed a slightly negative trend for respondents' perceptions of RTTT from prior to implementation in 2011 to 2014. Qualitative data collected from this survey item buttressed the findings from the quantitative data for the fourth research question. Responses to the open-ended question showed limited diversity of content or theme, with the majority of respondents reporting their change in professional classification had no impact on their perceptions of RTTT. Among the respondents that reported a change in their perceptions of the initiative due to their change in professional classification, the majority found their new professional classification to negatively impact their views of RTTT. These findings support the data previously discussed which showed a decrease in overall perceptions of RTTT from 2011 to the date of the current study.

Implications for Policy and Practice

In the following section, three main implications from the results of this study are presented. School-, school district-, and state-level educational leaders should consider these implications as they continue to implement RTTT reforms as well as consider new reforms in the future.

1. Similar to the recommendation made by Windish (2012), additional communication about the specifications of RTTT is needed. Educational leaders continue to express uncertainty related to the details of the reforms that are implemented in their school districts and their state. State- and school district-level educational leaders must continue to provide professional development opportunities for school-level staff to develop a more concrete understanding of new policies. No educational reform effort will be successful if state and school district personnel are ineffective at communicating the specifics of the policy or proposal to the individuals responsible for implementing them at the school level.
2. Administrative and instructional personnel continue to be skeptical of the implications of RTTT three years after it was introduced and adopted by the states. As the impact of RTTT is tracked over the coming years, state and federal legislators should be cognizant of the perceptions and opinions of these leaders as they seek to make further improvements to the quality of public education. Teachers and administrators are acutely aware of the problems facing students in the classrooms and may be a significantly under-utilized resource in policy efforts to drive student achievement.

3. Political leaders should be more mindful of research-based findings related to how students learn, as well as the factors that limit achievement and growth and incorporate those factors into education reforms in the future. As is evidenced by the close alignment of qualitative and quantitative data from this survey to findings from existing research in the field, the knowledge exists to truly drive student achievement, but that information must be actualized in major policy proposals to effect real-world change. Recent research (Fritz, 2014) has identified revisions to the RTTT evaluation system that may produce a tool better equipped to impact teacher effectiveness and drive student learning.
4. Given the empirical evidence that exists in the literature, supported by data presented in the current study, the challenges faced by children living in poverty are well documented by researchers and well understood by educational professionals. State and federal lawmakers interested in improving educational outcomes for all students should make concerted efforts to address the prevalence and persistence of child poverty in the United States. Social policy changes, such as raising the minimum wage, reconfiguring the Earned Income Tax Credit, or other make-work-pay policies (Sawhill & Karpillow, 2014, p. 3) have the potential to lift an estimated 1.8 million people out of poverty. While educational reform policies may be effective at improving life outcomes for children in poverty over the long-term, changes in social policies that support those in poverty have a greater likelihood of short- and intermediate-term benefits for poor families, and more

specifically for poor children, whose academic achievement is frequently handicapped by limited family resources.

Recommendations for Future Research

The following section presents recommendations for future research based upon the findings of the current study.

1. Researchers, including this one, often rely upon free and reduced lunch status to gauge the level of poverty in public schools. There are serious limitations to the use of this tool in measuring and responding to poverty. Since socio-economic status is such a powerful predictor of student achievement in school, a more accurate measure should be developed to assist educational leaders, political leaders, and future researchers in understanding and addressing the needs of those affected by poverty.
2. More research is needed to understand, and improve, the way information flows through large organizations, such as school districts. Failure to effectively communicate pertinent information, including goals and objectives, from the top of an organization, to its front line employees can have a negative impact on the effectiveness and efficiency of the entire organization. Researchers in the field of educational leadership should attempt to better understand the ways in which an organization's structure can preclude the effective flow of information throughout and propose mechanisms by which to improve organizational efficiency.

Limitations of this study

There are multiple limitations that should be considered by those seeking to interpret the findings from this study. Great care was taken by this researcher throughout the design process for this study to maximize its generalizability, however, several limitations to the design became apparent as the study commenced and those issues should be carefully considered in light of the following:

The sample for this study was made up of graduate students in education and educational leadership at a large, urban university in the state of Florida. As such, the perceptions enumerated in this study may not be representative of educational leaders in more rural parts of the country, or of educational leaders who do not pursue advanced degrees in education.

While RTTT been the focus of much discussion and analysis since it was first proposed, much of the actual impact of its reforms may not be fully evident for a number of years to come. It may be too early to tell if the RTTT reforms will be successful at improving student learning outcomes, and as time elapses the perceptions of educational leaders, related to the reforms, may be significantly different than they were at the time of this study.

Conclusion

By conducting this study, the researcher expanded the research base on RTTT as well as the literature on teacher compensation and evaluation reforms. This study was conducted to better understand the fairness and impact of RTTT reforms, from the perspective of educational leaders in central Florida. This was accomplished by using a

researcher-created survey to gather qualitative and quantitative data during the final year of RTTT funding.

Windish's (2012) study found significant differences between administrative and instructional personnel's perceptions of the RTTT initiative prior to implementation. The results from this study indicate that the professional perceptions of these professionals are more homogenous and generally more negative in the fourth year of the Federal program than they were in the first.

It is also important to note that when considering the effects of poverty on student achievement, educational leaders are not only acutely aware of the problems facing this student population, but can also clearly articulate those challenges using empirically-based language from research in the field. This appreciation for, and understanding of, poverty's negative impact on students does not appear to be shared by the legislators shaping educational reforms. This researcher would challenge that were this not the case, educational reforms enacted at the federal and state level would be more closely aligned with the literature on the subject and the knowledge of practitioners in the field.

APPENDIX A
SURVEY OF THE FAIRNESS AND IMPACT OF TEACHER EVALUATION AND
COMPENSATION COMPONENTS OF RACE TO THE TOP

Survey of the Fairness and Impact of Teacher Evaluation and Compensation Components of RTTT

The Race to the Top (RTTT) initiative awarded funds to the state of Florida to reform schools. Your input is needed to understand the professional perceptions of the teacher evaluation and compensation components of this program on student achievement and growth.

Directions: Please fill in or select the appropriate response for each item. PLEASE DO NOT COMPLETE THIS SURVEY MORE THAN ONCE, EVEN IF YOU RECEIVE IT MULTIPLE TIMES.

What is your gender?

- Female
- Male
- Prefer not to disclose

In which Graduate Degree Program are you enrolled?

- Ed.D. in Education
- Executive Ed.D. in Educational Leadership
- Ed.S. in Educational Leadership
- M.Ed. in Educational Leadership
- Modified Core in Educational Leadership
- Other

Current Professional Classification:

- School District-based Administrator
- School-based Administrator (principal, assistant principal, dean)
- Instructional (classroom teacher, counselor, dean, specialist)
- School District-based Instructional Coach
- School-based Instructional Coach
- Other (Please Specify) _____

Current School Level where Employed or Interned:

- Elementary
- Middle
- K-8
- High
- School District
- Higher Education (College or University)
- Other (Please Specify) _____
- N/A

Current Percentage of Free/Reduced Lunch at School where Employed or Interned:

- 0 -24
- 25 - 49
- 50-74
- 75-100
- I Don't Have Enough Information
- N/A

Current School District where Employed, if applicable:

- Brevard County
- Flagler County
- Lake County
- Orange County
- Osceola County
- Polk County
- Seminole County
- Florida Virtual School
- Other (Please Specify) _____
- N/A

For the following 3 items, please select the response that best fits your role prior to implementation of RTTT (in 2011).

Please select the response below that best matches your pre-RTTT Professional Classification:

Pre-RTTT School Level:

- Elementary
- Middle
- K-8
- High
- School District
- Higher Education (College or University)
- Other (Please Specify) _____
- N/A

Pre-RTTT Percent of Free/Reduced Lunch at the school where employed or interned:

- 0 - 24
- 25 - 49
- 50 - 74
- 75 - 100
- I Don't Have Enough Information
- N/A

From where have you received your information on RTTT? Select all that apply.

- School District
- Graduate Classes
- State Conferences
- Educational Journals/Publications
- Email Communication from RTTT
- Professional Organizations
- Guest Speakers
- Collective Bargaining Unit
- FLDOE
- Email Blasts
- Colleagues
- Webinars
- Media/News
- U.S. DOE
- Other (Please Specify) _____

Rate your knowledge of Race to the Top using the following scale:

- Expert Knowledge (Can facilitate a seminar on RTTT)
- Great Knowledge
- Moderate Knowledge
- Little Knowledge
- No Knowledge (I have not heard of RTTT)

Based on your knowledge of RTTT, rate the FAIRNESS of the initiative concerning the following two items:

	Extremely Fair	Fair	Unfair	Extremely Unfair	I Don't Have Enough Information	Not Applicable
Teacher Evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teacher Compensation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

To what extent has your perception of RTTT changed from prior to implementation in 2011 to today?

	Much More Favorable	Somewhat More Favorable	No Change	Somewhat Less Favorable	Much Less Favorable	I Don't Have Enough Information	Not Applicable
Compared to 2011, today my perception of RTTT is...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In your experience, to what extent have the RTTT teacher evaluation and compensation components improved the quality of public education in the state of Florida?

	Greatly Improved	Improved	Somewhat Improved	Not At All Improved	I Don't Have Enough Information
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rate the IMPACT of the following RTTT components on student achievement and growth.

	Strong Positive Impact	Positive Impact	No Impact	Negative Impact	Strong Negative Impact	I Don't Have Enough Information
The first 50% of Teacher Evaluation/Appraisal is based on student performance on a Statewide Assessment (VAM).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The first 50% of Teacher Evaluation/Appraisal, for those who teach a subject or level in which students are not tested, is based on school-wide or team performance (VAM).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The second 50% of Teacher Evaluation/Appraisal is based on administrator observations of core effective practices and at least one additional metric.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers may be able to optionally participate in a separate performance pay scale (Performance Pay).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at the lowest performing schools may be offered recruitment and retention salary enhancements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Have you been assessed under RTTT's requirements for professional evaluation and compensation?

- Yes
- No

Answer question if "Yes" is selected for "Have you been assessed under RTTT's requirements for professional evaluation and compensation?"

Do you believe your evaluation was fair?

- Yes
- No

How has your professional perception of RTTT's FAIRNESS changed from 2011 to today?

(Narrative Response)

How has your professional perception of RTTT's IMPACT on student achievement/growth changed from 2011 to today?

(Narrative Response)

Has your professional classification changed since 2011?

- Yes
- No

Answer question if "Yes" is selected for "Has your professional classification changed since 2011?"

How has your change in professional classification impacted your perception of RTTT?

(Narrative Response)

How does your perception of RTTT compare with other professionals with whom you have had related discussions?

(Narrative Response)

In your experience, how does school poverty relate to teachers' and administrators' evaluations under the new performance evaluation system?

(Narrative Response)

Thank you for taking the time to complete this electronic survey!

APPENDIX B
WRITTEN CONSENT TO EDIT AND REPLICATE SURVEY AND FOLLOW-UP
PROTOCOL

Saturday, March 22, 2014 at 10:21:42 PM Eastern Daylight Time

Subject: RE: Request for permission to utilize your dissertation instruments
Date: Saturday, March 22, 2014 at 10:12:35 PM Eastern Daylight Time
From: Daniel Windish
To: orin
CC: Rosemarye Taylor

Orin,

Yes, you have my permission to use and adjust both my survey instrument and follow-up interview protocol. I have attached my final copy of my dissertation in Word format which will give you electronic versions of both of the items you requested.

My survey was an electronic survey designed in Survey Monkey. Starting on page 154 of the document, what you will see represents what the survey looked like in Survey Monkey. Feel free to adjust as you see fit when you redesign it electronically for your use. The instrument was called the Electronic Survey of the Implementation and Impact of Teacher Evaluation and Compensation Components from the Race to the Top Grant. After a few failed attempts with other abbreviations, I came up with STECC-RTTT as the abbreviation.

My protocol for interviewing begins on page 168. This also may need some adjusting, as the country is now more familiar with RTTT and Florida is familiar with the Student Success Act.

I wish you luck with your endeavors and look forward to seeing the results of your study.

-Dan

From: orin <orin@knights.ucf.edu>
Sent: Friday, March 21, 2014 4:12 PM
To: Daniel Windish
Cc: Rosemarye Taylor
Subject: Request for permission to utilize your dissertation instruments

Dr. Windish,

My name is Orin Smith. I am an Ed.D. Student of Dr. Rosemarye Taylor, at UCF. I would like to replicate the study you completed for your dissertation in 2012 but from a post-implementation perspective.

After speaking with Dr. Taylor, I would like to ask if I may utilize the survey instrument and follow-up protocol from your study?

If you are amenable to this, may I also ask if you have an electronic version of those documents that you could provide me? Dr. Taylor and I agree they are excellent tools and believe it would greatly

Page 1 of 2

APPENDIX C
INSTITUTIONAL REVIEW BOARD APPROVAL



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

Approval of Exempt Human Research

From: **UCF Institutional Review Board #1
FWA00000351, IRB00001138**

To: **Orin Smith**

Date: **May 05, 2014**

Dear Researcher:

On 5/5/2014, the IRB approved the following activity as human participant research that is exempt from regulation:

Type of Review: Exempt Determination
Project Title: CENTRAL FLORIDA EDUCATIONAL LEADERS'
PROFESSIONAL PERCEPTIONS OF RACE TO THE TOP
COMPONENTS CONCERNING TEACHER EVALUATION
AND COMPENSATION
Investigator: Orin Smith
IRB Number: SBE-14-10268
Funding Agency:
Grant Title:
Research ID: N/A

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

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

Signature applied by Joanne Muratori on 05/05/2014 10:34:03 AM EDT

IRB Coordinator

APPENDIX D
INFORMED CONSENT

Date

Dear Recent Doctoral Graduates and Current Doctoral Students:

You are invited to participate in a study of educational leaders in Central Florida and their professional perception of certain Race to the Top components concerning teacher evaluation and compensation. This study will add to the field of knowledge of how such components impact student achievement and student growth.

Students enrolled in either Educational Leadership or Education graduate and doctoral programs were selected to participate. Your program coordinator has used your knights email to distribute this survey. The link found below will bring you to the actual electronic survey found on Qualtrics®.

Although asked for general demographic information, your responses will be kept anonymous, as the researcher will not have access to your email address nor your name nor other personally identifying information. The researcher is only interested in your honest, professional perceptions of the Race to the Top components concerning teacher evaluation and compensation.

I will be available to explain this research study to you. Whether or not you take part is up to you. You can agree to take part now and later change your mind. Whatever you decide, it will not be held against you. There are no anticipated risks or benefits to participating in this study. Please feel free to ask all the questions you may have prior to deciding whether or not to participate. This survey should take approximately 5-10 minutes to complete.

If you have any questions about this study, please contact me at orin@knights.ucf.edu. My faculty advisor, Dr. Rosemarye Taylor, may also be contacted by phone at (407) 823-1469 or by email at rosemarye.taylor@ucf.edu. Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed to the UCF Institutional Review Board Office at the University of Central Florida, Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone number is (407) 823-3778.

By clicking on this link, http://ucf.qualtrics.com//SE/?SID=SV_2a84Ao9vMRywnwp, you are giving your informed consent to participate in the survey.

Thank you in advance for taking the time to participate.

Best Regards,

Orin T. Smith
Doctoral Candidate, University of Central Florida
ESE Paraprofessional, Dr. Phillips High School
Orange County Public Schools
orin@knights.ucf.edu
(407) 409-1296

REFERENCES

- American Institutes of Research (n.d.). *Florida Value-Added Model Technical Report*. Retrieved from <http://www.fldoe.org>
- Aritomi, P., Coopersmith, J., & Gruber, K. (2009). *Characteristics of public school districts in the United States: Results from the 2007-08 schools and staffing survey*. (NCES 2009-320). Washington, D.C.: National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED505835.pdf>
- Boser, U. (2012). *Race to the top: What have we learned from the states so far? A state-by-state examination of race to the top performance*. Center for American Progress. Retrieved from <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED535605>
- Brown, I. I., & Crumpler, T. (2013). Assessment of foreign language teachers: A model for shifting evaluation toward growth and learning. *The High School Journal*, 96(2), 138-151. doi: 10.1353/hsj.2013.0003
- Buddin, R., McCaffrey, D. F., Kirby, S. N., & Xia, N. (2007). *Merit pay for Florida teachers: Design and implementation issues*. Working paper, RAND Education.
- Burnett, A., Cushing, E., & Bivona, L. (2012). *Uses of multiple measures for performance-based compensation*. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education, Center for Educator Compensation Reform. Retrieved from <http://eric.ed.gov/?id=ED533704>

- Consortium for Policy Research in Education (2012). History of teacher pay. Retrieved at <http://cpre.wceruw.org/tcomp/general/teacherpay.php>
- Darling-Hammond, L. (2010). *The flat world and education: How American's commitment to equity will determine our future*. New York, NY: Teachers College Press.
- Dee, T. & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446. doi: 10.1002/pam.20586
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Petegem, P. V. (2012). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education*, 36(1), 1-11. doi: 10.1016/j.tate.2013.06.011
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys. The tailored design method*. Hoboken, NJ: Wiley.
- Dolton, P., & Mercenaro-Gutierrez, O. (2013, October). *Global Teacher Status Index*. Varkey GEMS Foundation. Retrieved from <https://www.varkeygemsfoundation.org/sites/default/files/documents/2013GlobalTeacherStatusIndex.pdf>
- Education Commission of the States (2007). Evaluating Teacher Effectiveness: Assessing teacher quality and which teachers are most likely to improve student achievement. *The Progress of Education Reform 2007*, 8(4). Retrieved from <http://files.eric.ed.gov/fulltext/ED512136.pdf>

- Florida Department of Education (n.d.). *U.S. Department of Education: Race to the top definitions*. Retrieved from <http://www.fldoe.org/arra/pdf/usdef.pdf>
- Florida Department of Education (2011). Review and approval checklist for RTTT teacher evaluation systems 6-1-2011. Retrieved from <https://www.fldoe.org/arra/pdf/ReviewApprovalChecklist.pdf>
- Florida Department of Education (2012). Review and approval checklist for instructional personnel and school administrator evaluation systems. DOE Form no. EQEVAL-2012. Retrieved from <http://www.fldoe.org/profdev/pa.asp>
- Florida: State-reported ARP: SY 2012-2013 (2014). Race to the Top. Retrieved from <https://www.rtt-apr.us/state/florida/2012-2013/intro>
- Gallagher, K. S., Goodyear, R., Brewer, D. J., & Rueda, R. (2012). *Urban education: A model for leadership and policy*. New York, NY: Routledge.
- Gawlik, M. A., Kearney, C. P., Addonizio, M. F., & LaPlante-Sosnowsky, F. (2010). Teacher quality in Michigan: A school-level analysis of the Detroit metropolitan region. *Education and Urban Society*, 44(4), 412-447. doi: 10.1177/0013124510392568
- Goldhaber, D. D., Goldschmidt, P., & Tseng, F. (2013). Teacher value-added at the high-school level: Different models, different answers? *Educational Evaluation and Policy Analysis*, 35(2), 220-236. doi:10.3102/0162373712466938
- Goldhaber, D., & Walch, J. (2012). Strategic pay reform: A student outcomes-based evaluation of Denver's ProComp teacher pay initiative. *Economics of Education Review*, 31(6), 1067-1083. doi: 10.1016/j.econedurev.2012.06.007

- Green, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Education Evaluation and Policy Analysis, 11*(3), 255-274. doi: 10.3102/01623737011003255
- Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measure of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education, 119*(3), 445-470. doi: 10.1086/669901
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Herlihy, C., Karger, Z., Pollard, C., Hill, H. C., Kraft, M. A., Williams, M., & Howard, S. (2014). State and local efforts to investigate the validity and reliability of scores from teacher evaluation systems. *Teachers College Record, 116*(1), 1-28.
Retrieved from
<http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=94262510&site=ehost-live>
- Howell, W. G., Peterson, P. E., & West, M. (2011). Meeting of the minds. *Education Next, 11*(1). Retrieved from <http://educationnext.org/meeting-of-the-minds/>
- Jensen, E. (2009) *Teaching with poverty in mind: What being poor does to kids' brains and what schools can do about it*. Alexandria, VA: ASCD.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133. doi: 10.1177/1558689806298224

- Jones, M. D. (2013). Teacher behavior under performance pay incentives. *Economics of Education Review*, 37, 148-164.
- Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615-631. doi: 10.3386/w12155
- Kennedy, J. F. (1961). Proclamation 3422 – American education week, 1961. *The American Presidency Project*. Retrieved from <http://www.presidency.ucsb.edu/ws/?pid=24146>
- Knoepfel, R. C., Logan, J. P., & Keiser, C. M. (2005, November). *Measuring teacher quality: Continuing the search for policy-relevant predictors of student achievement*. Paper presented at the annual meeting of University Council for Educational Administration, Nashville, TN.
- Kolbe, T., & Rice, J. K. (2012). And they're off: Tracking federal race to the top investments from the starting gate. *Educational Policy*, 26(1), 185-209. doi: 10.1177/0895904811428975
- Konstantopoulos, S. (2009). Effects of teachers on minority and disadvantaged students' achievement in early grades. *The Elementary School Journal*, 110(1), 92-113. Retrieved from: <http://www.jstor.org/stable/10.1086/958845>
- Ladd, H. F. (2012). Presidential address: Education and poverty: Confronting the evidence. *Journal of Policy Analysis and Management*, 31(2), 203-227.
- Lagana-Riordan, C., & Aguilar, J. P. (2009). What's missing from no child left behind? A policy analysis from a social work perspective. *Children & Schools*, 31(3), 135-144. doi: 10.1093/cs/31.3.135

- Laine, S., Potemski, A., & Rowland, C. (2010). Compensation reform in the schools. *School Administrator*, 67(3), 10-14. Retrieved from <http://www.aasa.org/SchoolAdministratorArticle.aspx?id=12526>
- Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990-2009 reading and math achievement gaps and trends. *Education Evaluation and Policy Analysis*, 36(2), 209-231. doi: 10.31.02/0162373711431604
- Levine, M., & Levine, A. (2012). Education deformed: No child left behind and the race to the top. 'This almost reads like our business plans'. *American Journal of Orthopsychiatry*, 82(1), 104-113. doi: 10.1111/j.1939-0025.2011.01142.x
- Lubienski, S. T., & Crane, C. C. (2010). Beyond free lunch: Which family background measures matter? *Education Policy Analysis Archives*, 18(11). Retrieved from <http://epaa.asu.edu/ojs/article/view/756>
- Mangiante, E. M. (2011). Teachers matter: Measures of teacher effectiveness in low-income minority schools. *Educational Assessment, Evaluation and Accountability*, 23(1), 41-63. doi: 10.1007/s11092-010-9107-x
- Master, B. (2014). Staffing for success: Linking teacher evaluation and school personnel management in practice. *Education Evaluation and Policy Analysis*, 36(2), 207-227. doi: 10.3102/0162373713506552
- McNeil, M. (2014). Race to top reports detail winners' progress, challenges. *Education Week*, 33(26), 25-30. Retrieved from <http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=95101402&site=ehost-live>

- Mitra, D. (2011). Pennsylvania's best investment: The social and economic benefits of public education. *Education Law Center*. Retrieved from http://www.elc-pa.org/wp-content/uploads/2013/11/BestInvestment_Full_Report_6.27.11.pdf
- National Center for Children in Poverty (2014). *Child Poverty*. Retrieved from <http://www.nccp.org/topics/childpoverty.html>
- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for education reform*. Washington, DC: United States Department of Education.
- National Governors Association. (2011, December). *New models of teacher compensation: Lessons learned from six states* (Issue Brief). Washington, DC: NGA, Center for Best Practices. Retrieved from <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED532521>
- No Child Left Behind Act, 20 U.S.C. § 6301 (2001).
- Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis*, 30(2), 111-140. doi: 10.3102/016237308317680
- Podgursky, M., & Springer, M. G. (2007). Credentials versus performance: Review of the teacher performance pay research. *Peabody Journal of Education*, 82(4), 551-573. doi: 10.1080/01619560701602934
- Popp, P. A., Grant, L. W., & Stronge, J. H. (2011). Effective teacher for at-risk or highly mobile students: What are the dispositions and behaviors of award-winning

teachers? *Journal of Education for Students Placed at Risk*, 16(4), 275-291. doi: 10.1080/10824669.2011.610236

Praetorius, A. K., Pauli, C., Reusser, K., & Rakoczy, K. (2013). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2-12. doi: 10.1016/j.learninstruct.2013.12.002

Robbins, S. P., & Judge, T. A. (2010). *Essentials of Organizational Behavior*. Upper Saddle River, NJ: Pearson.

Sanders, W. L. & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Sawhill, I. & Karpilow, Q. (2014) *Raising the minimum wage and redesigning the EITC*. Center on Children and Families, The Brookings Institute. Retrieved from <http://www.brookings.edu/~media/research/files/papers/2014/01/30%20raising%20minimum%20wage%20redesigning%20eitc%20sawhill/30%20raising%20minimum%20wage%20redesigning%20eitc%20sawhill.pdf>

Schafer, W. D., Lissitz, R. W., Zhu, X., Zhang, Y., Hou, X., & Li, Y. (2012). Evaluating teachers and schools using student growth models. *Practical Assessment, Research & Evaluation*, 17(17), 1-21. Retrieved from <http://pareonline.net/getnv.asp?v=17&n=17>

School Improvement Grants; American Recovery and Reinvestment Act of 2009 (ARRA); Title I of the Elementary and Secondary Education Act of 1965, as Amended (ESEA) (2010). Federal Register, 75(208), 66363-66371.

School Leader Evaluation Model: Observation and evaluation forms procedures for leadership practice effective July 1, 2012 (2012). Orange County Public Schools.

Retrieved from

https://www.ocps.net/es/hr/PDS/leadership/Documents/OCPS_Principal_Leadership_Evaluation_Proposal.pdf

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417-453.

Shields, R. (2012). *Strategic design of teacher compensation*. Education Resource Strategies. Retrieved from

<http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED541221>

Smarick, A. (2011). Diplomatic mi\$\$ion: President Obama's path to performance pay.

Education Next, 11(1), 56-63. Retrieved from

<http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=508207784&site=ehost-live>

Race to the Top Fund (2014). U.S. Department of Education. Retrieved from

<http://www2.ed.gov/programs/racetothetop/index.html>

United States Census Bureau (2014). *Income, poverty, and health insurance in the United States: 2012 – Highlights*. Retrieved from

<https://www.census.gov/hhes/www/poverty/data/incpovhlth/2012/highlights.html>

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2012). *School and staffing survey (SASS), 2011-2012*.

Retrieved from

http://nces.ed.gov/surveys/sass/tables/sass1112_2013311_d1s_002.asp

Windish, D. H. (2012). *Central Florida educational leaders' professional opinions of the race to the top grant components concerning teacher evaluation and compensation prior to implementation*. (Doctoral dissertation, University of Central Florida).

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practice*. Santa Monica, CA: Rand.

Retrieved from

<http://www.rand.org/content/dam/rand/pubs/reports/2006/R3139.pdf>