

Implementation of Reform with a Performance-Based Teacher Evaluation System: A Case Study
of One School District

Michelle Morgan

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Doctor of Education
in
Educational Leadership and Policy Studies

M. David Alexander, Chair

John K. Burton

Richard G. Salmon

Horace A. Seibert

October 24, 2014

Blacksburg, Virginia

Keywords: teacher evaluation, student achievement, student academic progress, annual goals,
performance-based teacher evaluation, and policy implementation

Implementation of Reform with a Performance-Based Teacher Evaluation System: A Case Study
of One School District

Michelle Morgan

Abstract

This dissertation focused on the new performance-based teacher evaluation system implemented in the Commonwealth of Virginia. Eight school districts¹ were granted an implementation waiver for one year making the 2013-2014 school year the implementation year for those school districts. A mixed methodology study was designed to understand teacher perceptions in one waiver school district in the Commonwealth of Virginia. The purposes of this study were to collect empirical evidence on the implementation of the new performance-based teacher evaluation system by comparing teachers' and principals' perceptions at three elementary schools to teachers' perceptions in the school district, identified as District Z and were to add to the research base on policy implementation theories and teacher evaluation.

Findings and analysis of research questions were based on the responses of 357 teacher participants, two district directors, and three principals. Policy implementation research suggested access to resources and support along with an alignment of beliefs and values resulted in more effective implementation practices. Conclusions of this study, while limited to the district studied, suggested professional development on the new performance-based teacher evaluation system, including the Uniform Performance Standards and SMART goals, would provide teachers and principals with the increased knowledge to use as a resource during new implementation phases. Additionally, involvement in the design and modification of the teacher evaluation system would encourage participation and provide opportunities for teachers and principals to develop shared values regarding aspects of the evaluation system.

¹ In Virginia, school districts are referred to as school divisions. Most of the literature and most states refer to school systems as school districts. This study will refer to school systems as school districts.

Dedication

A thankful heart is not only the greatest virtue, but the parent of all other virtues.

-- Cicero

I dedicate this work to the many people who have supported and encouraged me throughout my life. I will forever be grateful for the sacrifice my loved ones have made so that I could know success. Thank you for making dreams a reality.

Acknowledgements

Many people have provided support and guidance to me over the past three years. My personal growth and learning is a direct result of the substantial contributions of professors, colleagues, fellow students, friends and family, and loved ones. My good fortune fills me with gratitude.

Dr. M. David Alexander served as my advisor and has been a source of knowledge and expertise. He was and is one of my strongest supporters. His commitment to my success was a gift I will always be grateful to have received.

The members of my dissertation committee provided much appreciated time and effort to this project and to me. Dr. Richard Salmon and Dr. John Burton provided welcomed advice and encouragement, which led to a much better final product. Dr. Horace Seibert provided support and offered practical suggestions. Dr. Min Sun and R. Brock Mucheson provided resources and technical support. I am grateful for the opportunity to learn from each of the members of my committee.

I acknowledge the participants of this study for their time and effort. I know time is always in short supply and is a precious commodity for educators. The participants of this study added to the research base on teacher evaluation and policy implementation. Many thanks and much gratitude!

I appreciate the support of my professional colleagues, without which, this project would not be complete. Finally, I would like to thank my colleagues who were part of the Virginia Tech doctoral cohort. I learned so much from each of you and will always be grateful for the time we spent together. Working and learning with each of you made me a better leader! For that and so much more, I will always be grateful.

Table of Contents

Abstract.....	ii
Dedication.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Tables.....	vii
Chapter One Introduction.....	1
Background.....	1
National Developments Impacting Teacher Evaluation.....	3
ESEA Flexibility Waiver in Virginia.....	5
Development of <i>Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers</i>	5
Statement of the Problem.....	7
Purpose of the Literature Review.....	8
Significance.....	9
Research Questions.....	10
Chapter Two Review of Relevant Literature.....	12
The Research Base for Standard 7: Student Academic Progress.....	12
Emerging Themes.....	20
Additional Research on Teacher Evaluation, Value-Added Modeling, and Student Academic Progress.....	20
Emerging Themes.....	30
Policy Implementation Theories and Research.....	30
Emerging Themes.....	33
Synthesis and Conclusions.....	33
Chapter Three Methodology.....	35
Introduction.....	35
Population.....	37
Selection of Subset Teachers and Principals.....	39
District Z Study Population Including Subset Teacher and Principals.....	44
Data Collection.....	44
Data Analysis.....	47
Summary.....	50
Chapter 4 Presentation and Analysis of Findings.....	51
Introduction.....	51
Procedures.....	51
Procedures with District Z Directors.....	51
Procedures with District Z Teacher Data Set.....	52
Procedures with Principals of Subset Schools A, B, and C.....	52
Findings and Analysis of District Z Director Interviews.....	53
Findings and Analysis of District Z and Schools A, B, and C Teacher and Principal Data Set.....	56
Summary.....	96

Chapter Five Introduction, Executive Summary, Conclusions, Limitations/Delimitations, and Recommendations	98
Introduction.....	98
Executive Summary	98
Conclusions.....	99
Limitations/Delimitations	104
Recommendations for Further Research.....	106
References.....	107
Appendix A IRB Approval	112
Appendix B School Approval Letter	114
Appendix C Level of Confidences Table.....	115

List of Tables

Table 3.1 <i>Federal Ethnicity and Race Report for District Z Student Enrollment Reported in Percentages</i>	38
Table 3.2 <i>Descriptions of Staffing of School A, School B, and School C</i>	39
Table 3.3 <i>Federal Reporting Characteristics of School A, School B, and School C Student Enrollment, Students with Disabilities, and Economically Disadvantaged Students Reported in Percentages</i>	40
Table 3.4 <i>Federal Ethnicity and Race Report for Student Enrollment at School A, School B, and School C Reported in Percentages</i>	42
Table 3.5 <i>2013 SOL Performance of Students at School A, School B, and School C</i>	43
Table 3.6 <i>2014 SOL Performance of Students at School A, School B, and School C</i>	43
Table 3.7 <i>Proposed Study Teacher Participants</i>	44
Table 3.8 <i>Survey Instrument Question Categories Aligned with Research Data Sets</i>	49
Table 4.1 <i>Research Question 1: Preparation Actions of District Z During 2012-2013</i>	55
Table 4.2 <i>Research Question 2: Professional Development/Training Provided for Principals and Teachers during 2012-2013</i>	56
Table 4.3 <i>District Z Teacher Data Sets and Survey Instrument Participation Rates</i>	57
Tables 4.4 <i>Description of Teacher and Principal Data Sets</i>	59
Table 4.5 <i>Average Hours Worked in a Week and Average Total Hours Worked on Teacher Evaluation During the 2013-2014 School Year</i>	60
Table 4.6 <i>Principal Time Reported on Teacher Evaluation Activities and Professional Development</i>	61
Table 4.7 <i>Who Conducted Observations and the Frequency Reported by each Teacher Data Set</i>	61
Table 4.8 <i>Nominations of Colleagues Providing Support with Teacher Evaluation in the 2013-2014 School Year</i>	65
Table 4.9 <i>Professional Development Mean Hours Reported by Topic and Extent of Usefulness Reported by Percentage</i>	66
Table 4.10 <i>Extent of Agreement with Evaluation Outcome Statements Reported in Percentages</i> 70	
Table 4.11 <i>Importance of Purposes of Teacher Evaluation Reported in Percentages</i>	76
Table 4.12 <i>Extent of Foci of Teacher Evaluation in 2013-2014 Reported in Percentages</i>	79
Table 4.13 <i>Extent of Emphasis on Sources of Evidence in Teacher Evaluation Reported in Percentages</i>	83
Table 4.14 <i>Extent of Involvement in the Design and Modification of the Teacher Evaluation System Reported in Percentages</i>	87

Table 4.15 <i>Level of Confidence in Activities of the Teacher Evaluation System Reported in Percentages</i>	91
Table 4.16 <i>Extent of Performance Standards Consistency Reported in Percentages</i>	94
Table C1 <i>Level of Confidence in Activities of the Teacher Evaluation System Reported in Percentages</i>	115

Chapter One

Introduction

Background

The process of teacher evaluation has undergone a significant shift in the past few years. Most notably linking student achievement to teacher evaluation has received considerable notoriety from research journals, workshops, and other reputable press agencies. Many states have responded with legislation designed to link teacher evaluation to student achievement outcomes. The Commonwealth of Virginia passed Section 22.1-253.13.5 of the Virginia Code, which stated,

Consistent with the finding that leadership is essential for the advancement of public education in the Commonwealth, teacher, principal, and superintendent evaluations shall be consistent with the performance standards included in the Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers, Principals, and Superintendents. Evaluations shall include student academic progress as a significant component and an overall summative rating. Teacher evaluations shall include regular observation and evidence that instruction is aligned with the school's curriculum.

Evaluations shall include identification of areas of individual strengths and weaknesses and recommendations for appropriate professional activities. (Retrieved from:

<http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+22.1-253.13C5>)

This version of the 2012 Virginia Senate Bill No. 1223 was adopted March 20, 2013, and required school districts to implement evaluations consistent with the Virginia Department of Education (VDOE) guidelines, which specified the performance standards and weighted one standard - student academic progress - 40 percent of a teacher's evaluation.

Legislative action in many states and policy shifts at the federal level were based on current research, which suggested traditional teacher evaluation methods have not provided administrators and teachers with the information needed and required to improve the quality of instruction and increase student achievement. Weisberg, Sexton, Mulhern, and Keeling (2009) drafted a report to examine "our pervasive and longstanding failure to recognize and respond to variations in the effectiveness of our teachers" (p. 4). While Marzano and Toth (2013) repeated the need for change in teacher evaluation, they suggested teacher evaluation could increase student achievement only by focusing on improving teachers' instructional skills. Further,

Marzano and Toth suggested that attempts to alter teacher evaluation that did not focus on this critical component, improving teachers' instructional skills, would continue to fail to improve teacher quality and increase student achievement.

Several non-profit educational foundations supported improvement in the teacher evaluation process and have committed substantial effort and resources to improve the quality of teaching and improve student-learning outcomes. These organizations joined other researchers (Weisberg, Sexton, Mulhern, and Keeling, 2009 & Marzano and Toth, 2013) in recognizing the need for change in teacher evaluation, change that informed teaching practice and resulted in increased student achievement. Battelle for Kids (2011) published *Selecting Growth Measures: A Guide for Education Leaders* to provide educators with practical information to consider when making decisions about measuring student academic growth. The organization supported the use of student growth measures and Value-Added Modeling (VAM) as a useful tool to inform and improve educator practice and support professional growth and provided practical guidance regarding limitations and considerations of models. Two types of models were explained, including simple growth models and VAM, and each type was evaluated against seven key considerations: 1) intended uses; 2) inputs for analysis; 3) measurement error and uncertainty; 4) results and outputs; 5) communications, training, and support; 6) experience, expertise, and capacity; and 7) cost (pp. 3-4). Battelle for Kids suggested educators should spend considerable time and effort in determining the goals of an evaluation tool and use the seven considerations to make an informed decision about the type of growth measure to use.

Additionally, the Bill & Melinda Gates Foundation funded the Measures of Effective Teaching (MET) project (2013), and MET recently published the results of a large, three-year national study. The MET project involved over 3,000 volunteer teachers from seven school districts in the United States. The project studied classroom observation instruments, student perception surveys, and student achievement gains as tools to improve teaching practice. MET project published three important findings from the national study. First, the measures of effective teaching used in the project could identify the teachers whose students achieved higher on state assessments and other more challenging assessments in math and English (p. 4). Next, a balanced weighting system in the teacher evaluation model that included student achievement, classroom observation, and student surveys was more stable from year to year (p. 4). Finally, the reliability of classroom observations increased when another observer scored the same lesson as

opposed to the same observer scoring a second lesson (p. 5). The MET project researchers planned to continue to support professional growth of teachers by engaging more than 300 MET project teachers to build a video library of effective teaching practices for educators to use in professional development (p. 21).

The main stream press, including newspapers such as *The Wall Street Journal* and *The Washington Post*, has also called for reform of teacher evaluation and published how states are implementing a VAM of teacher evaluation in which student achievement or academic progress represents a portion of a teacher's overall evaluation. The use of VAM as a measure of teacher evaluation has exploded across the country with the support of both political parties. The requirements of evaluating and supporting teacher and principal effectiveness as part of Race to the Top and the No Child Left Behind (NCLB) Elementary and Secondary Education Act (ESEA) contributed substantial pressure to implement such practices.

Howell, Peterson, and West (2011) noted that Democrats and Republicans disagreed about very few educational reform issues. However, the two political parties have significant differences regarding school spending and teacher tenure based on the survey responses of the 2010 Education Next – Program on Education Policy and Governance (PEPG). The PEPG survey results also measured an increase in public support for basing a part of a teacher's salary on student academic progress, from 43 percent in 2009 to 49 percent in 2010. While school reform remains unlikely if the two political parties strongly disagree on the options, the PEPG survey results in 2010 demonstrated fewer differences than might be expected. The PEPG survey results were predictive of widespread state legislation regarding teacher evaluation and the use of student achievement and academic progress as an indicator of effectiveness.

National Developments Impacting Teacher Evaluation

The federal legislation NCLB was enacted in 2001 as part of landmark educational reform under President George W. Bush. The purpose of the NCLB Act of 2001 was “to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind” (Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/index.html>). NCLB required states to implement research based instructional strategies and programs to increase student achievement and to measure that achievement with rigorous accountability systems. Schools and school districts failing to meet benchmark achievement levels were required to implement corrective action plans. Dissatisfaction with NCLB and the failure of the legislation to achieve

its goals resulted in growing concerns among states and policy makers throughout the years of implementation.

The devastating economic recession of 2008 pushed the federal government to enact the American Recovery and Reinvestment Act (ARRA) of 2009. The ARRA legislation was enacted with a primary goal of stimulating the economy in the short term, with specific funds designated to invest in education. Education funds were to be used to provide a State Fiscal Stabilization Fund grant to help sustain and create jobs and advance education reforms, including evaluating and supporting teacher effectiveness. The United States Department of Education (USDOE) (2009) published a guidance document for state educational agencies (SEA) that provided suggested uses of ARRA funds that would drive school reform and improvement. The USDOE document provided five categories that were reflective of the priorities of ARRA including increasing teacher effectiveness and equitable distribution of effective teachers (p. 2). Further, the USDOE document specified:

Establish and implement a fair and reliable teacher evaluation system that provides ongoing feedback to teachers about their performance based on objective measures of student achievement outcomes and multiple classroom observations, that gives guidance for improving instructional practices, and that is used to inform teacher professional development and advancement. (USDOE, 2009, p. 4)

Many SEAs received the needed funds to develop new performance-based teacher evaluation systems when a priority linking teacher evaluation with student achievement was established for the ARRA funds. Additionally, this priority would add critical national support to teacher evaluation reform already under way for many SEAs, including Virginia.

Under the leadership of President Barack Obama, the Secretary of Education, Arne Duncan, was directed to provide states with flexibility and relief to NCLB. This relief was offered by the USDOE to SEA through the ESEA Flexibility Request. The ESEA Flexibility Request was developed to provide SEA and local education agencies (LEA) with flexibility regarding specific requirements of the NCLB Act of 2001. This flexibility, when approved by the USDOE, would be granted to states who substituted rigorous plans to improve educational outcomes in three categories: 1) transitioning to college- and career-ready standards and assessments; 2) developing systems of differentiated recognition, accountability, and support; and 3) evaluating and supporting teacher and principal effectiveness. SEA flexibility waivers

were evaluated by external peer reviewers and staff reviewers based on whether and how each request would support comprehensive improvements in the three designated areas of improvement. According to the USDOE (2013), only five states have not requested or been approved for an ESEA Flexibility Waiver including: California, Montana, Nebraska, North Dakota, and Vermont. Three additional states, Illinois, Iowa, and Wyoming, have requested ESEA Flexibility Waivers, which were under review.

ESEA Flexibility Waiver in Virginia

The VDOE submitted the final version of the ESEA Flexibility Request on July 24, 2012. The VDOE's preparatory work on the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* and the established federal commitment to teacher evaluation and student achievement outcomes enabled Virginia to meet one of the waiver requirements. When the flexibility waiver was submitted, the VDOE indicated the SEA had adopted all of the guidelines consistent with Principal 3 including 3.A Develop and adopt guidelines for local teacher and principal evaluation and support systems and 3.B Ensure LEAs implement teacher and principal evaluation and support systems. Specifically, the waiver noted the Governor's proposed legislation, now approved, in which an overall summative rating for teachers and principals included student academic progress as a significant component of the evaluation. The waiver included a history of teacher evaluation in the commonwealth and the path to the current revised teacher evaluation program. The ESEA Flexibility Request (2012) included a timeline of the development of the new performance-based teacher evaluation system, noting the Virginia Board of Education (VBOE) approved the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* and the *Virginia Standards for the Professional Practice of Teachers* in April 2011 and authorized implementation of the new evaluation system effective July 1, 2012.

Development of *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*

The inclusion of student academic progress as a weighted component of teacher evaluation in Virginia was a result of political forces at the national and state level accepting public support of this idea. The new performance-based evaluation model used by districts throughout the Commonwealth was outlined in the *Guidelines for Uniform Performance*

Standards and evaluation Criteria for Teachers, codified by law in 2013, and required for approval in the federal ESEA Flexibility Waiver of the Secondary Education Act of 1965. The VaBOE approved the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* on April 28, 2011, to become effective July 1, 2012. School boards and districts were authorized to implement prior to the effective date. As mentioned, one of the catalysts behind the development and implementation of the new evaluation model was to meet the requirements pursuant to the ESEA Flexibility Request to grant SEA relief from certain requirements of NCLB. The VDOE submitted final revisions to the ESEA Flexibility Request on July 24, 2012, which were approved by the USDOE.

The *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* represented the product of a state level Work Group. This Work Group was tasked with conducting a comprehensive review of teacher evaluation and included members from K-12 and higher education, professional organizations, and VaDOE staff. James Stronge and Terry Dozier, two expert consultants from the Center for Innovative Technology (CIT), led the Work Group. The Work Group met three times, August 2010, October 2010, and December 2010. A subcommittee of the Work Group met in March 2011 to review documents and present to the VBOE the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* and *Virginia Standards for the Professional Practice of Teachers*. The document outlined the seven standards all teacher evaluations must be consistent with pursuant to state law. Further, the document explained the evaluation criteria and the weighting of each of the standards, with standards one through six weighted equally at 10 percent each and standard seven, student academic progress, weighted at 40 percent of the summative evaluation.

The *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* suggested student learning account for 40 percent of the evaluation. The document suggested 20 percent, half of the student academic progress measure be comprised of student growth measures as provided by VDOE, where available and appropriate. The document suggested the other 20 percent be comprised of alternative measures (*Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, 2011, p. 42). Because only 30 percent of Virginia public school teachers have an SOL assessment as a direct measure of student achievement and student growth, student achievement data will not be provided to many teachers based on the SOL testing results. Multiple measures of student academic progress were

suggested for those teachers who would not be provided data from the state. While student growth percentiles were recommended as one half of the 40 percent measurement of student academic progress, the use of student growth percentiles was not implemented in most school districts in the Commonwealth because those measures were not available. The primary approach suggested to link teacher evaluation to student academic progress was student achievement goal setting. Criterion referenced tests, norm-referenced tests, standardized achievement tests, school adopted interim/common/benchmark assessments, and authentic measures (learner portfolio, recitation, performance) were suggested as effective measurement tools for identifying student academic progress.

As stated, all school boards and school districts were required to implement the standards on or before July 1, 2012. If a school district failed to implement teacher evaluation with standards consistent to those in the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* by the July 1, 2012, implementation deadline, those school boards and districts were required to write a corrective action plan indicating how they would reach the requirement of Standard 7 – Student Academic Progress as weighted at 40% of a teacher’s and principal’s evaluation. This corrective action plan waived implementation of the new performance-based teacher evaluation system for one year based on a description of how the school district would take the necessary steps to comply with the requirements by July 1, 2013.

Statement of the Problem

The document VaDOE *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, the approved ESEA Flexibility Waiver, and the passage of Section 22.1-253.13.5 of the Virginia Code required the use of consistent standards, as identified in the performance document, and that those standards be weighted with standard seven representing 40 percent of a teacher’s summative evaluation. While school boards and districts' retained local authority to modify the evaluation document, the standards and the weighting were not optional components of the policy. Because of this school boards and school districts were afforded the opportunity to review the documents and develop a local evaluation tool that was consistent with the performance standards and weighting through a corrective action plan.

Only eight school districts in the Commonwealth elected to defer implementation until July 1, 2013, including: Alexandria City Public Schools, Appomattox County Public Schools, Bedford County Public Schools, Botetourt County Public Schools, Campbell County Public

Schools, Frederick County Public Schools, Loudon County Public Schools, and Spotsylvania County Public Schools. All eight school districts completed the TEPC data collections questions indicating lack of compliance with state law regarding implementation of the consistent standards and weighting of Standard 7 – Student Academic Progress and each of the eight districts completed corrective actions plans as required by the VaDOE.

Because these eight school districts requested and were granted an implementation waiver, the 2013-2014 school year will be the implementation year in each of these district's teacher evaluation policies with a mandatory requirement of student academic progress weighted at 40 percent of the teacher's overall evaluation. Shifting from a traditional teacher evaluation system to a performance-based teacher evaluation system represented a major shift for educators. Implementation of new policies remains an expensive undertaking and requires substantial resources, both dollars and human, to be successful. Determining how the new requirement was implemented in a waiver district would provide important information regarding policy implementation. Further, a case study in one of the waiver districts examining teachers' perceptions of the new evaluation policy would provide information regarding implementation and would be useful for policy makers and educational leaders particularly when making decisions regarding allocations of capital, both human and fiduciary.

Purpose of the Literature Review

The purpose of the literature review that follows was to provide a comprehensive analysis of the research cited as the basis for the VDOE's *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, review additional research on performance-based teacher evaluation systems, and review policy implementation research. This research was compiled in the VDOE's document *The Research Base for the Uniform Performance Standards for Teachers* and approved by the VBOE in 2011. *The Research Base for the Uniform Performance Standards for Teachers* was organized with a general introduction and purpose of teacher evaluation. Each of the seven standards of teacher evaluation was described and a summary of the research base for each standard was included. Standard 7: Student Academic Progress contained eight research references as the basis for inclusion of this standard in the evaluation tool. Each of the eight research references was analyzed and included in the literature review. The significance of the eight research articles cannot be overstated, combined the eight research articles serve as the foundation of the reasoning behind the state's implementation model and 40 percent of a

teacher's evaluation based on student academic progress. According to the VaDOE, the teaching standards required in the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* were linked to the performance standards for teachers with the Interstate Teacher Assessment and Support Consortium (INTASC) and the National Board for Professional Teaching Standards (NBPTS); however, "INTASC and NBPTS do not include measures of student academic progress in their standards/core propositions" (VDOE, 2008, *The Research Base for Uniform Performance Standards for Teachers*, p. 12). Therefore, a Summon search was conducted through the Virginia Tech Library to explore additional research articles that provided other views on the use of student academic progress as a component of teacher evaluation. Policy implementation literature, focusing on theories of production, was reviewed to gain practical information about how individuals and organizations influence policy implementation. Because the purpose of policy is to change behavior, understanding the influences encountered by individuals and organizations implementing new policy would be paramount for policy makers and practitioners. Literature on policy implementation in educational settings was selected by conducting a Summon search through the Virginia Tech Library.

Significance

The empirical research included in *The Research Base for Uniform Performance Standards for Teachers* (2008) was reviewed and analyzed because that research was identified by the VaDOE as the evidence supporting implementation of each of the seven performance standards. One of the primary purposes of the teacher performance standards identified in *The Research Base for Uniform Performance Standards for Teachers* was to optimize student learning and growth (p. 2). The research supporting the use of student academic progress as a significant component of teacher evaluation should support that purpose as identified in the adopted teaching performance standards. Further, this research was analyzed to identify practical applications for administrators and policy makers who are tasked with maximizing both learning for all students and professional growth for teachers. This focus will provide information for school leaders, and offer insight to policy makers as they consider the effects of law and state regulations. The Summon search conducted through the Virginia Tech library produced much research on VAM and teacher evaluation. But that same Summon search demonstrated a major research gap on the use of teacher created and administrator approved goals as the tool and mechanism for measuring student academic progress.

“Scholars have long recognized that local factors dominate the policy implementation process” (McLaughlin, 1987, 1990 as cited by Spillane, Gomez, and Mesler, 2009, p. 409). Therefore, policy implementation research focusing on education will provide valuable insight for policy makers and practitioners because practitioners’ behaviors and actions about a policy could facilitate successful implementation or contribute to roadblocks that potentially stall implementation. Further, an examination of theories of policy implementation provided a theoretical focus for the research topic, which could inform the actions of policy makers and practitioners by offering practical solutions that would guide school districts and leaders as they engage in teacher evaluation reform.

Research Questions

Research cited in the VaDOE document *The Research Base for Uniform Performance Standards for Teachers* (2008), additional research on the use of student achievement data, and policy implementation research was reviewed for confirmation that the following research questions need additional study:

The study will address the following research questions:

Questions for District Z Directors

1. What actions did District Z take during the 2012-2013 school year, the waiver year, to prepare for implementation of the new performance-based teacher evaluation system as outlined in *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*?
2. What professional development and/or training were provided for principals and teachers during the waiver year to prepare for implementation?

Questions for District Z Teacher Data Set (including teachers at School A, School B, and School C)

3. What are teachers’ perceptions of the new evaluation policy?
 - a. Did access to support, including professional development, increased knowledge and/or materials, and assistance from other educators affect teacher perceptions of the evaluation system?
 - b. If so, in what ways did access to support affect teacher perceptions?

Questions for Subset Teachers and Principals at School A, School B, and School C

4. How do teachers' perceptions at School A, School B, and School C compare with District Z teachers' perceptions?
5. What are the principals' perceptions of the new teacher evaluation system?
6. How do principals' perceptions compare with teacher perceptions?

The information provided from these research questions will offer practical guidance for policy makers and practitioners charged with drafting, implementing, and monitoring new reform efforts designed to target enhanced teacher performance with increased student achievement as an outcome.

Chapter Two

Review of Relevant Literature

The Research Base for Standard 7: Student Academic Progress

The Virginia Teacher Evaluation Work Group drafted the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* referencing research on teacher evaluation and student achievement. This research was summarized in the VDOE document *The Research Base for the Uniform Performance Standards for Teachers*. In this document the Work Group tied specific literature and research to each of the seven teacher performance standards. Standard 7: Student Academic Progress cited eight works as the research base supporting implementation of this standard. Each of the cited references was included in this literature review.

Sanders and Rivers (1996) published results of statistical analyses of student achievement for two school districts in Tennessee, one of the earliest states to experiment with student achievement as a measurement of teacher performance. In this study, the researchers conducted four separate analyses to investigate the cumulative effects of teachers on student academic achievement over grade levels. Sanders and Rivers suggested a teacher did have a cumulative effect on student achievement after an analysis of the different academic outcomes for students with similar ability and achievement levels. Further, they suggested that regardless of the quality of teachers a student had in later grades the effects of both very effective and ineffective teachers were measurable two years later. This finding trumpeted the importance of teacher quality by ensuring stakeholders understood the substantial impact over time of a teacher's effectiveness. Another finding was the lowest achieving students were the first to benefit as the effectiveness of the teacher increased. Sanders and Rivers recommended improvements to education and student achievement by focusing on student assignment in teacher's classes and improving formative teacher evaluation.

Aaronson, Barrow, and Sander (2007) published the results of a study analyzing the performance of math teachers and ninth grade students with the Chicago Public Schools. Teacher performance was measured using math achievement data from the Iowa Test of Basic Skills (ITBS) in eighth grade and the Test of Achievement and Proficiency (TAP) for ninth grade students. The researchers were able to match student performance data and records, classroom scheduling, and teacher administrative records as part of the statistical analysis. The resulting

data confirmed the quality of instruction provided by the teacher impacts the level of student achievement on standardized math assessments. Specifically, a one standard deviation improvement in math teacher quality resulted in approximately a .13, .20, and .13 grade equivalents increase for low-, middle-, and high-ability students (p. 126). These improvements resulted in 24 percent, 29 percent, and 6 percent average test scores gains between eighth and ninth grade for students (p. 126). Similar to the Sanders and Rivers study from 1996, Aaronson et al. found teachers rated highest in instructional quality had the largest gains in student achievement for those students at the lower end of the student performance distribution. Aaronson et al. argued that traditional human capital models did not explain the differences in student achievement and suggested teacher pay should be based on student achievement, particularly over the current model which emphasized education level and number of years of teaching. Aaronson et al. cautioned a model using achievement for teacher compensation would be ripe with implementation problems and would require extensive examination of the measurement tool used to evaluate student achievement (p. 132).

Hanushek, Kain, O'Brien, and Rivkin (2005) studied the variation in teacher quality using a semi-parametric approach based on value added to student achievement in a large urban Texas school district. Student achievement was measured by the criterion referenced Texas Assessment of Academic Skills (TAAS) tests, which were used to assess basic proficiency in math. However, the student achievement scores were standardized to permit comparisons of teachers within and between schools given very different initial test score distributions. This research highlighted, not just the differences that exist in teacher quality, but the importance of differentiation among those differences. Specifically, Hanushek et al. studied the size of the variation in teacher quality, the value of teacher experience and education to the variation in teacher quality, and the value of student teacher matching (p. 28). Hanushek et al. found the inexperience of first year teachers had a significant impact on student achievement. However, the experience/inexperience effects disappeared quickly after the first year of teaching. Another finding of the study was teachers who were able to increase achievement among lower achieving students were able to increase achievement in other areas of the achievement distribution (p. 28). Finally, Hanushek et al. found, when teacher quality was held constant in the value added model, students achieved more when taught by a teacher of the same race (p. 28). The researchers made significant progress in dissecting teacher attrition in an urban school district, dispelling

commonly held beliefs. Hanushek et al. found higher performing teachers were the least likely to exit the public schools completely. In fact, the findings of this study demonstrated it was the least effective teachers who were leaving teaching. Additionally, Hanushek et al. found those teachers who left the urban district for suburban school districts, which presumably have fewer minority students and higher student achievement, were as effective as those teachers who remained in the Texas urban school district at increasing student achievement. In other words, the best teachers were not the most likely to leave education or the most likely to leave schools with the most challenges to increasing student achievement. While this finding was limited to the population studied, this was an important finding as it dispelled a popular myth that the best teachers were leaving education.

Hattie (2003) studied the differences between teachers categorized as expert and experienced in a study of teachers applying for National Board for Professional Teaching Standards (NBPTS) certification throughout the United States, focusing on a subset of teachers who passed and did not pass the rigorous, comprehensive assessment of teaching ability. Hattie referenced previous work in which he conducted meta-analysis of over 500,000 research studies enabling him to identify the amount of variance among six influences on student achievement including students, home, school, principals, peer effects, and teachers. Significantly, Hattie's results yielded the largest controllable variance for student achievement was attributable to teachers. "It is what teachers know, do, and care about which is very powerful in this learning equation" (Hattie, 2003, p. 2). His research strongly advised that educators focus on this area, the influence of the teacher, instead of wasting time with other areas that have minimal influence or were completely out of control of educators. After a careful review of literature, Hattie claimed there were five major dimensions of expert teachers, distinguishing expert teachers from experienced teachers. The five dimensions of expert teachers were as follows: identifying essential representations of their subjects, guiding learning through classroom interactions, monitoring learning and providing feedback, attending to affective attributes, and influencing student outcome (Hattie, 2003, p. 5). Hattie dissected the five major dimensions of expert teachers into 16 characteristic attributes of expertise, furthering delineating the differences between expert and experienced teachers. The 16 characteristic attributes described the activities expert teachers did in each of the five dimensions.

The validity of the 16 attributes synthesized from over 500,000 research studies was then tested on a sample of 300 teachers with results reported on two groups of those teachers. Hattie measured the 16 attributes by having trained individuals examine student tasks, conduct observations, conduct interviews with the teachers and selected students, administer surveys, and collect and examine artifacts from the observed lessons. Hattie found differences in the mean ratings of expert and experienced teachers. Hattie graphed the effect size by dividing the difference in mean attribute ratings by the pooled standard deviation revealing the magnitude of the differences further. Finally, Hattie used a logical regression analysis to identify the best subset of dimensions. Three dimensions, challenge, deep representation, and monitoring and feedback, accounted for 80% of the difference between expert teachers and experienced teachers, signifying the importance of these dimensions. Hattie (2003) concluded expert teachers differ greatly from experienced teachers and educators would do well to focus on identifying, valuing, and encouraging excellent teachers (p. 15).

The overall theme of Hattie's study was to encourage educators to focus on what matters, which was the quality of the teacher. But more importantly, if stakeholders want to improve the quality of teachers and increase student achievement, the focus must be on the things that matter with such specificity that the focus, effort, and energy is spent on those qualities that actually do improve teaching. The results of this research suggested we do not have a comprehensive understanding of the dimensions and attributes that differentiate quality teaching. An implication of Hattie's research was that teacher quality, while assessed regularly in educational settings through annual observation and evaluation, might not be accurately measured using current evaluation practices. Further, without that comprehensive understanding of the dimensions and effective ways to differentiate performance on those dimensions the likelihood of improved teacher quality was diminished.

Rowan, Chiang, and Miller (1997) presented the findings of a study of the effects of teachers on student achievement. A unique feature of this research, at the time of publication, was that Rowan et al. studied how teachers' ability, motivation, and work situations combined to effect a student's achievement (p. 256). The sample population of this study included 5,381 students from 410 schools, 382 public schools and 28 Catholic schools. Student achievement was measured using students' National Educational Longitudinal Study (NELS) 10th grade mathematics achievement test. Rowan et al. asserted the NELS math test contained more

questions requiring higher order skills (p. 265). A number of variables were controlled in this study including courses taken by students, course track placement, student sex, students' socioeconomic status (SES), student motivation, and student educational expectations. Teacher ability was controlled based on a response to a single item designed to measure a teacher's math knowledge, whether or not the teacher was a math major, and the use of instructional strategies designed to promote critical thinking and problem solving. Teacher motivation was controlled using survey responses designed to measure the teachers' sense of efficacy and a survey measuring the teachers' expectations for students. Work situations were identified as school restructuring variables and were controlled using a survey assessment of the extent to which teachers rated control over decision-making, staff collaboration, and supportive leadership. Additionally, principals were required to identify the extent to which certain school restructuring variables were present in the school. The major findings of the study were as follows:

- A moderate correlation between a students' SES and achievement was noted prior to entering high school; however, achievement after entering high school was more highly correlated with prior achievement than with SES;
- Prior student achievement in math and other subjects as measured by the course track and taking variables positively impacted 10th grade math achievement;
- Student achievement was higher for students whose math teacher answered the one item math question correctly, and students who were taught by a teacher with a math degree performed better;
- The teachers use of instructional strategies utilizing critical thinking and problem solving did not positively impact student achievement;
- The teachers' expectations for specific students positively impacted student achievement, while the teachers' general motivation did not;
- Two school restructuring variables, teacher control over decision making and common planning, resulted in higher levels of mathematics achievement;
- And finally, larger schools produced positive effects on achievement. (p. 268 – 272).

A final analysis was conducted to determine any interaction effects from the variables. One noteworthy finding in this analysis was that as the average ability of students increased the effects of a teacher's ability and motivation decreased (p. 272). This finding was consistent with other researchers (Sanders & Rivers, 1996) in that lower achieving students make substantial

gains in performance based on the quality of the teacher. Rowan et al. (1997) noted that schools with low achieving students struggle to recruit effective teachers and maintain high morale among those teachers, making recruitment and morale a high priority for policy makers (p. 274). This research conclusion and suggestion could have focused the attention of policy makers on ways to recruit effective teachers and identify ways to keep those teachers in the schools where they were most needed.

Fuchs, Deno, and Mirkin (1984) examined the effects of data based program modification (DBPM) on the reading achievement of targeted students for 39 New York City Public School teachers. DBPM was a teaching method in which student goals, in this case the Individual Education Plan (IEP) goals, were monitored using a specific protocol at various intervals and adjustments to instruction were made based on individual student results. Participating teachers were assigned to a treatment and control group and were requested to select three to four students to participate in the program. The treatment teachers wrote curriculum based IEP goals and objectives, developed instructional goals designed to meet the IEP goal, monitored goals each week to assess progress towards the goal, and implemented instructional change if students did not meet the specified weekly performance standard. Control teachers were asked to set IEP goals and monitor those goals as the teacher preferred. Treatment teachers received training in DBPM while control teachers received training in general teaching strategies, managing student behaviors, the use of equipment, and managing support staff. While the content of the training was substantially different and the treatment teachers received more training time, each group of teachers received ongoing, job-embedded professional development. Further, the statistical analysis of the difference in the number training hours between the two groups of teachers was deemed insignificant.

Fuchs et al. (1984) found the students of teachers implementing the DBPM outperformed, on both reading progress and IEP goal progress, the students of teachers who used traditional methods of monitoring IEP goal progress (p. 456). Additionally, Fuchs et al. found the use of DBPM affected how the teachers taught. Specifically, treatment teachers increased the structure of lessons while control teachers decreased the structure lessons (p. 457). This finding was supported in previous research, which stated learning-disabled students' achievement increased as the teaching and classroom structure increased. Teachers utilizing DBPM were more realistic about student performance and achievement and responded to student needs more quickly by

making changes to goals and adjusting instruction. Students of the treatment teachers were more knowledgeable of their goals and were able to use data to make predictions about their own performance (p. 458). Fuchs et al. hypothesized that students' knowledge of performance could affect school performance and could be a significant marker of individual student achievement (p. 458). These findings suggested the importance of a structured plan of cyclical instruction. An important feature of this study was the inclusion of the student's assessment of learning and progress.

Stecker, Fuchs, and Fuchs (2005) reviewed research on curriculum-based measurement (CBM) and published a synthesis of the results of the research studies conducted in the past 25 years. The purpose of CBM was to provide special education teachers with an efficient and adequate measurement system to track student progress on basic skills and improve the quality of instruction for students through curriculum monitoring with instructional adjustments based on individual progress (p. 795). CBM was a formalized measurement system with similarities to Fuchs et al. (1984) DBPM model. Stecker et al. conducted this review of the research to identify the factors that impact student achievement when teachers used CBM to monitor student progress in reading and mathematics. Research results were reported for special education and general education students. Stecker et al. concluded that a teachers' use of CBM lead to significant student achievement growth for special education students if the teachers implemented program modifications when student performance results met the criteria for changing the instructional program. Another conclusion from the body of research was that when teachers used database decision rules those teachers were more responsive to student needs (p. 816). In other words, when the teachers followed the program with fidelity, assessing the required number of times, charting the results using a student goal line, and modifying the program when student performance did not exceed the established benchmark, those teachers provided individualized instruction for the students that resulted in student growth. A third conclusion presented by Stecker et al. was that the use of computer applications to chart and interpret data supported the implementation of the CBM decision rules and led to student growth. The use of computer applications led to increased teacher satisfaction and Stecker et al. believed the software reduced the time required of teachers to complete CBM procedures (p. 816). Computer application programs also provided teachers with detailed skills analysis of student performance, and skills analysis was found to be help teachers when they planned instructional

modifications if a student did not reach the performance standard and a decision rule was implemented. But it was noted that skills analysis without ongoing support would not lead to student growth. Stecker et al. also reviewed the research of CBM performance results with general education students. Teachers who used CBM with general education students also experienced greater student achievement than teachers who used their own methods for monitoring student performance. While the results of research studies involving general education students included more variables and were not considered large gains, the use of CBM did produce positive growth for all groups of students. Stecker's et al. review of CBM research confirmed that teachers need support when implementing instructional change but the payoff for continued support would lead to individually responsive instruction.

Stronge, Ward, Tucker, and Hindman (2007) identified key differences between the most effective and least effective teachers based on student gains as measured by the Virginia Standards of Learning (SOL) assessments. The research study used a regression analysis of student performance data on the SOL tests of 1,936 students from 85 classrooms in an urban school district located in Virginia, controlling for student-level and class/school-level variables, to identify the teachers whose students performed beyond expectation, at expectation, and below expectation on the SOL tests (p. 171). Based on this range of teacher performance expectation, 11 teachers, five top performing teachers and six low performing teachers, were selected to highlight the differences in instructional practices and characteristics among the two groups of teachers. Trained observers evaluated the 11 teachers' instructional practices during observations, and the observers did not know the teachers were categorized based on the effective and ineffective student achievement gains on SOL scores. While a limitation of this study was the small sample size of 11 teachers, three differences were identified between the two groups of teachers. Teachers whose students performed higher than expected were able to differentiate instruction and materials, asked a greater number of higher level Blooms' questions, and had fewer incidents of disruptive behavior during observations (p. 180). Stronge et al. assumed "assessments were used that were closely aligned with the curriculum taught by the teachers, which allowed for a meaningful interpretation of student learning gains" (p. 180). Stronge et al. believed curriculum based student achievement information could be used to identify the practices of effective and ineffective teachers in an effort to improve student achievement (p. 181). Further, Stronge et al. supported the use of student achievement data as an

indicator of effectiveness for schools, teachers, and principals (p. 181). Important findings from this research study would anchor the development of the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, as evidenced by the inclusion of Standard 7: Student Academic Progress as a weighted component of the new evaluation model.

Emerging Themes

Three key themes emerged from the cited references. The first theme was that differences exist among teachers and the quality of instruction provided to students varies based on student achievement data. The second theme among the research was that lower achieving students appear to benefit the most in achievement gains when the quality of instruction was improved. Finally, some teaching strategies and programs, with structured monitoring and feedback for students, and professional development for teachers led to increased student achievement both with general education and special education students.

Additional Research on Teacher Evaluation, Value-Added Modeling, and Student Academic Progress

Much of the research supporting the use of student achievement as a component of teacher evaluation was critiqued in the literature through a statistical model designed to quantify an individual teacher's contribution to each student's achievement, known as VAM. According to Koretz (2008), VAM in the research literature often, "refer to two very different quantities" (p. 19). The first quantity is the amount of total growth in student achievement that occurred during a specified time period, usually the beginning of the school year until state accountability testing at the end of the school year. The second quantity Koretz believed VAM referred to was how much value the individual teacher added to the individual student's achievement during the specified time period. The distinctions provided by Koretz were important to understand how VAM could be applied practically and in determining which quantity was referred to the literature.

Braun (2005) presented information for policy makers and educators about the role of VAM in teacher evaluation systems. Braun conceded the appeal of VAM was based partly on the simplistic logic that student learning provides substantial information about the overall quality of the teacher. Further, Braun noted the improvement of VAM over a model in which all students must achieve a specified criterion, such as with Adequate Yearly Progress (AYP) a provision of

the NCLB Act. Braun supported the use of VAM because VAM provided valuable information about teachers that could be used in conjunction with observations, interviews, and artifacts to make decisions about professional development and school improvement needs. Undoubtedly, the information provided by VAM brings objectivity to teacher evaluation that was appealing to policy makers.

This policy information report focused not only on the strengths of VAM but on the weaknesses as well. Braun (2005) defined VAM as the difference between a given teacher's performance from the performance of the typical teacher, with respect to the average growth realized by the students in their classes, stating further that VAM "is often referred to as a measure of teacher effectiveness" (p. 7). Because the statistical model of VAM attempted to establish a causal relationship, Braun noted the primary problem with VAM is the fundamental lack of randomness found within schools and classrooms. "The fundamental concern is that, if making causal attributions is the goal, then no statistical model, however complex, and no method of analysis, however sophisticated, can fully compensate for the lack of randomization" (Braun, 2005, p. 8). Braun cited additional concerns based on lack of randomization with VAM including inappropriate attribution, the negative impact of missing data, differences among classes at the secondary levels, and the variance of teacher effects from year to year. Most notably, Braun concluded VAM should not be the sole criteria used when making consequential decisions about teacher effectiveness, but when VAM was used with other measures, this tool could assist administrators in identifying which teachers were most effective and which teachers would need support.

The perspective presented by Braun (2005) on VAM could easily be interpreted as a negative review of the use of VAM in teacher evaluation. However, that assumption would not be valid. Braun carefully evaluated and explained the use of VAM, focusing primarily on the use of Educational Value-Added Assessment System (EVAAS). While the statistics of VAM as the sole determinant of teacher effectiveness would not be advisable, Braun noted the value of the quantitative information as one component administrators and policy makers could use in an effort to increase student achievement.

Like Braun (2005), Koretz (2008) suggested potential benefits to the use of VAM, particularly when compared to the current models of evaluating student achievement – the status model and the cohort-to-cohort change models. Koretz also suggested that the use of VAM was

based on logic that made sense to educators and the public, specifically a given teacher's effectiveness would be evaluated based on the growth of the students and how much the given teacher's efforts contributed to that growth. According to Koretz, a primary problem with VAM rested with the sheer complexity of the technical application and how so few practitioners and policy makers understood how technical decisions could influence the results. It was suggested more research on the types of statistical models used in VAM could clarify some of the uncertainties of VAM. Further, Koretz stated, "we must accept the fact that value-added models, taken by themselves, are not an adequate measure of overall educational quality" (p. 39). Noting further, the information provided by VAM offered another source of information about student achievement, which should not be considered in isolation as a measure of teacher and/or school effectiveness.

Koretz (2008) identified appropriate uses for VAM in education. It was suggested VAM would be more effective with elementary reading and math because of the use of vertically scaled tests. Additionally, Koretz concluded states would need to develop assessments based on a vertical scale. Certainly, this conclusion implied substantial economic ramifications for states that have seen education funding expenditures decreased as a result of the national economic crisis in 2008. Another conclusion Koretz reported was that more testing data would be needed to reduce the imprecision of VAM, specifically Koretz suggested including more years of testing data and more data from other sources. Koretz concluded VAM should be used to compare classes and schools where students were performing at similar levels to increase the trustworthiness of the results. Koretz concluded the use of test scores as the only element included in any accountability system would be a mistake especially given the heightened importance of those tests for schools and teachers. Koretz's final suggestion was to recommend more effort and resources be placed on evaluating the accountability systems states and policy makers have established to assess the performance of schools and teachers, suggesting that the failure to carefully evaluate the intended and unintended outcomes of any accountability system places desired educational outcomes at risk. Koretz and Braun (2005) both supported the use of VAM with full disclosure of the limitations of the models. Neither researcher suggested the wholesale adoption of VAM as the sole determiner of a measure of teacher quality and effectiveness.

Braun, Chudowsky, and Koenig (2010) served as members of a steering committee tasked with planning and organizing a workshop on the research base surrounding the use of value-added methods in teacher evaluation in an effort to inform policy makers. The workshop was sponsored by the National Research Council (NRC) and the National Academy of Education and funded by the Carnegie Corporation. Braun et al. drafted a written report that summarized the discussions of the workshop. Braun et al. defined value added as “a methodology referring to efforts to measure the effects on the achievement of students of their current teachers, schools, or education programs” (p. 1).

Braun et al. (2010) classified the concerns of the workshop presenters with VAM into three categories:

- 1) concerns about the models themselves including transparency and robustness to violations of assumptions; 2) concerns about the test data that feed into the models including reliability, validity, and scaling; and 3) concerns about statistical characteristics of the results including precision and bias. (Braun, et al., p. 56)

Braun et al. reported the expert presenters at the workshop offered a range of perspectives on the use of VAM. The most conservative perspective presented suggested VAM was more problematic than current status measures and should only be used for low-stakes purposes, such as identifying which teachers need professional development. While the more liberal perspective recommended using VAM as part of teacher evaluation, the presenters still recommended the use of VAM only in combination with other indicators of teacher effectiveness. Despite the concerns expressed by presenters at the workshop, Braun et al. suggested an overall enthusiasm among the presenters for the use of VAM, even if only in limited capacities, because of the valuable information that could potentially be provided to policy makers.

A substantial hurdle VAM must effectively neutralize was implementation expense in an industry that remains strapped for cash. Braun et al. (2010) reported on implementation requirements and suggestions that would maximize the benefits of VAM for school districts. In order to implement VAM, a school, district, and/or state must have access to a longitudinal database that tracks individual students over time and links those students to the teachers, a means for managing missing data that accounts for student mobility, and highly qualified technical staff to administer and monitor the system (p. 58). These three requirements represented a substantial financial commitment in both human and technical capital.

Alternatively, the preferred VAM would also need a sequence of tests vertically aligned to the standards, curriculum, and instructional strategies, a reporting system that presented information in a way that users could understand and analyze, an ongoing training program for teachers and administrators, and a way to monitor and evaluate the program's effects and adapt as those effects necessitated (p. 59). Given the urgent need for increased resources in school districts across the nation, it remains to be seen if the will and financial resources will be made available to support successful implementation of VAM based on the required and suggested conditions as presented by Braun et al. (2010).

White (2004) published the results of a study of the relationship between teacher evaluation scores and student achievement. The Coventry School District in Rhode Island used a standards based teacher evaluation developed from the evaluation standards and rubric used in Danielson's *Enhancing Professional Practice: A Framework for Teaching* (Danielson, 1996). The purpose of White's study was to describe the relationship between a teacher's overall evaluation rating and student achievement. The research findings indicated the teacher evaluation tool that was used identified the teachers who produced higher than expected student achievement in reading. Further, White found student reading achievement was greater if the teacher was rated Distinguished, as opposed to the student reading achievement of teachers who were rated Proficient. However, no relationship existed between the evaluation scores and math achievement. White concluded standards-based teacher evaluation ratings were an indicator of teacher quality as related to student achievement even when controlling for teacher experience. White noted a substantial limitation of the findings of this study was the small sample size. Because the selected methodology required consecutive years of student assessments in reading and mathematics and a teacher evaluation rating, one of the restrictions of the methodology resulted in a small sample size for this study. While the findings did find a relationship between teacher quality, as evidenced by the higher evaluation rating and higher student reading achievement, those findings should be cautiously interpreted.

Nye, Konstantopoulos, and Hedges (2004) used data from the Tennessee Class Size Experiment or Project Student-Teacher Achievement Ratio (STAR) to conduct a randomized experiment of within school variance of teacher effects. Nye et al. suggested the results of this study "provide stronger evidence about teacher effects" (p. 253) because of the use of random assignment of students in Kindergarten through third grades to treatment conditions of small

class size, regular class size, or regular class size with a classroom aide. This assertion was made because the two primary types of research used to measure the effects of a teacher were education production function studies and variation studies that controlled for student background. Because neither type of research utilized random assignment, Nye et al. believed the results of the current study using Project STAR student achievement data could provide important information to policy makers regarding teacher effectiveness. A major limitation of this study was the sample size because of the methodology. Specifically, randomization requirements demanded that only those schools with at least four classes at a given grade level could be included in the sample. Results of the study should be interpreted based on this limitation.

Nye et al. (2004) found substantial differences in the student achievement of teachers experiencing the same treatment condition within schools. Specifically, Nye et al. reported a teacher in the highest quartile of effectiveness produced approximately one third of a standard deviation increase in reading and a one half standard deviation increase in mathematics achievement, when compared to a teacher in the lowest quartile of effectiveness (p. 253). Student achievement gains were consistently higher in reading and mathematics as the teacher's effectiveness increased, in other words, as the teacher's placement on the standardized effectiveness ranking increased so did student achievement. Nye et al. suggested policy makers should consider interventions designed to identify and improve teacher effectiveness particularly when considering most student achievement measures utilized by states were more aligned with curriculum and instruction than the Stanford Achievement Test (SAT) used with Project STAR. Given that current reform measures focused on school effects and sanctions such as school choice, Nye et al. suggested interventions targeting teacher effectiveness could lead to increased student achievement. Another finding of Nye et al. (2004) was teacher effects were greater in schools with a lower SES than a higher SES, suggesting which teacher a student received in lower SES schools could potentially impact that students' achievement greatly. This finding supported Nye's et al. earlier conclusion that policy makers should focus on identifying and improving teacher effectiveness over implementing sanctions against schools.

Views on the use of student achievement and performance represented a growing body of research. Notably, researchers voiced concerns about the stability of the statistical models, the overall fairness of use, and a concern that use of VAM could deter educators from focusing on

educational practices proven to improve instruction. Darling-Hammond, Beardsley, Heartel, and Rothstein (2011) identified three major problems with VAM when they were promoted as measuring teacher effectiveness. First, Darling-Hammond et al. noted, “value-added models of teacher effectiveness are highly unstable” (p. 2). This assumption was based on the fluctuations in student performance for the same teachers from year to year, and fluctuations in student performance based on the type of assessment used to measure performance growth. Second, Darling-Hammond et al. noted that the students assigned to a teacher significantly impacted the teacher's value-added rating (p. 3). Students were not assigned to teachers randomly. Further, the higher the grade level of a student in the American education system the less random student assignment becomes. Specifically, students in middle and high school are regularly grouped by achievement levels based on the diversity of course selections offered. For example, a student in 11th grade in the United States would likely have the option of enrolling in English 11, Advanced English 11, and Advanced Placement (AP) English 11. It is easy to understand why a given teacher would want to teach AP English 11, as opposed to the other levels, if that teacher's evaluation was based on a VAM of student academic progress. Darling-Hammond et al. surmised a teacher's evaluation was influenced by the students assigned to that teacher despite efforts of some VAM to control for prior achievement and student demographic variables (p. 3). Finally, Darling-Hammond et al. noted “value-added ratings cannot disentangle the many influences on student progress” (p. 4). Student achievement was influenced by a number of factors that were not directly attributable to a given teacher. Studies suggested value-added ratings did not align with ratings received by trained observers and administrators. Based on these three reasons, Darling-Hammond et al. suggested that VAM were not appropriate as a primary means of evaluating teachers. Darling-Hammond et al. promoted the use of other mechanisms such as video taped lessons and authentic artifacts of teaching as a more appropriate measure of a teacher's ability to impact student achievement and academic progress.

Similarly, Haladyna (2011) believed a more appropriate accountability system must consider the actions of all who have an impact on student achievement including students, parents, teachers, school leaders, local and state school boards, the governor, and state and federal lawmakers. Measuring student achievement was a burden that must be shared with all stakeholders having an impact on that achievement, not just the teacher. Because students were not equal and came to school with varying levels of ability and support, evaluating teachers

based on student achievement tests would not be a fair or equitable measure of a teacher's ability. "Imagine coaching a baseball team with the worst players, and being fired because they lose games" (Haladyna, p. 2). Haladyna noted problems with using single tests to measure performance including cheating on the test, narrowing the curriculum, and the tests were not designed to measure teacher effectiveness. Like Darling-Hammond et al. (2011), Haladyna suggested multiple measures to assess teacher effectiveness and avoiding student achievement tests as a means of measuring teacher performance.

Baker, Barton, Darling-Hammond, Haertel, Ladd, Linn, Ravitch, Rothstein, Shavelson, and Shepard (2010) evaluated the evidence regarding the use of VAM and teacher effectiveness, and concluded that while VAM was preferred to current methods of identifying the effectiveness of teachers, states and school districts should use caution when weighting student achievement heavily as a component of teacher evaluations. Alternatively, the researchers suggested standards-based evaluations of teaching practice such as those used by the National Board of Professional Teaching Standards and the beginning teacher assessment systems in Connecticut and California, as both evaluation systems offered substantial promise for increasing teacher quality. These models were noted for the use of research-based techniques about teaching and learning including systematic observation protocols, observations and videotapes of classroom instruction, teacher interviews, and artifacts such as lesson plans, assignments, and student work samples.

Jacob and Lefgren (2008) presented the results of the empirical quantitative study addressing how well principals, in a midsize school district in the western part of the United States, were able to "distinguish between less and more effective teachers, where effectiveness was measured by the ability to raise student math and reading achievement" (p. 103). The unit of analysis for the study was the teacher with the sample consisting of 221 elementary teachers in grades two through six who were teaching a core subject during the 2002-2003 school years. Principals were given a survey requiring them to rate the teachers in their own schools, not only on their overall effectiveness, but also on specific performance characteristics. These characteristics consisted of classroom management, dedication and work ethic, parent satisfaction, ability to raise math and reading achievement, and relationship with administrators. The results of the survey were then compared to student standardized test scores in the areas of

math and reading to establish whether the principals' ratings of their teachers were correlated to the trends seen in the testing data as related to increasing student achievement.

Jacob and Lefgren (2008) specifically found principals were highly skilled at identifying those teachers who produce the largest and smallest standardized achievement gains in their schools. A further finding was that principals were much less skilled at distinguishing the teaching characteristics of teachers in the middle of the distribution. Finally, Jacob and Lefgren found that when comparing subjective principal assessments of teachers to traditional determinants of teacher compensation, education and experience, subjective principal assessments were a better predictor of future student achievement. This was important because despite concerns of bias in subjective assessments the findings suggested that using principals' subjective assessments to determine compensation and promotion could increase student achievement (p. 103).

Finnigan and Gross (2007) conducted a research study examining the influence of accountability policies on teacher motivation. The 1995 School Reform Act in Illinois established an accountability system utilizing the same premises as the NCLB Act of 2001, specifically sanctions and consequences motivate schools, administrators, and teachers to improve performance and increase student achievement. The findings of this study were based on data from teachers in 10 low-performing schools in the Chicago Public Schools. Each of the 10 schools was designated with a probation status rating based on the criteria of the 1995 School Reform Act. Finnigan and Gross found a connection between morale and motivation, specifically in schools that struggled the most, teacher motivation decreased as a result of the sanctions. Additionally, Finnigan and Gross found teachers' expectations of improving decreased the longer a school and teachers remained on probation status. Finnigan and Gross suggested policy makers should consider carefully the negative impact of sanctions on teacher motivation when developing school accountability measures. While this study focused on school accountability and the impact of threats of sanctions on teachers' motivation levels, the organizational theories of motivation and expectancy could easily be applied to the use of student achievement as part of a teacher evaluation system.

Lockwood, McCaffrey, Hamilton, Stecher, Le, and Martinez (2007) used four years of student level data linked to teachers from a large school district to examine VAM's ability to consistently evaluate teacher effectiveness when using two different subscales of a single

standardized math assessment, specifically the Procedures and Problem Solving subscales of the Stanford 9 mathematics assessment. Lockwood et al. found the teacher's effectiveness as measured with VAM was dependent on the subscale of the Stanford 9 mathematics assessment that was used. Specifically, student performance differed greatly on the Procedures and Problem Solving subscales, generating concern when using the results of VAM to make critical decisions about a teacher's performance. Lockwood et al. noted that if consistencies among scores were limited to those different from zero, "only 26% for year 2 and 16% for year 3 have classifications insensitive to weighting" (p. 59).

While the findings noted above raised caution regarding both the content and type of assessment given, Lockwood et al. (2007) found VAM to be consistent regardless of the model or form used when removing the effects of student background variables beyond the control of teachers, specifically SES, race, limited English proficiency, special education status, gender, and age. This finding was encouraging because policy makers and educators must be confident the results of VAM provide information relevant to the specific teacher's performance if that information was to be effectively used in evaluation or as a tool to identify how to improve teacher performance. Lockwood et al. agreed with Nye, et al. (2004) that when the state assessments were aligned with the school curricula, VAM could offer more sensitivity to teacher effects and suggested future research could examine the validity of this hypothesis. Lockwood et al. concluded VAM results should be interpreted with caution, especially if those results would be used for high stakes purposes. Further, Lockwood et al. cautioned that policy makers must carefully examine the assessment used in the VAM, how the assessments aligned with the curriculum and instructional approaches, and how the selected model of VAM used was sensitive to the assessments to ensure results provided consistent information about teachers.

Marzano and Toth (2013) noted that existing practices of teacher evaluation failed to focus on the specific teaching skills needed to improve instruction and increase student achievement. Ericsson, Krampe, and Tesch-Romer (1993) defined deliberate practice as a method to improve performance on a task in which an individual would practice the specific skill, receive targeted feedback on performance, and then conduct further practice. Marzano and Toth purported that Ericsson's et al. model of deliberate practice could be used as the foundation of any evaluation system and would meet the educational needs of current economies. While the primary purpose of this policy report was to promote a teacher growth model similar to the

Marzano Teacher Evaluation Model drafted by the authors of the report, several key features of teacher evaluation were identified as fundamental requirements for teacher and student growth. First, Marzano and Toth suggested an evaluation model must have the specificity to outline exactly what a teacher needs to improve upon and communicate the ways to make those improvements. In order to provide this high level of specific feedback, a substantial amount of training would be required for all observers, whether administrators, mentors, or peers. Next, Marzano and Toth suggested the teacher evaluation model must measure teacher growth through the use of research-based classroom strategies aligned to the curriculum. Finally, Marzano and Toth supported the use of student achievement as a measurement to demonstrate teacher growth and effectiveness. Each of these fundamental requirements must be implemented with deliberate practice, as defined by Ericsson et al. (1993) with a deliberate practice model, of practice – feedback – practice, to ensure increased teacher effectiveness and student achievement.

Emerging Themes

The additional research on teacher evaluation, VAM, and student academic progress produced three emerging themes. First, VAM when used with other performance indicators could improve the overall effectiveness of a teacher evaluation system. Next, teacher evaluation models that included research based performance standards could improve teacher quality. Finally, researchers identified specific concerns with the use of VAM in teacher evaluation that included the concerns about the model used, the test data, and the interpretation of the results.

Policy Implementation Theories and Research

Much of the research on policy implementation in education was grounded in production theory, which assumed the alignment of organizational and individual values and beliefs resulted in increased performance (Akerlof & Kranton, 2005; Youngs, Frank, Thum, & Low, 2012). Additionally, new theories focusing on how organizations prepare and implement those policies offered valuable information on the importance of setting, formal and informal leaders, and the role of resources and support for schools as they implement a new policy (Coburn, 2001; Little, 1982; Spillane et al., 2009; Sun, Frank, Penuel, & Kim, 2013). Little (1982) summarized organizational perspectives of policy implementation as requiring detailed attention to the specific characteristics of the social organization where change would be attempted. Given this consensus among the researchers, policy makers and practitioners should not underestimate the

importance of individual behavior, the setting, and the organizations ability to manage implementation with appropriate information and utilization of resources.

Akerlof & Kranton (2005) presented research arguing for the use of identity, as defined as an employees attachment to the organization and willingness to behave in certain ways, as an important component in a new economic model of work incentives. While the research focused on a military model, Akerlof and Kranton believed this conclusion was applicable to schools and businesses. Akerlof and Kranton found a worker who identified with an organization required less incentive pay, translating to less reward and punishment to achieve high performance on a job. Workers were classified as insiders and outsiders based on like-mindedness and responsibility felt toward the organization. The researchers concluded an organization would benefit greatly from investments in the identity of its workers, specifically investments targeted to align the identity of workers with the identity of the organization.

Youngs, Frank, Thum, and Low (2012) studied mentor and collaboration programs to determine what influence those programs produced in the commitment level and instructional ability of new teachers. Youngs et al. provided additional validity to the theory that like-mindedness of beliefs and practices among new teachers and mentors, subgroups, or other colleagues promoted student learning. Further, Youngs et al. found a lack of alignment among new teachers and those providing the influence resulted in negative effects on the commitment level, quality of instruction, and effectiveness of new teachers.

In addition to theories focusing on the relationship of the individuals and the organization, theories focusing on how organizations prepare for implementation of policy were identified as an important component in the successful implementation of policies. Coburn (2001) analyzed how a school and the teachers implemented new reading policies. Coburn found teachers were continuously constructing meaning from the messages provided in their environment. Those messages shaped their beliefs and practices as they implemented the new reading policy. Both formal and informal actors within the school community influenced teachers, and while some values and beliefs were shared, it was possible for a teacher to be influenced differently even if exposed to the same influence (Coburn, p. 162). Coburn concluded that formal policy was one of many influences on teacher behavior, finding that influences could come from a variety of sources.

Little (1982) conducted a qualitative study designed to identify and characterize the workplace characteristics of successful, adaptable schools. Little found schools, as organizations, were a powerful influence on individual behavior and the success of the school. Moreover, in successful, adaptive schools there was a deep commitment for professional improvement of the organization, as opposed to the professional improvement of individuals (p. 338). Little found two primary expectations among successful schools including a norm of collegiality and a norm of continuous improvement. Little concluded norms of collegiality, receptivity to staff development, and instructional leadership were critical practices of schools with a powerful influence on the behavior of individuals and the success of the school.

Spillane, Gomez, & Mesler (2009) suggested the influence of organizations, both the school and LEA, on policy implementation could only be understood by examining work practice, resource allocation, and how individuals access and activate organizational resources. Further, Spillane et al. concluded policy implementation theory should focus on the interactions of individuals and organizations as they occurred because much action was dependent on the emergent situations encountered by individuals. Human capital, social capital, technology, and organizational routines were identified as the key resources individuals could access and activate, and each of these key resources could substantially influence policy implementation (Spillane et al. p. 410). This research connected the influence of the organization and its resources and how teachers access those resources. Further, Spillane et al. defined organization as departments, schools, district offices, and state and federal offices; suggesting that practice and policy implementation could be examined fully only when the influences of all organizations were considered.

Sun, Frank, Penuel, & Kim (2013) investigated the role of formal and informal leadership as a function of external policy implementation. Sun et al. found formal leaders influence general practices of teachers by setting standards, selecting materials, and assessing student progress, while informal leaders influenced teachers pedagogical practices when teaching reading. This finding that alignment of values, beliefs and practices of formal and informal leaders enhanced teacher performance and potentially student achievement was substantiated in previous research (Akerlof & Kranton, 2005). Sun et al. concluded policy makers should consider the influence of leadership from a variety of sources within organizations. Additionally, the influence of aligned leadership from a variety of sources magnified the powerful influence on behavior and change.

The researchers offered practical suggestions to support successful implementation of external reforms at schools including coordinating formal and informal leaders' influences, providing guidelines for collaboration among teachers and leaders, and targeted professional development emphasizing diverse skills and content designed to expand the groups' knowledge base (p. 634). Sun's et al. study suggested that individual actors within organizations influenced different areas of a teacher's practice, and regardless of the status of the influence, formal and informal leaders influenced policy implementation through their interactions with teachers.

Emerging Themes

The policy implementation research was organized around two policy implementation theories. This research produced two emerging themes. First, identity and like-mindedness of beliefs and practices promoted student learning. Suggesting new policy would be more positively perceived at schools when teachers thought their own beliefs and values were aligned with those beliefs and values of the school. This positive perception of teachers could equate to greater ease when implementing new policies. The final theme was that individuals continuously interact with the organization, and organizations have many different influences.

Synthesis and Conclusions

Years of research have convinced educators and policy makers that effective teachers contribute to higher student achievement. Stronge et al. (2007) added information to the existing research by focusing on the differences between effective and less effective teachers. The key question for educators and policy makers remains how best to measure teacher effectiveness and how best to utilize that information and data to improve teacher effectiveness.

Researchers (Stronge et al, 2007, Hattie, 2003, and Baker et al. 2010) concluded there were differences between the instructional behaviors of effective teachers and those categorized as less effective. The Stronge et al. (2007) study produced three key findings of the differences among effective and less effective teachers through the extensive, time-intensive observation and interview protocols of the 11 teachers. Additional research clarified the role of student achievement data in teacher evaluation. Specifically, the consensus among reviewed literature suggested student achievement data were just one of the ways to measure teacher effectiveness, particularly if VAM was used to produce the student achievement data. If a school district

adopted a conservative view and implemented VAM, the school district would be wise to use VAM in conjunction with other methods to measure a teacher's effectiveness.

Few school districts and states have committed the funding to education that would be required to reduce the span of control regarding teacher observation and evaluation, enabling such extensive observation of all teachers. Yet, Virginia law requires teacher evaluation to be consistent with the standards outlined in the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, and that student academic progress be a significant component of a teacher's overall summative evaluation rating. With the implementation of the new model requiring 40 percent of the evaluation to be based on student academic progress, school districts are using teacher goals approved by site principals as the tool to measure the 40 percent component of the evaluation. This review of literature provided little guidance regarding the use of teacher goals to measure student academic progress. Additional research regarding the use of teacher goals, specifically SMART goals as a measure of student academic progress, could contribute greatly to the body of research on teacher quality and student achievement. The literature reviewed confirmed policy was never implemented in a vacuum but was influenced by the individual, the setting, the organization, resources, and access to resources. A review of research on policy implementation suggested policy gained meaning when policy was practiced. Further, policy theories outlined the importance of the alignment of values and beliefs of teachers and organizations combined with access to information, resources, and the ability to activate those resources as critical features in the practice of any policy. This literature review suggested the need to understand teachers' values and beliefs as related to the new teacher evaluation system and the need to identify what factors influenced those values and beliefs in an effort to understand policy implementation.

Chapter Three

Methodology

Introduction

The proposed study was a subset of a larger research project studying the implementation of performance-based teacher evaluation system in Virginia, which was directed by Professor Min Sun at Virginia Polytechnic Institute and State University. The purposes of this study were to collect empirical evidence on the implementation of the new performance-based teacher evaluation system by comparing teachers' and principals' perceptions at three elementary schools to teachers' perceptions in the school district, identified as District Z, and were to add to the research base on policy implementation theories and teacher evaluation. Section 22.1-253.13.5 of the Virginia Code required the use of consistent standards, as identified in the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*, and that student academic progress be a significant component of the evaluation and overall summative rating.

School districts in Virginia retained authority to modify the evaluation used for teachers as long as the standards and weighting remained consistent with the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* document. Eight school districts out of 134 in Virginia elected to apply for an implementation waiver for the 2012-2013 school year. District Z was a waiver district with regard to implementation of the new teacher evaluation system. As a result, the 2013-2014 school year was the implementation year of the new evaluation system.

Policy implementation research confirmed two emergent themes in the review of relevant literature. Specifically, like-mindedness of beliefs and values within an organization leads to increased student performance and teacher effectiveness, and a teacher's access and activation of organizational resources impact policy implementation. Based on these themes, it would be expected teachers who work at schools where beliefs were aligned and teachers who received support, including support from administration and colleagues, additional resources, and professional development, would experience more favorable perceptions about the new performance based teacher evaluation system.

This study adopted a mixed-method design with both quantitative and qualitative information used to describe and understand the new teacher evaluation system in District Z and

at three elementary schools in District Z. The proposed study examined closely teacher perceptions at three elementary schools and compared those perceptions to District Z teacher perceptions. Additionally, the teacher perceptions at the three elementary schools were compared to the perceptions of the principals of the three elementary schools. Two of the elementary schools, School A and School B, were selected for additional study based on similarities in local, state, and federal reporting categories. The third elementary school, School C, was selected for additional study based on differences from School A and School B on the same state and federal reporting categories. The careful selection of the three elementary schools added depth and richness to the proposed study because teachers' perceptions were analyzed at schools with similar and dissimilar characteristics. A mixed-methods design was selected to provide structure and empirical measurement with a description of preparation activities in the district and principal perceptions on matched survey instrument questions. McMillan and Wergin (2010) noted that a mixed-method design "combines the best features of both approaches" (p. 134). The purpose of the study was to answer a series of research questions at the district level, site level, and administrative level in an effort to add to the existing research on policy implementation and teacher evaluation.

The study addressed the following research questions:

Questions for District Z Directors of Human Resources and Instruction

1. What actions did District Z take during the 2012-2013 school year, the waiver year, to prepare for implementation of the new performance-based teacher evaluation system as outlined in *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*?
2. What professional development and/or training were provided for principals and teachers during the waiver year to prepare for implementation?

Questions for District Z Teacher Data Set (including teachers at School A, School B, and School C)

3. What are teachers' perceptions of the new evaluation policy?
 - a. Did access to support, including professional development, increased knowledge and/or materials, and assistance from other educators affect teacher perceptions of the evaluation system?
 - b. If so, in what ways did access to support affect teacher perceptions?

Questions for Subset Teachers and Principals at School A, School B, and School C

4. How do teachers' perceptions at School A, School B, and School C compare with District Z teachers' perceptions?
5. What are the principals' perceptions of the new teacher evaluation system?
6. How do principals' perceptions compare with teacher perceptions?

All teachers in District Z were requested to participate in the proposed study by completing an electronic survey instrument designed to answer research question three. The survey instrument results of the subset group of teachers from School A, School B, and School C were the focus of research questions four and five. The principals of School A, School B, and School C were interviewed to answer research questions five and six. During the interviews, principals were asked matched questions from the teacher survey instrument to compare with teacher responses on the electronic survey instrument. Specifically, ten teacher survey instrument questions were asked in principal interviews. The Directors of Instruction and Human Resources at District Z were interviewed to answer research questions one and two.

Population

The population for this study will be PreK-12 teachers, three principals, and two directors, in District Z. District Z was located within a rural county, County Z, in southwestern Virginia. County Z was located between two urban areas, and the 2010 population census data reported 68,676 residents (United States Department of Commerce, 2010). According to the County Z's comprehensive plan (2007), agriculture remained the primary industry, but the area had broadened its economic base to include commercial growth in other industries. County Z's top employment sectors excluding agriculture, according to the fourth quarter 2006 results, were identified as Education/Health Care with 20.10%, Professional, Scientific, and Technical Services with 16.04%, and Construction with 14.18%; accounting for slightly over 50% of the County Z's workforce (Bedford County 2025 Comprehensive Plan, 2007).

The most recent District Z Virginia state report card reflected enrollment of 10,513 students for the 2012-2013 school year. District Z has three high schools, three middle schools, one alternative education center, one science and technology center, and 15 elementary schools. Based on federal ethnicity and race reporting, District Z enrollment included .5% American Indian or Alaska Native students, 1% Asian students, 7% Black or African American students,

.5% Native Hawaiian/ Other Pacific Islander students, 85% White students, 3% Hispanic or Latino students, and 3% Two or More Race Categories students².

Table 3.1

Federal Ethnicity and Race Report for District Z Student Enrollment Reported in Percentages

	American Indian or Alaskan Native	Asian	Black or African American	Native Hawaiian/Other Pacific Islander	White	Hispanic or Latino	Two or More Race Categories
District Z	.5	1	7	.5	85	3	3

During the 2013-2014 school year, District Z employed 753 teachers. All teachers in District Z will be asked to participate in the proposed study, excluding the teachers at one of the high schools. The teachers at one of the high schools were excluded from participation in the proposed study because the researcher was employed at that school in a supervisory capacity. This connection would likely influence survey instrument results of the teachers; therefore, this group of 73 teachers were excluded from participation in the proposed study. The exclusion of teachers at one of the high schools in District Z reduced the study’s overall population by 73 teachers. The total sample teacher population for District Z was 680 teachers, ranging in grade level from pre-kindergarten to 12th grade who were assigned to one of 22 schools, including 15 elementary, three middle, two high, one alternative, and one career and technology center.

The proposed study analyzed the survey instrument results of the teacher participants at the 22 schools as the District Z data set, after excluding the subset teacher survey instrument results from the three elementary schools. This large data set was analyzed to answer research question three. The survey instrument results of teachers from three of District Z’s elementary schools were used as the second data set to answer research questions four and five. The proposed study used student enrollment and district staffing as the primary indicators to select three elementary schools for additional study; however, federal and state reporting criteria was examined to select two similar elementary schools and one dissimilar school. After identifying the elementary schools with similar membership and staffing, the state and federal reporting criteria including the federal ethnicity and race reports, students with disabilities, students

² District Z Student Information System Enrollment Summary School/District Federal Ethnicity and Race Report (September 30, 2012).

receiving free/reduced lunch, student achievement results on 2013 and 2014 Standards of Learning (SOL) assessments were analyzed to ensure two similar elementary schools and one dissimilar elementary school were selected for additional study. The principals of the selected elementary schools, School A, School B, and School C, were interviewed to answer research questions five and six. The district personnel were interviewed to discover how the district prepared for implementation during the waiver year and answer research questions one and two.

Selection of Subset Teachers and Principals

A unique feature of the proposed study was the selection of three elementary schools for additional study. The teachers at the three elementary schools represented the subset teacher population. School A, School B, and School C were identified as three elementary schools with similar enrollment and district staffing levels. School A enrolled 436 students, School B enrolled 505 students, and School C enrolled 461 students. According to the District Z staffing standards, School A, School B, and School C employed a principal, one assistant principal, and one counselor. This staffing established an important similarity because of the additional support for the teachers, two administrators and one counselor, at School A, School B, and School C. During the 2013-2014 school year, School A employed 35 teachers, School B employed 37 teachers, and School C employed 32 teachers suggesting similar numbers of teachers for both principals and assistant principals to supervise and evaluate.

Table 3.2

Descriptions of Staffing of School A, School B, and School C

	Teachers	Principal	Assistant Principal	Counselor
School A	35	1	1	1
School B	37	1	1	1
School C	32	1	1	1

VDOE requires schools to report information annually as a regular part of the data collection process. The VDOE annually reported data was reviewed to determine if the three schools selected based on student enrollment and District Z staffing standards would be the three

elementary schools in District Z selected as the sites selected for additional study³. School A and School B were identified as two elementary schools with similar reporting characteristics on the VDOE annually reported categories. School C was identified as the third elementary school selected for additional study because of differences from School A and School B as identified on two of the VDOE annually reported categories. School A and School B were identified as the two most similar elementary schools based on two of the reporting criteria required by the VDOE in District Z, the percentage of students with disabilities and the percentage of economically disadvantaged students. While student enrollment at each of the three elementary schools was similar, School C differed from School A and School B in the other categories. School C was identified as the elementary school in District Z with similar enrollment and staffing, but differences from School A and School B based on the VDOE reporting criteria of students with disabilities and economically disadvantaged students. Table 4: Reporting Characteristics of School A, School B, and School C demonstrates the similarities in reporting criteria for two of the elementary schools and demonstrates the differences between School A and School B with School C.

Table 3.3

Federal Reporting Characteristics of School A, School B, and School C Student Enrollment, Students with Disabilities, and Economically Disadvantaged Students Reported in Percentages

	Enrollment	Students with Disabilities	Economically Disadvantaged Students
School A	436	14%	64%
School B	505	12%	53%
School C	461	6%	13%

After a review of this data, School A, School B, and School C were evaluated further to identify any additional similarities and differences. School A, School B, and School C represent three different school feeder zones for District Z; School A is located within the town limits of

³ Virginia Department of Education Educational Information Management 2012-2013 End-of-Year Student Record Collection – 10 District Z.

the county seat, School B is located in the southwestern part of the county, and School C is located in the northeastern part of the county. School A and School B are approximately 25 miles apart, School A and School C are approximately 17 miles apart, and School B and School C are approximately 43 miles apart. School A and School B differ in enrollment numbers by 69 students and disadvantaged population by 11%. School C student enrollment was greater than School A by 25, but School C student enrollment was less than School B by 44 students. However, these differences in student enrollment at the three elementary schools did not change the allocation of the additional staffing of an assistant principal and a counselor at School A, School B, and School C.

The Federal Ethnicity and Race Report indicated substantial differences in the ethnic and racial breakdown of the schools⁴. School A enrollment included 1% American Indian or Alaska Native students, 2% Asian students, 21% Black or African American students, 0% Native Hawaiian/ Other Pacific Islander students, 68% White students, 2% Hispanic or Latino students, and 6% Two or More Race Categories students. School B enrollment included 1% American Indian or Alaska Native students, 0% Asian students, 1% Black or African American students, 0% Native Hawaiian/ Other Pacific Islander students, 88% White students, 6% Hispanic or Latino students, and 4% Two or More Race Categories students. School C enrollment included 1% American Indian or Alaska Native students, 7% Asian students, 3% Black or African American students, 0% Native Hawaiian/ Other Pacific Islander students, 83% White students, 3% Hispanic or Latino students, and 3% Two or More Race Categories students. Table 3.4 displays the federal ethnic and race report for students enrolled at School A, School B, and School C.

⁴ District Z Student Information System Enrollment Summary School/District Federal Ethnicity and Race Report (September 30, 2012).

Table 3.4

Federal Ethnicity and Race Report for Student Enrollment at School A, School B, and School C Reported in Percentages

	American Indian or Alaskan Native	Asian	Black or African American	Native Hawaiian/Other Pacific Islander	White	Hispanic or Latino	Two or More Race Categories
School A	1	2	21	0	68	2	6
School B	1	0	1	0	88	6	4
School C	1	7	3	0	83	3	3

The VDOE School Report Cards for District Z schools were reviewed to examine student achievement at School A, School B, and School C to identify further similarities and differences. Because the proposed study will focus on teacher perceptions of the new performance-based teacher evaluation system, student achievement on state assessments at School A, School B, and School C could influence teacher perceptions at the schools. A review of School A, School B, and School C report cards was conducted to identify student achievement trends at the three schools. VDOE School Report Cards indicated similar performance at School A and School B on SOL tests in the 2013 and 2014 administration of the tests, except for 2014 Math achievement. The VDOE School Report Cards also indicated School C scored consistently higher than School A and School B during those same administrations of the tests, excluding the 2013 and 2014 History achievement⁵. Table 3.5 and Table 3.6 display the average 2013 and 2014 SOL performance of students at School A, School B, and School C.

⁵ Virginia Department of Education School Report Cards Retrieved from <https://p1pe.doe.virginia.gov/reportcard/> on March 18, 2014.

Table 3.5

2013 SOL Performance of Students at School A, School B, and School C

	English	Math	History	Science
School A	78	53	82	82
School B	81	49	88	89
School C	83	76	87	91

Table 3.6

2014 SOL Performance of Students at School A, School B, and School C

	English	Math	History	Science
School A	64	57	81	75
School B	66	75	88	84
School C	83	76	87	91

The achievement results of School A and School B did not demonstrate substantial performance differences regardless of the differences in economically disadvantaged students and the federal ethnic and racial reporting categories, except for a substantial increase in student math achievement for School B in 2014. The achievement results of School C demonstrated consistently higher student scores in all tested subjects, excluding a one percent average increase on the 2013 and 2014 History SOL for School B over School C. Despite the differences in economically disadvantaged students and the federal ethnic and racial reporting categories for School A and School B, the two schools were similar in student enrollment, staffing standards, percentage of students with disabilities, and student achievement as measured by the 2013 and 2014 SOL tests. While School C enrolled similar numbers of students qualifying School C for the same staffing standards as School A and School B, differences were noted between School C's federal ethnic and racial reporting categories, the students with disabilities and economically disadvantaged student categories when compared to School A and School B. School C's annual report card reflected higher student achievement on the annual SOL tests for the past two years,

excluding a one-point difference on both the 2013-2014 History SOL tests. When all data sources were reviewed, School A and School B were identified as the two most similar elementary schools, and School C was identified as the elementary school that was most different from School A and School B, excluding student enrollment numbers and district staffing standards.

District Z Study Population Including Subset Teacher and Principals

The teacher population for the proposed study included 680 teachers at 22 schools. After the researcher identified the subset schools for additional study, the District Z teacher population was reduced by 104 teachers. The District Z teacher data set included 576 possible participants; the School A teacher data set included 35 possible participants; the School B teacher data set included 37 possible participants; and the School C teacher data set included 32 possible participants. The selection of the elementary schools identified the three principal participants. The Directors of Human Resources and Instruction were identified as research participants because of an administrative role in the district during the waiver year. Table 3.7 provides details of the teacher population sample for the proposed study.

Table 3.7

Proposed Study Teacher Participants

Research Data Set	Teacher Survey Instrument Population
District Z	576
School A	35
School B	37
School C	32
Total	680

Data Collection

Data was collected from all teachers in District Z who elected to participate in the proposed study. The District Z data set was used for comparative analysis with School A, School B, and School C. The proposed study provided comparative analysis of data collected from

teachers at School A, School B, and School C who participated in the proposed project using an electronic survey instrument. The survey instrument that was developed by Professor Min Sun and R. Brock Mucheson was delivered using Qualtrics⁶, the survey instrument software program used by Virginia Polytechnic Institute and State University.

Approval was obtained to conduct research through the Virginia Polytechnic Institute and State University Institutional Research Board (IRB). After receiving IRB approval, approval was requested to conduct research through District Z's Supervisor of Assessment and Planning. Once approval was secured a school specific list of all teachers that included name, email address, personnel status title, assignment title, and assignment location was requested from the Human Resource Department. This school specific list of teachers was used to create the District Z Teacher Data Set, the School A, School B, and School C Teacher data set. The information provided by the District Z Human Resources Department was used to build email groups of teachers who were asked to participate in the proposed study by completing the survey instrument. Additionally, the District Z Teacher data set was used to customize the electronic teacher survey instrument. This customization feature allowed participants to select the specific school to which they were assigned during the 2013-2014 school year. This customization feature allowed participants to nominate site-specific colleagues, all teachers and administrators at each individual school site, as individuals who provided help or assistance with implementation of the new performance-based teacher evaluation system during the 2013-2014 school year. The researcher customized the survey instrument by importing school specific lists of teachers, provided by the Human Resources Department in Microsoft Excel, into the Qualtrics survey instrument.

The principals of School A, School B, and School C were interviewed for responses to research questions five and six. The school site principals provided information and documentation of the site specific activities conducted during the new teacher evaluation implementation year and principals' perceptions were compared to teachers' perceptions at the three elementary schools selected for further study. Data was collected from the Directors of Human Resources and Instruction. The District Z directors were interviewed for responses to

⁶ Qualtrics was the software delivery system used by Virginia Polytechnic Institute and was used to create and distribute the survey instrument developed by Professor Min Sun and R. Brock Mucheson.

research questions one and two. The district directors provided information and documentation of the district level activities conducted in the year prior to implementation and the implementation year, school years 2012-2013 and 2013-2014 respectively. As a waiver school district, it was anticipated the district engaged in preparatory activities to prepare for full implementation.

After securing approval from the Virginia Polytechnic Institute and State University IRB and District Z, principals at each of the participating school sites were notified of the project through email. This email contained information about the proposal, informed consent, and included timelines for surveying teachers and conducting interviews. Principals were asked to encourage teachers to participate in the proposed study. This introduction and informational email provided the site-based administrators with information to share with staff and informed site-based administrators where to direct questions. After the site-based administrative contact was made, all participating teachers were sent an introductory email that included the purpose of the proposed project, assurance of confidentiality, general information about the survey instrument, when to expect the survey instrument electronic link, and information about informed consent. Participating teachers were also told to whom questions should be directed.

Approximately one week after the teacher introductory email, the survey instrument was distributed electronically through the Qualtrics software program to teachers using the district's webmail database that was collected from the district. The survey instrument was administered during a predetermined timeline identified as the optimal window for teacher experiences with the new performance-based teacher evaluation system and participation in the study. An implementation window was developed to maximize responses from participants. Follow-up emails were sent in an effort to encourage participation.

The survey instrument was used previously in two school districts as part of the larger research study. During these administrations of the survey instrument, previous participants averaged 14 minutes to complete the survey instrument. It was estimated that teachers would be able to complete the electronic survey instrument in 30 minutes or less. It was expected the short time commitment for completion would increase the response rate. Research participants have the option to answer all, some, or none of the survey instrument questions. Despite the low time commitment required of participants, missing data was expected on responses to questions on the survey instrument. However, the researcher planned to identify missing data through the data

cleaning process, and included the responses of teachers electing to complete the questions. This strategy would result in differences when reporting the number of teachers responding to questions. This strategy would also eliminate the need to make future contact with research participants.

School A, School B, and School C principals will be contacted by telephone and requested to participate in further study. The criteria used for selection of the elementary schools, information about the proposal, informed consent, and expected time commitment required by the principal were shared with the selected principals. The principals were encouraged to participate in the study and informed that disaggregated data gathered from each school site would be provided to the principals. Principals could use the results and findings of the proposed study to set goals and plan professional development. After securing consent, the interviews were scheduled according to a mutually agreed upon time for the researcher and principal. Because the interviews took place after the principals had completed most of the requirements of the new teacher evaluation system, the responses from principals should accurately reflect perceptions based on the current implementation process.

The two District Z directors were notified by telephone of the request to seek additional information for the proposed study. The criteria used for selection, information about the proposal, informed consent, and time commitment required by the director were shared. The directors were encouraged to participate in the study and informed that disaggregated data gathered from District Z teacher data sets would be provided to the district. The district could use the results and findings of the proposed study to set goals and plan professional development. After securing consent, the interviews were scheduled according to a mutually agreed upon time for the researcher and the directors. The director interviews were not dependent on the teacher evaluation implementation schedule because the district directors would not be conducting teacher evaluations and could discuss new teacher evaluation preparation activities without regard to the actual evaluation timeline schedule.

Data Analysis

A mixed-method design, with both quantitative and qualitative information, was utilized to describe and understand the new performance-based teacher evaluation system as implemented in District Z and at three elementary schools in District Z. Descriptive statistical analysis was used to analyze teacher responses to the survey instrument. A qualitative case study

design was used to analyze the three principal interviews and two director interviews. The interviews identified how the district prepared for implementation and offered comparative analysis of principals' perceptions of teacher evaluation at each of the three schools during the implementation year. Principals were asked ten matched questions, from the teacher survey instrument, in an interview format. District Z directors were asked two open-ended questions to explain what was done at District Z to implement the new teacher evaluation system.

Content validity of the survey instrument was a priority for the developers, Professor Min Sun and R. Brock Mutcherson. After the survey instrument was piloted in another school district in Virginia, the developers met with a small group of respondents and received oral and written feedback regarding survey instrument language, format, length, and content and construct of the questions. The developers also met with a second group of administrators from another school district after the pilot to review the same items. Prior to distributing the survey instrument to District Z teachers, the survey instrument was tested by sending test survey instruments to a group of six professional educators. These test administrations of the survey instrument identified problems with customization of schools and teaching staff and the compilation of data. Each of the problems was corrected, and another test was distributed to the same group of six professional administrators to confirm that the problems identified were corrected.

The survey instrument questions were designed to address multiple characteristics of the new teacher evaluation system and were designed to use in the larger research project. Because the proposed study was a subset study of the larger research project, the proposed study will only use those survey question and responses that aligned with this study's research questions. The selected research questions were Likert-type responses designed to identify the extent or importance, agreement with or value, and degree of usefulness of the question and included a variety of options with responses. Extent of agreement with responses were: 1) Not at all; 2) Some extent; 3) Moderate extent; 4) Great extent; 5) Not applicable; and 6) Not sure. Value responses were: 1) Not at all important; 2) A little important; 3) Somewhat important; 4) Very important; 5) Not applicable; and 6) Not sure. Degree of usefulness with responses were: 1) Not at all useful; 2) Somewhat useful; 3) Useful; 4) Very useful; 5) Not sure. After careful analysis, the survey instrument questions were classified into eight primary categories for data analysis. The eight primary categories of questions were then aligned with the questions asked of District Z teachers, School A, School B, and School C principals, and District Z directors. Table 3.1:

Survey Instrument Question Categories Aligned with Research Data Sets identified the categories of the survey instrument and the subset of the proposed study population that would answer the research questions.

Table 3.8

Survey Instrument Question Categories Aligned with Research Data Sets

	District Z Teacher Data Set (including teachers at School A, School B, and School C)	School A, School B, and School C Teacher and Principal Data Subset	District Z Directors Data Set
Demographic	X		
Teaching Role/Assignment	X		
Time spent on duties/evaluation	X	X	
Observations	X	X	
Teacher evaluation	X	X	X
Professional Development	X	X	X
School Leadership	X	X	
Teaching Characteristics	X		

The demographic, teaching role/assignment, and teaching characteristics categories were included to explain identifying characteristics of the District Z teacher data set and School A, B, and C teacher data sets.

Descriptive statistics, including frequencies, means, and percentages, were applied to survey instrument responses of the District Z data set; excluding the three elementary schools to answer research questions three. The descriptive statistic information of School A, School B, and School C was compared with the larger District Z data set. Demographic data was analyzed and included in the descriptive analysis to provide a rich description of the teacher participants. The survey instrument responses were analyzed using the researcher’s computer software IBM SPSS Statistics 22 and exported to Microsoft Excel: mac 2011. The IBM SPSS Statistics 22 program disaggregated the descriptive statistics and reported the statistics by question in tables to identify any patterns and to compare and contrast teachers’ perceptions of the new evaluation system.

Exporting the data from IBM SPSS Statistics 22 to Microsoft Excel: mac 2011 enabled the data to be analyzed in a format compatible with Microsoft Word: mac 2011.

A qualitative case study approach was used to address research questions one, two, five, and six. The interviews of district directors addressed research questions one and two and were expected to validate the actions taken by the district to prepare for the implementation of the new teacher performance-based evaluation system. Rossman and Rallis (2012) suggested a purpose of case study research was to seek in-depth understanding of a phenomenon through a close examination of a particular case. The principals of School A, School B, and School C were asked a series of selected matched questions from the survey instrument during an interview to address questions five and six. Principal responses to interview questions were reported and compared with the teacher responses to the matched questions in an interview format. The interview dialogue was transcribed utilizing a phenomenological analysis strategy to seek meaning and address the research questions. The principals and directors were asked to check the transcribed interviews for accuracy and given the opportunity to make any changes, additions and/or deletions. The analysis strategy of meaning condensation (Rossman & Rallis, p. 292-293) was employed to construct a narrative description and identify themes for each of the research questions under consideration.

Summary

The purposes of the proposed study were to collect empirical evidence on the implementation of the new performance-based teacher evaluation system by comparing teachers' and principals' perceptions at three elementary schools to the perceptions of teachers in the school district, identified as District Z, and were to add to the research base on policy implementation theories. This chapter described the methodology, mixed-methods design, which was used in this study. This chapter described the selection of the population and the criteria used to select three elementary schools from all of the schools in District Z based on state and district criteria and local staffing standards. The data collection process and the methods used to analyze the data, including the teacher survey instrument and the principal and director interviews, were explained. The proposed study offers deep insight into one district's implementation of the new performance-based teacher evaluation system and adds to the existing body of research on policy implementation and teacher evaluation.

Chapter 4

Presentation and Analysis of Findings

Introduction

District Z implemented the new performance-based teacher evaluation system during the 2013-2014 school year, after spending the 2012-2013 school year on a VDOE approved implementation waiver. The compilation of data presented in this study examined teachers' perceptions of the new performance-based teacher evaluation system in District Z. Teachers' perceptions of the new evaluation system at three elementary schools were compared to teachers' perceptions in District Z. Additionally, principals of the targeted elementary schools were interviewed using matched questions from the teacher survey instrument to address research questions and principal responses were compared with teacher responses on the matched questions. District Z Directors of Human Resources and Instruction were interviewed to provide a description of what the district did during the waiver year (school year 2012-2013) to prepare for implementation. Findings from this study will add to the existing research on policy implementation and teacher evaluation. Further, the findings from this study may provide deep insight for education policy makers and leaders regarding perception and how perception can impact implementation of policy in a public school setting.

Procedures

IRB approval was given on April 17, 2014, (Appendix A) which allowed the researcher to begin making contact with staff members from the district and seek permission to conduct the study. District Z approval to conduct research was granted on April 18, 2014 (Appendix B).

Procedures with District Z Directors

District Z directors of Instruction and Human Resources were contacted by telephone and requested to participate in the study. The criteria used for selection, information about the proposal, informed consent, and time commitment were discussed. The directors provided verbal consent and interviews were scheduled. Prior to beginning the interview, director participation was requested and granted through the completion of the informed consent for the study. The directors were asked questions to inform the study regarding District Z's preparation for implementation of the new performance-based teacher evaluation during the 2012-2013 school

year, the waiver year. The interviews were audio recorded using the researchers personal laptop computer and the interviews were transcribed. Directors were given an opportunity to review the transcription to ensure accuracy of content. Information reported by directors during the interview was reported in the findings and analysis.

Procedures with District Z Teacher Data Set

Upon approval from District Z, information was collected to build the teacher population for the survey instrument. Additionally, District Z provided the VDOE annually reported data, which was evaluated to select three elementary schools to represent the subset teachers and principals at School A, School B, and School C. Principals at the 22 schools in District Z were notified of the research project through email. Principals were provided information about the project and notified where to direct questions. District Z teachers, excluding the teachers at one high school, were sent an introductory email explaining the purpose of the project, an assurance of confidentiality, general information about the survey instrument, when to expect the email with the survey instrument electronic link, and information about informed consent. The electronic survey instrument was distributed to teachers using Qualtrics, the survey instrument database used by Virginia Polytechnic Institute and State University, on April 28, 2014, through email. General information about participation, informed consent, and the purpose of the study were included in this email with the link to the survey instrument. The survey instrument link directed participating teachers, including subset teachers, to the informed consent, which required approval before participants could begin the survey instrument. Qualtrics, the electronic survey software program, was used to send emails to teachers who had not participated in the survey each week during the survey instrument data collection window. The survey instrument was closed for District Z teachers on May 30, 2014. Data was collected from District Z teachers for five weeks.

Procedures with Principals of Subset Schools A, B, and C

The principals of School A, School B, and School C were contacted by telephone and requested to participate in further study. The criteria used for selection, information about the study, informed consent, and expected time commitment from each principal was discussed. Interviews were scheduled with the three principals. Each of the principals consented to participate by completing the informed consent prior to beginning the interview. Principals were

asked selected matched questions from the teacher survey instrument in an interview. The interviews were audio recorded using the researcher's personal laptop computer. Interviews were transcribed and each principal was given an opportunity to review the transcript to ensure accuracy of content. Principal responses to the interview questions were analyzed to address research questions.

Findings and Analysis of District Z Director Interviews

The directors of human resources and instruction for District Z were interviewed on May 12, 2014. The interviews took place separately at the location chosen by the director. The findings reported below include information in response to the research questions and in response to follow-up and/or clarification questions.

The director of human resources for District Z will be referred to as Director A while reporting the findings. Director A has worked in human resources for 28 years, serving District Z as the director of human resources for three years and as the assistant director of human resources for four years. Before working with District Z, Director A worked in the business industry in the human resources field. Director A has taught high school and coached and has experience as an adjunct professor. The director of instruction for District Z will be referred to as Director B while reporting the findings. Director B has 21 years experience in education administration and teaching, including five years as the director of instruction, 10 years as a principal, and six years teaching middle school social studies.

The research questions addressed in the interviews were:

Questions for District Z Directors (Director A and Director B)

1. What actions did District Z take during the 2012-2013 school year, the waiver year, to prepare for implementation of the new performance-based teacher evaluation system as outlined in *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers*?
2. What professional development and/or training were provided for principals and teachers during the waiver year to prepare for implementation?

Responses were coded and central themes were identified for each research question as noted in Table 4.1 and 4.2.

District Z used the waiver year as intended to prepare for implementation of the new performance-based evaluation system. A 16-member committee, consisting of five district administrators, two school site administrators, five teachers, two parents, one professional association representative, and one school board member, was established to develop the district's evaluation tool. From that 16-member committee, two subcommittees were formed. One subcommittee focused on the new performance-based teacher evaluation and used the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* published by the VDOE as the framework for creating the district's evaluation manual. The smaller subcommittee explored the development and implementation of a teacher career ladder that would establish a differentiated pay scale for teachers. Both directors stressed the committee's focus on Standard 7: Student Academic Progress, specifically noting the committee's strong desire to identify what student academic progress would look like and how to measure that progress. The full committee, District Z administrators, and principals reviewed the document prior to school board adoption. The District Z School Board approved the draft document in May 2013, approximately 10 months after the committee was formed. Directors A and B responses to research question one were recorded as central themes in Table 4.1

Table 4.1

Research Question 1: Preparation Actions of District Z During 2012-2013

	<p style="text-align: center;">Director A and B Central themes in response to research question 1</p>
<p>Preparation Actions of District Z during waiver year, 2012-2013</p>	<ul style="list-style-type: none"> ▪ District Z developed a committee and initial work began in the summer of 2012. ▪ Committee included two subcommittees – teacher evaluation and teacher career ladder pilot. ▪ Committee used VDOE <i>Guidelines</i> document as a resource and reference, adopting all standards and weighting. ▪ Teacher evaluation subcommittee focused on Standard 7: Student Academic Progress (What does this look like? How do we measure it?) ▪ Teacher evaluation subcommittee focused on role of observations and other performance documentation methods and data sources ▪ Committee was comprised of central office leadership, principals, teachers, school board member, parent ▪ Similarities were noted with new evaluation system and District Z’s previous evaluation system (both evaluation systems were developed from by the same researcher). ▪ District Z teacher evaluation handbook received full committee review and full review by district administration and principals, prior to adoption by the District Z School Board.

Research question two was asked to identify the professional development and/or training provided for principals and teachers during the waiver year to prepare for implementation. In response, Directors A and B reported the evaluation process was reviewed by District Z administration at a district administrative meeting. This review focused on the differences between the new performance-based teacher evaluation system in comparison to the district’s previous teacher evaluation system. Directors A and B did not indicate any teacher professional development occurring during the waiver year. Further, Directors A and B did not indicate any

principal professional development beyond review at an annual conference and support by District Z leaders on an as needed basis. Directors A and B responses to research question two were recorded as central themes in Table 4.2.

Table 4.2

Research Question 2: Professional Development/Training Provided for Principals and Teachers during 2012-2013

Director A and B Central themes in response to research question 2	
Professional Development/Training provided for principals and teachers during the waiver year, 2012-2013 ⁷	<ul style="list-style-type: none"> ▪ District leadership provided was available to answer questions and assist principals with implementation. ▪ Evaluation process was reviewed by District Z administration at annual administrative summer conference, which included District Z principals, assistant principal, and administrative office staff. Review included information about the new evaluation process and forms, notation of changes in the new evaluation tool, and a review of the new evaluation manual.

Findings and Analysis of District Z and Schools A, B, and C Teacher and Principal Data Set

The District Z teacher data set included teacher participants at all schools in District Z during the 2013-2014 school year, excluding the teachers at one high school. An email was sent to teachers on April 28, 2014, that included the link to Qualtrics, the survey instrument software program. Teachers were able to participate in the research study at any time during the study window, which ended on May 30, 2014. School A, School B, and School C principals were contacted by phone and requested to participate in the study by answering matched survey instrument questions in an interview. The findings and analysis reported include descriptive statistics to the questions that were matched and asked of principals of School A, School B, and School C during the interviews. The District Z teacher data set findings were reported in four groups, District Z teachers, School A teachers, School B teachers, and School C teachers. The subset teacher data set (Schools A, B, and C) were excluded from the District Z teacher data set

⁷ District Z professional development was not defined as part of the research study. Director participants were allowed to respond to the research question based on their own, unspecified, definition of professional development.

to ensure accurate comparison among the four groups. Combining Schools A, B, and C with the District Z teacher data set would have skewed the findings of the District Z teacher data set. Table 4.3 includes the sample teacher population for each data set and the participants' response rate for each data set.

Table 4.3

District Z Teacher Data Sets and Survey Instrument Participation Rates

Research Data Set	Teacher Survey Instrument Population	Teacher Survey Instrument Response Rate by Research Data Set	Percentage of Teacher Survey Instrument Response Rate by Research Data Set
District Z	576	275	48%
School A	35	15	43%
School B	37	37	100%
School C	32	30	94%
Total	680	357	53%

Missing data was expected on responses to questions on the survey instrument. The researcher identified missing data through the data cleaning process, and included the responses of all teachers electing to complete the survey instrument questions regardless of how many questions an individual teacher may have completed. This strategy resulted in differences when reporting the number of teachers responding to questions, which were noted in the table for each survey instrument matched question. This strategy was selected because it eliminated the need to make future contact with research participants.

All District Z teacher data sets responses to survey instrument questions were reported in tables 4.4 through 4.16. Principals of School A, B, and C were asked matched questions from the teacher survey instrument during individual interviews and those responses were reported with the matched teacher responses.

The research questions addressed using all District Z teacher data sets and the principals of School A, School B, and School C were:

Questions for District Z Teacher Data Set (including teachers at School A, School B, and School C)

3. What are teachers' perceptions of the new evaluation policy?
 - a. Did access to support, including professional development, increased knowledge and/or materials, and assistance from other educators affect teacher perceptions of the evaluation system?
 - b. If so, in what ways did access to support affect teacher perceptions?

Questions for Subset Teachers and Principals at School A, School B, and School C

4. How do teachers' perceptions at School A, School B, and School C compare with District Z teachers' perceptions?
5. What are the principals' perceptions of the new teacher evaluation system?
6. How do principals' perceptions compare with teacher perceptions?

The survey instrument included 28 questions targeting teacher perceptions of the new performance-based teacher evaluation; many of the questions had multiple parts. However, the proposed study did not report results on all of the questions. Analysis of the matched questions and descriptive statistics were provided in Tables 4.4 through 4.16.

The survey instrument included questions that described the research population characteristics. Principals of School A, B, and C were asked matched questions to provide descriptive characteristics during interviews. Tables 4.4 provided statistical information describing the teacher and principal population by data set.

Tables 4.4

Description of Teacher and Principal Data Sets

Data Set	N	Gender		Percentage of N teaching an SOL content subject		Percentage of N's educational attainment	
		Percentage of N		Yes	No	Bachelors	Masters+
		Female	Male				
District Z	275	83.7	16.3	58	42	53.2	46.8
School A	15	73.3	26.7	46.7	53.3	80	20
School B	37	77.8	22.2	48.6	51.4	56.8	43.2
School C	30	75.9	24.1	62.1	37.9	46.4	53.6

Data Set	N	Number of years working at this school				Number of years working as a full-time teacher in K-12 schools			
		Minimum	Maximum	Mean	S.D.	Minimum	Maximum	Mean	S.D.
District Z	275	0	40	9.12	7.4	0	40	13.8	9.3
School A	15	1	30	11.33	8.6	2	33	18	10.4
School B	37	1	40	10.89	10.7	1	40	17.3	11.4
School C	30	1	30	10.90	11.0	1	41	17.5	12.6

Principal Data Set	Gender N		Principal Education Attainment		Number of years in education		
	Female	Male	Masters	Specialist +	Teacher	Assistant Principal	Principal
School A	X		X		26	2	5
School B		X	X		13	5	22
School C	X		X		9	5	14

In Table 4.5, the District Z teacher data sets reported the average hours worked each week and the average hours worked annually on teacher evaluation activities. It was important to note that all District Z teacher data sets reported working on average far more hours than the standard workweek of 40 hours. School A teachers reported on average more hours working each week than teachers at School B, School C, and District Z. School A teachers worked at the elementary school with the highest number of disadvantaged students, students with disabilities, and the most diverse student population. Alternatively, School B teachers reported the highest average of hours spent working on teacher evaluation activities. However, the difference reported by all District Z teacher data sets on the average hours worked in teacher evaluation during the 2013-2014 ranged 3.72 hours, suggesting that all District Z teacher data sets spent a similar amount of time during the implementation year on teacher evaluation activities.

Table 4.5

Average Hours Worked in a Week and Average Total Hours Worked on Teacher Evaluation During the 2013-2014 School Year

Data Set	Average Hours Worked in a Week		Average Hours Worked on Teacher Evaluation during the 2013-2014 School Year	
	N	Hours	N	Hours
District Z Teachers	271	55.56	267	24.30
School A Teachers	15	62.27	15	24.87
School B Teachers	37	54.89	36	26.75
School C Teachers	30	53.63	29	23.03

Table 4.6 reported the average number of hours principals spent on teacher evaluation activities and the average number of hours principals participated in teacher evaluation professional development activities. Principals reported substantially more time on teacher evaluation activities during the 2013-2014 school year as compared to the District Z teacher data sets, which would be expected given that teacher evaluation remains a primary responsibility of principals. School C principal reported almost twice as many hours working on teacher evaluation activities when compared to the School A and School B principals. During the interview, School C principal reported doing most teacher evaluation work at home during the evenings and on weekends, which could have accounted for the substantial difference in time reported by the three principals.

Principal interviews affirmed the findings of the District Z Directors that professional development opportunities on the implementation of the new performance-based teacher evaluation system were low. School B and C principals indicated participating six hours and three hours of professional development respectively. School A principal indicated participating in 25 hours of professional development; however, it should be noted School A principal reported attending a department of education professional development. Further, all three

principals indicated the need for more professional development on the new teacher evaluation system during interviews.

Table 4.6

Principal Time Reported on Teacher Evaluation Activities and Professional Development

Data Set	Time Spent on Teacher Evaluation Activities School Year 2013-2014	Number of Hours Principal Participated in Teacher Evaluation Professional Development School Year 2013-2014
Principal School A	288	25
Principal School B	252	6
Principal School C	540	3

Table 4.7 reported information about who conducted observations and how frequent those observations were during the 2013-2014 school year for each teacher data set. All District Z teacher data sets reported peers and colleagues conducted the most frequent observations once or twice per year. Additionally, all District Z teacher data sets reported principals and assistant principals conducted observations once or twice per year or once or twice per semester as the next most frequent observers of instruction during the 2013-2014 school year.

Table 4.7

Who Conducted Observations and the Frequency Reported by each Teacher Data Set

District Z	Frequency of Observations				
	Once or twice per year	Once or twice per semester	Monthly	Weekly	Daily
Principal	110	75	27	6	2
Assistant Principal	89	52	13	3	0

(table continued)

Table 4.7 (cont.)

District Z	Frequency of Observations				
	Once or twice per year	Once or twice per semester	Monthly	Weekly	Daily
Department Chair	31	15	11	3	3
Master/Mentor or Instructional Coach	32	23	5	4	1
Peers/Colleagues	129	27	5	6	9
Other	19	4	4	5	0

School A	Frequency of Observations				
	Once or twice per year	Once or twice per semester	Monthly	Weekly	Daily
Principal	5	6	2	1	1
Assistant Principal	2	4	2	3	1
Department Chair	0	2	0	1	1
Master/Mentor or Instructional Coach	4	2	1	0	0
Peers/Colleagues	6	2	1	0	3
Other	0	0	0	0	0

(table continued)

Table 4.7 (cont.)

School B		Frequency of Observations				
	Once or twice per year	Once or twice per semester	Monthly	Weekly	Daily	
Principal	9	25	3	0	0	
Assistant Principal	9	20	4	2	0	
Department Chair	5	1	0	0	0	
Master/Mentor or Instructional Coach	4	0	0	0	0	
Peers/Colleagues	19	1	0	1	4	
Other	1	0	0	0	0	

School C		Frequency of Observations				
	Once or twice per year	Once or twice per semester	Monthly	Weekly	Daily	
Principal	13	15	0	1	0	
Assistant Principal	14	11	1	1	0	
Department Chair	0	0	0	0	0	
Master/Mentor or Instructional Coach	1	0	0	0	0	
Peers/Colleagues	16	4	0	0	0	
Other	4	1	0	0	0	

A unique feature of this study was a participants' ability to nominate colleagues at the school site and within the district as individuals who provided support regarding implementation of the new teacher evaluation system. The percentages of nominations by each teacher data set were reported in Table 4.8. Seventy-seven percent of District Z teachers nominated at least one colleague as an individual at his or her school site who provided support regarding

implementation of the new evaluation system. Further, 13% of District Z teachers nominated at least one other colleague, a representative who was not assigned to his or her school site or was not listed on the roster, as an individual who provided support regarding implementation of the new evaluation system. Sixty percent of School A teachers nominated at least one colleague within his or her school who provided support with teacher evaluation, and 13% nominated at least one other colleague, a representative who was not assigned to his or her school site or was not listed on the roster, as an individual who provided support regarding implementation of the new evaluation system. The percentage of School A teachers who nominated at least one colleague at his or her school as providing support with teacher evaluation was 17 percent less than District Z teachers; however, the same percentage of School A teachers nominated at least one other colleague as the percentage of District Z teachers. Seventy percent of School B teachers nominated at least one colleague at the school site who provided support with teacher evaluation, and 16% of School B teachers nominated at least one other colleague indicating similar perception with District Z teachers on the availability of support from colleagues both within the school and outside the school. Seventy-seven percent of School C teachers nominated at least one colleague who provided support with teacher evaluation at the school site, and 10% nominated at least one other colleague who provided support with teacher evaluation, indicating similar perceptions of collegial support at the school site and slightly lower perceptions of support from other colleagues.

Table 4.8

Nominations of Colleagues Providing Support with Teacher Evaluation in the 2013-2014 School Year

Data Set	Nomination of Colleagues														
	Colleague 1		Colleague 2		Colleague 3		Colleague 4		Colleague 5		Other Colleague ⁸ 1		Other Colleague ²		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
District Z	266	77	205	49	131	29	77	16	44	28	10	35	13	7	2
School A	15	60	9	33	5	13	2	6	1	0	-	2	13	1	6
School B	37	70	26	38	14	27	10	8	3	1	2	6	16	0	-
School C	30	73	22	67	20	37	11	17	5	3	10	3	10	1	3

The findings of the mean number of hours and the usefulness of professional development on three topics, content area, teacher evaluation, and using data to adjust instruction for District Z teacher data sets were reported in Table 4.9. District Z teachers reported an average of 3.80 hours of professional development on the teacher evaluation and six percent rated the professional development as very useful. District Z teachers reported an average of 8.36 of professional development on using data to adjust instruction and 20.4 percent rated the professional development as very useful. District Z teachers participated in more hours of professional development focused on using data to adjust instruction and rated that topic much higher in usefulness. While School A teachers reported less time on teacher evaluation and more time on using data to adjust instruction professional development hours than District Z teachers, the professional development focused on using data to adjust instruction was perceived very useful by more School A teachers. School B teachers reported an average of 1.87 hours of professional development on teacher evaluation, and none of the teachers perceived the professional development on teacher evaluation was very useful. School B teachers reported an average of 7.23 hours of professional development on using data to adjust instruction, and 21.2% perceived the professional development on using data to adjust instruction was very useful. School B teachers reported less professional development on both topics compared with School A and District Z teachers. Further, the percentage of teachers who perceived the professional

⁸ Other Colleagues were identified as individuals who provided support but were not included in the school site roster and/or were individuals who worked at another school/district location.

development on both topics was very useful was less than District Z teachers and School A teachers, except the teacher evaluation professional development was rated very useful by none of School B teachers which was the same as School A teacher perceptions. School C teachers reported an average of 1.87 hours of professional development on teacher evaluation, and 8.3% rated that professional development as very useful. School C teachers reported an average of 3.86 hours of professional development on using data to adjust instruction, and 12% rated that professional as very useful. The number of hours reported at School C was lower than District Z, School A, and School B, excluding the number of hours of professional development on teacher evaluation at School B. The effectiveness of the teacher evaluation professional development was perceived as very useful by a higher percentage of School C teachers than any other teacher data set. School C reported the lowest average of hours of professional development on using data to adjust instruction, and the percentage of teachers who rated that professional development very useful was lower than any other teacher data set. The number of hours of professional development and the effectiveness reported by teachers across all data sets affirmed the need for effective professional development as expressed by the principals.

Table 4.9

Professional Development Mean Hours Reported by Topic and Extent of Usefulness Reported by Percentage

Professional development topic	N	Mean Number of Hours Reported	District Z Teachers					
			N	Not at all Useful	Somewhat Useful	Useful	Very Useful	Not sure
Content area	225	13	223	9.4	21.5	35	27.8	6.3
Teacher evaluation	219	3.80	200	14.5	36	32	6	11.5
Using data to adjust instruction	222	8.36	230	11.3	30.9	30.4	20.4	7

(table continued)

Table 4.9 (cont.)

Professional development topic	N	Mean Number of Hours Reported	School A Teachers					
			N	Not at all Useful	Somewhat Useful	Useful	Very Useful	Not sure
Content area	11	10.82	12	8.3	50	25	16.7	0
Teacher evaluation	12	2.33	11	18.2	63.6	18.2	0	0
Using data to adjust instruction	12	16.50	12	8.3	25	41.7	25	0

Professional development topic	N	Mean Number of Hours Reported	School B Teachers					
			N	Not at all Useful	Somewhat Useful	Useful	Very Useful	Not sure
Content area	31	17.61	32	3.1	25	37.5	28.1	6.3
Teacher evaluation	30	1.87	28	10.7	57.1	17.9	0	14.3
Using data to adjust instruction	31	7.23	33	3	33.3	36.4	21.2	6.1

Professional development topic	N	Mean Number of Hours Reported	School C Teachers					
			N	Not at all Useful	Somewhat Useful	Useful	Very Useful	Not sure
Content area	23	14.43	26	11.5	19.2	30.8	38.5	0
Teacher evaluation	21	1.95	24	25	33.3	29.2	8.3	4.2
Using data to adjust instruction	22	3.86	25	12	40	36	12	0

Table 4.10 reported the extent of agreement among teacher participants with outcome statements regarding the new evaluation system, which directly addressed research questions 3 and 4. Twenty eight percent of District Z teachers agreed to a great extent the rating received on the final evaluation would accurately portray teaching; 28% of District Z teachers agreed to a great extent the procedures used to evaluate were fair; 23.9% of District Z teachers agreed to a great extent there was sufficient time to complete the evaluation; and 20.7 agreed to a great extent there were sufficient resources to complete the evaluation. When considering District Z

responses to these four perception questions, the responses indicated a quarter of the teachers agreed to a great extent with the outcome statements. The responses to these four questions represented the highest percentages of extent of agreement among the 12 outcome statements. Further, the Not at all extent category was ranked the lowest for each of the same four questions, 5.7%, 5.7%, 5%, and 5.9% respectively.

Despite a high percentage of nomination of colleagues who provided support with teacher evaluation during the school year, the percentage of District Z teachers with a great extent of agreement with outcome perception statements remained less than one-third of those responding. It was noted that the percentage of District Z teachers who agreed to a great extent with the outcome perception statements regarding the access to resources and time was less a one-quarter of those responding. A high percentage of District Z teachers nominated a colleague within his or her school indicating access to supportive colleagues was occurring but at the same time reported low hours of professional development on teacher evaluation and using data to adjust instruction. Access to resources, particularly support from other educators, was not enough to increase the percentage of teachers who agreed to a great extent with the perception outcome statements. But at the same time, the percentage of District Z teachers who did not agree with the perception outcome statements at all was rated the lowest among the 12 statements of outcomes.

Twenty percent of School A teachers agreed to a great extent the rating received on the final evaluation would accurately portray teaching, 6.7% agreed to a great extent the procedures used to evaluate were fair, 7.1% agreed to a great extent there was sufficient time to complete the evaluation, and 6.7% agreed to a great extent there were sufficient resources to complete the evaluation. The responses to these four perception outcome statements were not rated highest in the great extent of agreement category. Seven percent of School A teachers perceived the rating received would not portray teaching accurately at all, and 14.3% perceived they did not have sufficient time to complete the evaluation.

School A teachers reported participating in an average of 2.33 hours of professional development on teacher evaluation and none of the teachers rated those hours as very useful. School A teachers reported participating in an average of 16.5 hours of professional development focused on using data to adjust instruction and 25 percent rated those hours as very useful. School A teachers reported consistently lower percentages of a great extent of agreement with

the perception outcome statements and reported a perception of far less time and resources, when compared with District Z teachers.

Forty-two percent of School B teachers perceived to a great extent the rating received on the final evaluation would accurately portray teaching, 38.9% perceived to a great extent the procedures used to evaluate were fair, 16.7% perceived to a great extent there was sufficient time to complete the evaluation, and 11.1% perceived to a great extent there were sufficient resources to complete the evaluation. The perceptions of School B teachers were substantially higher in great extent agreement than District Z and School A teachers on perceptions of accurately portraying teaching and fairness of evaluation procedures. While School B teachers reported having a sufficient amount of time and resources at a higher percentage of a great extent of agreement as compared to School A teachers, School B teachers who perceived to a great extent of agreement there was a sufficient amount of time and resources was lower than District Z. Seventy percent of School B teachers nominated at least one colleague at the school site who provided support with teacher evaluation, and 16% of School B teachers nominated at least one other colleague indicating similar perception with District Z teachers on the availability of support from colleagues both within the school and outside the school.

Twenty-five percent of School C teachers agreed to a great extent the rating received on final evaluation would accurately portray teaching, 17.2% agreed to a great extent the procedures used to evaluate were fair, 13.8% agreed to a great extent there was sufficient time to complete the evaluation, and 17.2% agreed to a great extent there were sufficient resources to complete the evaluation. These perceptions reflected lower percentages of extent of agreement than District Z and School B teachers, but higher percentages of extent of agreement than School A teachers on the perception outcome statements. Seventy-seven percent of School C teachers nominated at least one colleague who provided support with teacher evaluation at the school site, and 10% nominated at least one other colleague who provided support with teacher evaluation, indicating similar perceptions of collegial support at the school site and slightly lower perceptions of support from other colleagues.

Table 4.10 also demonstrated an area of growth for all of the District Z teacher data sets. Notably, all teacher data sets responded with high percentages in the not at all extent of agreement for using feedback from observations to identify professional development programs for the next school year. Thirty two percent of District Z teachers, 40 percent of School A

teacher, 27 percent of School B teachers, and 41 percent of School C teachers perceived no agreement with the statement that feedback from observations was used to identify professional development programs for the next school year. This data suggested District Z administration should focus on providing appropriate feedback that could be linked to specific professional growth opportunities.

Principals of School A and B agreed to a moderate extent the rating teachers received on their final evaluation would accurately portray their teaching. While the Principal of School C agreed to a great extent the rating teachers received on their final evaluation would accurately portray their teaching. When comparing the principals' responses to the teachers' responses, simple agreement and/or disagreement was not applicable because teacher responses included multiple teachers. School A and B principals perceived a moderate extent of agreement and the greatest number of School A teachers, 46.7% agreed to a moderate extent their evaluation would accurately portray their teaching and the second highest percentage of School B teachers, 36.1%, agreed to a moderate extent their evaluation would accurately portray their teaching. School C principal perceived to a great extent the rating received on the final evaluation would accurately portray their teaching, but the largest percentage of School C teachers perceived at some extent, 39.3%, and a moderate extent, 28.6%. This analysis would suggest that a higher percentage of teachers did not agree with the principal's perception at School C.

Table 4.10

Extent of Agreement with Evaluation Outcome Statements Reported in Percentages

Extent of agreement with evaluation outcome statements	District Z Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Teachers were encouraged to find more effective ways to teach	257	6.2	27.2	37.7	28.8
I think the rating I receive on my final evaluation will accurately portray my teaching	261	5.7	25.3	41	28
Instruments used to evaluate me were precise	258	7.4	34.1	39.9	18.6

(table continued)

Table 4.10 (cont.)

Extent of agreement with evaluation outcome statements	District Z Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Procedures used to evaluate me were fair	261	5.7	25.3	41	28
The evaluation process was worth the effort for me	257	12.5	32.3	37.4	17.9
The evaluation process caused more administrative burden within the school	257	12.5	32.3	37.4	17.9
I received recognition based upon how my performance compared with my goals	258	32.9	33.7	23.6	9.7
Feedback included suggestion for improving my teaching	259	23.6	29.7	30.5	16.2
I used feedback from observations to identify professional development programs for next school year	259	32.8	30.5	25.1	11.6
Tying teacher evaluation to personnel decisions would improve our school	256	23.4	39.5	25.8	11.3
I had sufficient time to complete the evaluation	259	5	29.7	41.3	23.9
I had sufficient resources to complete the evaluation	256	5.9	32.4	41	20.7

(table continued)

Table 4.10 (cont.)

Extent of agreement with evaluation outcome statements	School A Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Teachers were encouraged to find more effective ways to teach	15	6.7	33.3	46.7	13.3
I think the rating I receive on my final evaluation will accurately portray my teaching	15	6.7	26.7	46.7	20
Instruments used to evaluate me were precise	15	6.7	46.7	40	6.7
Procedures used to evaluate me were fair	15	0	40	53.3	6.7
The evaluation process was worth the effort for me	15	20	26.7	46.7	6.7
The evaluation process caused more administrative burden within the school	15	7.1	28.6	42.9	21.4
I received recognition based upon how my performance compared with my goals	14	42.9	14.3	35.7	7.1
Feedback included suggestion for improving my teaching	15	40	13.3	46.7	0
I used feedback from observations to identify professional development programs for next school year	15	40	13.3	33.3	13.3
Tying teacher evaluation to personnel decisions would improve our school	14	14.3	42.9	35.7	7.1
I had sufficient time to complete the evaluation	14	14.3	42.9	35.7	7.1
I had sufficient resources to complete the evaluation	15	0	26.7	66.7	6.7

(table continued)

Table 4.10 (cont.)

Extent of agreement with evaluation outcome statements	N	School B Teachers			
		Not at all	Some extent	Moderate extent	Great extent
Teachers were encouraged to find more effective ways to teach	36	0	25	30.6	44.4
I think the rating I receive on my final evaluation will accurately portray my teaching	36	2.8	19.4	36.1	41.7
Instruments used to evaluate me were precise	36	2.8	27.8	44.4	25
Procedures used to evaluate me were fair	36	2.8	16.7	41.7	38.9
The evaluation process was worth the effort for me	36	8.3	30.6	33.3	27.8
The evaluation process caused more administrative burden within the school	35	20	31.4	34.3	14.3
I received recognition based upon how my performance compared with my goals	35	25.7	28.6	37.1	8.6
Feedback included suggestion for improving my teaching	36	8.3	44.4	33.3	13.9
I used feedback from observations to identify professional development programs for next school year	36	27.8	25	30.6	16.7
Tying teacher evaluation to personnel decisions would improve our school	35	17.1	40	28.6	14.3
I had sufficient time to complete the evaluation	36	2.8	33.3	47.2	16.7
I had sufficient resources to complete the evaluation	36	5.6	36.1	47.2	11.1

(table continued)

Table 4.10 (cont.)

Extent of agreement with evaluation outcome statements	School C Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Teachers were encouraged to find more effective ways to teach	29	3.4	34.5	34.5	27.6
I think the rating I receive on my final evaluation will accurately portray my teaching	28	7.1	39.3	28.6	25
Instruments used to evaluate me were precise	29	6.9	41.4	41.4	10.3
Procedures used to evaluate me were fair	29	6.9	34.5	41.4	17.2
The evaluation process was worth the effort for me	29	6.9	37.9	48.3	6.9
The evaluation process caused more administrative burden within the school	29	10.3	55.2	24.1	10.3
I received recognition based upon how my performance compared with my goals	29	44.8	24.1	27.6	3.4
Feedback included suggestion for improving my teaching	29	17.2	41.4	27.6	13.8
I used feedback from observations to identify professional development programs for next school year	29	41.4	20.7	24.1	13.8
Tying teacher evaluation to personnel decisions would improve our school	28	32.1	42.9	17.9	7.1
I had sufficient time to complete the evaluation	29	3.4	20.7	62.1	13.8
I had sufficient resources to complete the evaluation	29	3.4	24.1	55.2	17.2
I think the rating my teachers received will accurately portray their teaching performance				A B	C

To further understand teacher perceptions of the new performance-based evaluation, survey questions designed to understand the teachers' perceptions of the purpose, focus, what was emphasized, extent of involvement in the design and modification, extent of confidence in the components of the evaluation, and consistency with personal views were analyzed and

compared. The importance of purposes of teacher evaluation was reported in table 4.11. Fifty-six percent of District Z teachers perceived informing improvement in classroom instruction as a very important purpose of teacher evaluation, making this purpose the highest rated as very important. District Z teachers perceived the purpose to hold teachers accountable for student performance as the next very important purpose, rating this purpose at 51.3%. Sixty-nine percent of School A teachers perceived informing improvement in classroom instruction as a very important purpose and 57.1% perceived providing information for use in making decisions about reassignment and contract renewal as a very important purpose. Fifty percent of School B teachers perceived holding teachers accountable for student performance as a very important purpose of teacher evaluation, and 40.5% perceived informing improvement in classroom instruction as a very important purpose of teacher evaluation. Fifty percent of School C teachers perceived holding teacher accountable for student performance as a very important purpose, and 48.3% perceived informing improvement in classroom instruction as a very important purpose of teacher evaluation. Participants of this study perceived the purposes of teacher evaluation with great consistency. The only teacher data set rating another purpose very important was School A. It is important to note that 35.7% of School A teachers perceived holding teachers accountable for student performance as a very important purpose, making this the next very important purpose.

Principal agreement on the extent of the importance of purposes of teacher evaluation was varied. All three principals agreed holding teachers accountable for student performance was a very important purpose of teacher evaluation, representing the only purpose with principal consensus on the extent of importance. School A Principal perceived informing improvement in classroom instruction as a very important purpose of evaluation, and School A Teachers also perceived this purpose as very important with 69.2% responding. School C Principal perceived holding teachers accountable for student performance as a very important purpose of teacher evaluation, and the School C Teachers also perceived this purpose as very important with 50% responding. The School C Principal perceived informing improvement in classroom instruction and holding teachers accountable for student performance as two very important purposes of teacher evaluation, and most notably, School C Teachers perceived the same two purposes very important responding at 48.3% and 50% respectively. While discerning any pattern between principal perceptions was challenging, it was clear each of the principals identified a purpose of

teacher evaluation as very important and the teachers at the school rated one or both of the same purposes as very important with the highest percentage rate.

Table 4.11

Importance of Purposes of Teacher Evaluation Reported in Percentages

Purpose	N	District Z Teachers					
		Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To provide information for use in making decisions about reassignment and contract renewal	270	9.6	14.4	23.7	41.9	1.5	8.9
To provide evidence for use in determining compensation increases for teachers	271	37.6	10.7	12.5	14.0	12.2	12.9
To inform the selection of teachers' professional development	270	17.4	17.0	33.0	21.1	2.6	8.9
To inform improvement in classroom instruction	271	4.1	9.6	25.1	56.1	1.1	4.1
To hold teachers accountable for student performance	271	4.8	10.3	28	51.3	2.2	3.3
Other purpose	113	4.4	3.5	6.2	10.6	23	52.2

Purpose	N	School A Teachers					
		Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To provide information for use in making decisions about reassignment and contract renewal	14	14.3	0	28.6	57.1	0	0
To provide evidence for use in determining compensation increases for teachers	14	28.6	14.3	7.1	28.6	7.1	14.3
To inform the selection of teachers' professional development	14	14.3	7.1	71.4	7.1	0	0

(table continued)

Table 4.11 (cont.)

Purpose	School A Teachers						
	N	Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To inform improvement in classroom instruction	13	7.7	0	23.1	69.2	0	0
To hold teachers accountable for student performance	14	21.4	0	42.9	35.7	0	0
Other purpose	11	0	0	25	0	50	25

Purpose	School B Teachers						
	N	Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To provide information for use in making decisions about reassignment and contract renewal	36	11.1	13.9	27.8	33.3	2.8	11.1
To provide evidence for use in determining compensation increases for teachers	36	33.3	11.1	22.2	22.2	0	11.1
To inform the selection of teachers' professional development	35	14.3	14.3	45.7	14.3	0	11.4
To inform improvement in classroom instruction	37	5.4	10.8	37.8	40.5	0	5.4
To hold teachers accountable for student performance	34	8.8	8.8	32.4	50	0	0
Other purpose	9	0	0	11.1	0	44.4	44.4

Purpose	School C Teachers						
	N	Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To provide information for use in making decisions about reassignment and contract renewal	30	16.7	13.3	23.3	40	3.3	3.3
To provide evidence for use in determining compensation increases for teachers	30	26.7	16.7	13.3	13.3	23.3	6.7

(table continued)

Table 4.11 (cont.)

Purpose	School C Teachers						
	N	Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To inform the selection of teachers' professional development	29	10.3	13.8	51.7	10.3	6.9	6.9
To inform improvement in classroom instruction	29	10.3	13.8	24.1	48.3	0	3.4
To hold teachers accountable for student performance	30	3.3	10	36.7	50	0	0
Other purpose	10	10	10	0	40	10	30

Purpose	Principals					
	Not at all important	A little important	Somewhat important	Very important	Not applicable	Not sure
To provide information for use in making decisions about reassignment and contract renewal	A		B	C		
To provide evidence for use in determining compensation increases for teachers	A B	C				
To inform the selection of teachers' professional development		C		A B		
To inform improvement in classroom instruction			B	A C		
To hold teachers accountable for student performance				A B C		

Teacher and principal perceptions of the extent of focus on selected teacher behaviors was reported in Table 4.12. Teacher perception continued to demonstrate consistency among District Z teacher data sets. Fifty-two percent of District Z teachers' perceived instructional knowledge and skills were focused upon to a great extent, and 50.9% perceived classroom management skills were focused upon to a great extent. Seventy-seven percent of School A

teachers perceived the same two aspects of teacher evaluation were focused upon to a great extent. School C teachers perceived the same two aspects of teacher evaluation were focused upon to a great extent, with 73% perceiving instructional knowledge and skills to a great extent and 70% perceiving classroom management skills to a great extent. Fifty-eight percent of School B teachers perceived instructional knowledge and skills were focused upon to a great extent; however, School B teachers differed from the other teacher data sets on the next aspects. Fifty-seven percent of School B teachers perceived the ability to use data to inform instruction and impact on students' growth were focused upon to a great extent.

School A, B, and C Principals identified a teachers' impact on students' growth as a great extent of the focus of teacher evaluation, and this was the only focus all three principals identified as a great extent of the focus of teacher evaluation. It was clear to principal participants student growth was paramount with the new evaluation system. The teachers of each of the three schools perceived impact on students' growth as a focus of great extent at 53.8%, 56.8%, and 50% respectively. While only School B Teachers rated this focus as a great extent with the second highest percentage, the perception among the majority of teachers clearly recognized the new teacher evaluation system focused to a great extent on a teacher's impact on students' growth.

Table 4.12

Extent of Foci of Teacher Evaluation in 2013-2014 Reported in Percentages

Extent of Focus	N	Not at all	District Z Teachers			
			Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' individual attributes, personality, and dispositions	268	17.5	31	30.6	13.8	7.1
Teachers' subject content knowledge	270	4.4	17.4	31.5	44.1	2.6
Teachers' instructional knowledge and skills	269	1.9	13	30.9	52.4	1.9
Teachers' classroom management skills	271	1.1	11.4	34.3	50.9	2.2
Teachers' relations with parents and students	270	7	25.6	36.7	26.7	4.1
Teachers' relations with other teachers in the school	271	22.5	30.3	30.6	12.2	4.4
Teachers' relations with school administrators	271	17.7	32.8	29.5	14.8	5.2

Teachers' service to the profession, the school, and the district	271	15.5	29.2	31	21	3.3
Teachers' abilities to use data to inform instruction	271	4.8	17.7	31	42.8	3.7
Teachers' impact on students' growth	270	3.3	13.3	31.1	50.4	1.9
Other	67	14.9	1.5	7.5	11.9	64.2

Extent of Focus	School A Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' individual attributes, personality, and dispositions	13	7.7	7.7	61.5	23.1	0
Teachers' subject content knowledge	13	0	23.1	15.4	61.5	0
Teachers' instructional knowledge and skills	13	0	15.4	7.7	76.9	0
Teachers' classroom management skills	13	0	15.4	7.7	76.9	0
Teachers' relations with parents and students	13	15.4	15.4	38.5	30.8	0
Teachers' relations with other teachers in the school	13	30.8	38.5	23.1	7.7	0
Teachers' relations with school administrators	13	23.1	15.4	38.5	15.4	7.7
Teachers' service to the profession, the school, and the district	13	0	30.8	46.2	15.4	7.7
Teachers' abilities to use data to inform instruction	13	0	23.1	38.5	38.5	0
Teachers' impact on students' growth	13	0	15.4	30.8	53.8	0
Other	1	0	0	0	0	100

(table continued)

Table 4.12 (cont.)

Extent of Focus	School B Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' individual attributes, personality, and dispositions	36	19.4	19.4	50	0	11.1
Teachers' subject content knowledge	37	2.7	13.5	27	54.1	2.7
Teachers' instructional knowledge and skills	36	11.1	0	27.8	58.3	2.8
Teachers' classroom management skills	37	0	16.2	27	54.1	2.7
Teachers' relations with parents and students	37	8.1	18.9	43.2	27	2.7
Teachers' relations with other teachers in the school	37	16.2	27	40.5	13.5	2.7
Teachers' relations with school administrators	36	11.1	30.6	36.1	19.4	2.8
Teachers' service to the profession, the school, and the district	37	2.7	32.4	40.5	21.6	2.7
Teachers' abilities to use data to inform instruction	37	0	18.9	21.6	56.8	2.7
Teachers' impact on students' growth	37	2.7	8.1	29.7	56.8	2.7
Other	6	16.7	0	16.7	0	66.7

Extent of Focus	School C Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' individual attributes, personality, and dispositions	29	0	31	31	34.5	3.4
Teachers' subject content knowledge	30	3.3	16.7	20	60	0
Teachers' instructional knowledge and skills	30	3.3	16.7	6.7	73.3	0
Teachers' classroom management skills	30	3.3	16.7	10	70	0
Teachers' relations with parents and students	30	6.7	20	20	50	3.3
Teachers' relations with other teachers in the school	30	23.3	30	30	16.7	0
Teachers' relations with school administrators	29	24.1	34.5	20.7	20.7	0

(table continued)

Table 4.12 (cont.)

Extent of Focus	School C Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' service to the profession, the school, and the district	30	13.3	36.7	26.7	23.3	0
Teachers' abilities to use data to inform instruction	30	3.3	16.7	33.3	46.7	0
Teachers' impact on students' growth	30	3.3	16.7	30	50	0
Other	5	0	0	0	40	60

Extent of Focus	Principals				
	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Teachers' individual attributes, personality, and dispositions	C	A	B		
Teachers' subject content knowledge		C		A	B
Teachers' instructional knowledge and skills		C	B	A	
Teachers' classroom management skills	C	B	A		
Teachers' relations with parents and students	C		A	B	
Teachers' relations with other teachers in the school	C		B	A	
Teachers' relations with school administrators	C		A	B	
Teachers' service to the profession, the school, and the district	C	A	B		
Teachers' abilities to use data to inform instruction			B	A	
				C	
Teachers' impact on students' growth				A	B
					C

District Z teacher data sets and principal perceptions of the extent that a source evidence was emphasized were reported in Table 4.13. District Z, School A, School B, and School C teachers perceived the new teacher evaluation system emphasized formal observations and students' academic growth (state standardized tests and/or SMART goals) to the great extent. Further, there was very little difference in the percentage of teachers' ratings on formal

observations and students' academic growth. District Z, School A, and School B teachers perceived these two evidence sources were focused upon to a great extent within a range of five percent. School C teachers differed only in the number of teachers perceiving these two evidence sources were focused upon to a great extent, with the largest range difference of 36.7% for students' academic growth. The extent teachers perceived formal complaints and grievances were emphasized in the teacher evaluation was higher in the three-subset teacher data sets when compared to the District Z teacher data set.

School A, B, and C Principals identified formal observations and students' academic growth (state standardized tests and/or SMART goals) as emphasized to a great extent with the new teacher evaluation system, making these two evidence points the only two points all three principals agreed were emphasized to a great extent. This perception was aligned exactly with the highest percentage of responding teachers at each school. While the evaluation system recognized other data sources as evidence that could be used to assess a teaching standard, formal observations remained the most emphasized or the second most emphasized source of evidence of teaching effectiveness.

Table 4.13

Extent of Emphasis on Sources of Evidence in Teacher Evaluation Reported in Percentages

Extent of Emphasis	N	District Z Teachers				
		Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Formal observations by principals	271	2.6	14.4	21.8	57.9	3.3
Informal (walkthrough) observations by principals	270	16.3	27.8	33.7	17.8	4.4
Artifacts/actual examples of work provided by teachers	271	21.4	32.5	28.8	11.1	6.3
Evidence provided by colleagues, mentors, and other peers	271	26.6	32.5	25.5	5.5	10
Reports and feedback from parents	270	27	35.9	18.5	6.3	12.2
Responses on student surveys and other feedback from students	270	39.3	29.6	15.9	3.7	11.5
Students' academic growth (state standardized tests and/or SMART goals)	271	3.7	11.8	24.4	55.4	4.8
Formal complaints or grievances	261	29.5	23.4	12.6	6.5	28

(table continued)

Table 4.13 (cont.)

Extent of Emphasis	School A Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Formal observations by principals	14	0	0	35.7	57.1	7.1
Informal (walkthrough) observations by principals	14	0	28.6	50	21.4	0
Artifacts/actual examples of work provided by teachers	14	28.6	35.7	21.4	14.3	0
Evidence provided by colleagues, mentors, and other peers	14	28.6	28.6	28.6	7.1	7.1
Reports and feedback from parents	14	28.6	21.4	7.1	21.4	21.4
Responses on student surveys and other feedback from students	14	35.7	35.7	14.3	0	14.3
Students' academic growth (state standardized tests and/or SMART goals)	14	0	14.3	28.6	57.1	0
Formal complaints or grievances	14	0	42.9	14.3	21.4	21.4

Extent of Emphasis	School B Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Formal observations by principals	37	2.7	13.5	27	56.8	0
Informal (walkthrough) observations by principals	37	8.1	32.4	35.1	18.9	5.4
Artifacts/actual examples of work provided by teachers	36	16.7	38.9	19.4	8.3	16.7
Evidence provided by colleagues, mentors, and other peers	37	24.3	37.8	21.6	8.1	8.1
Reports and feedback from parents	37	27	27	29.7	8.1	8.1
Responses on student surveys and other feedback from students	37	40.5	35.1	13.5	2.7	8.1
Students' academic growth (state standardized tests and/or SMART goals)	37	5.4	8.1	24.3	62.2	0
Formal complaints or grievances	36	44.4	16.7	11.1	11.1	16.7
Formal observations by principals	29	0	13.8	31	51.7	3.4
Informal (walkthrough) observations by principals	30	10	23.3	33.3	26.7	6.7
Artifacts/actual examples of work provided by teachers	28	10.7	35.7	21.4	28.6	3.6

(table continued)

Table 4.13 (cont.)

Extent of Emphasis	School B Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Sure
Evidence provided by colleagues, mentors, and other peers	30	20	43.3	23.3	10	3.3
Reports and feedback from parents	30	16.7	23.3	30	20	10
Responses on student surveys and other feedback from students	30	36.7	16.7	23.3	10	13.3
Students' academic growth (state standardized tests and/or SMART goals)	30	6.7	6.7	43.3	36.7	6.7
Formal complaints or grievances	30	33.3	16.7	20	10	20

Extent of Emphasis	Principals				
	Not at all	Some extent	Moderate extent	Great extent	Not sure
Formal observations by principals				A B C	B
Informal (walkthrough) observations by principals			C	A	B
Artifacts/actual examples of work provided by teachers		C		A B	
Evidence provided by colleagues, mentors, and other peers		C	A B		
Reports and feedback from parents			B	A C	
Responses on student surveys and other feedback from students		A C	B		
Students' academic growth (state standardized tests and/or SMART goals)				A B C	
Formal complaints or grievances	A B	C			

Teacher and principal perceptions of the extent of involvement in the design and modification of the performance-based teacher evaluation system were reported in table 4.14.

There were no indicators of involvement in the design and modification of the teacher evaluation system rated to a great extent above 14%. To analyze this question, it was more revealing to focus on the percentage of teachers rating the extent of involvement at not at all. Of the six indicators, District Z teachers perceived the extent of involvement not at all at a highest percentage of 86.3% and the lowest percentage of 60.9%. School A teachers perceived the extent of involvement not at all at a highest percentage of 100% and the lowest percentage of 66.7%. School B teachers perceived the extent of involvement not at all at a highest percentage of 88.9% and a lowest percentage of 68.4%. School C teachers perceived the extent of involvement not at all at a highest percentage of 82.8% and a lowest percentage of 62.1%. Overwhelmingly, the majority of teachers in each of the data sets perceived their extent of involvement in the design and modification of the teacher evaluation system were not at all.

Similar to the teacher data sets, principals perceived very little involvement in the design and modification of components of the teacher evaluation system. School A Principal perceived a great extent of involvement and modification in the use of data to inform instructional decisions, in the selection of professional development to support teachers' implementation of the new evaluation, and in the procedures for how evidence of teacher effectiveness was collected. School B Principal perceived a great extent of involvement and modification in the procedures for how evidence of teacher effectiveness was collected. School C Principal perceived the extent of involvement and modification was not at all on each of the six components included on the survey instrument and asked in the interview. This perception was shared by the teacher data sets with over 60% of teachers at School A, School B, and School C perceiving no involvement in the design and modification of each of the components of the new teacher evaluation system. Importantly, School A, B, and C principals noted high levels of involvement in the implementation of the performance-based teacher evaluation system at the school site. This perception, if accurate, could indicate great differences in how the new performance-based teacher evaluation system was implemented at the three school sites. Further, if there were great differences at the three school sites targeted for additional study, did those implementation differences exist at the other schools in District Z?

Table 4.14

*Extent of Involvement in the Design and Modification of the Teacher Evaluation System
Reported in Percentages*

Extent of involvement in the design and modification	District Z Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Of the teacher evaluation criteria	263	80.2	11.8	6.1	1.9
Of the specific evidence used in evaluation	262	70.6	15.6	8	5.7
Of the use of data to inform instructional decisions	258	60.9	15.9	13.2	10.1
Of the selection of professional development to support teachers' implementation of the new evaluation	262	78.6	12.6	6.9	1.9
Of the procedures for how evidence of teacher effectiveness is collected	261	77.8	11.9	9.2	1.1
Of tying evaluation results with teacher personnel decisions (promotions, compensation, reassignment, and dismissal)	263	86.3	6.5	5.7	1.5

Extent of involvement in the design and modification	School A Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Of the teacher evaluation criteria	15	66.7	26.7	0	6.7
Of the specific evidence used in evaluation	15	80	13.3	0	6.7
Of the use of data to inform instructional decisions	15	66.7	20	6.7	6.7
Of the selection of professional development to support teachers' implementation of the new evaluation	15	73.3	20	0	6.7
Of the procedures for how evidence of teacher effectiveness is collected	15	66.7	26.7	0	6.7
Of tying evaluation results with teacher personnel decisions (promotions, compensation, reassignment, and dismissal)	15	100	0	0	0

(table continued)

Table 4.14 (cont.)

Extent of involvement in the design and modification	School B Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Of the teacher evaluation criteria	36	88.9	8.3	2.8	0
Of the specific evidence used in evaluation	36	68.4	19.4	2.8	8.3
Of the use of data to inform instructional decisions	35	68.6	17.1	2.9	11.4
Of the selection of professional development to support teachers' implementation of the new evaluation	36	75	16.7	5.6	2.8
Of the procedures for how evidence of teacher effectiveness is collected	36	80.6	13.9	5.6	0
Of tying evaluation results with teacher personnel decisions (promotions, compensation, reassignment, and dismissal)	36	88.9	11.1	0	0

Extent of involvement in the design and modification	School C Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Of the teacher evaluation criteria	29	82.8	10.3	6.9	0
Of the specific evidence used in evaluation	29	62.1	24.1	6.9	6.9
Of the use of data to inform instructional decisions	29	62.1	13.8	10.3	13.8
Of the selection of professional development to support teachers' implementation of the new evaluation	29	72.4	24.1	3.4	0
Of the procedures for how evidence of teacher effectiveness is collected	29	69	10.3	13.8	6.9
Of tying evaluation results with teacher personnel decisions (promotions, compensation, reassignment, and dismissal)	28	82.1	7.1	7.1	3.6

(table continued)

Table 4.14 (cont.)

Extent of involvement in the design and modification	Principals			
	Not at all	Some extent	Moderate extent	Great extent
Of the teacher evaluation criteria	A			
	B			
	C			
Of the specific evidence used in evaluation	B	A		
	C			
Of the use of data to inform instructional decisions	C	B		A
Of the selection of professional development to support teachers' implementation of the new evaluation	C	B		A
Of the procedures for how evidence of teacher effectiveness is collected	C			A
				B
Of tying evaluation results with teacher personnel decisions (promotions, compensation, reassignment, and dismissal)	A			
	B			
	C			

Teacher and principal perceptions regarding the level of confidence in components of the teacher evaluation system were reported in table 4.15. Analysis of perceptions of teacher data sets demonstrated various levels of confidence in activities of the teacher evaluation system. District Z teachers perceived their level of confidence at 57.7% to a great extent on communicating student performance with parents and at 53.9% to a great extent using student academic progress data to adjust teaching. School A teachers perceived their level of confidence at 60% to a great extent on communicating student performance with parents, and 53.3% to a great extent on communicating their teaching with parents. School B teachers perceived their level of confidence highest at 48.6% to a great extent for communicating student performance with parents; and perceived their level of confidence at 37.8% to a great extent on four components using academic progress to adjust teaching, using evaluation results to inform next year's teaching, communicating their teaching with parents, and communicating their teaching with the principal. School C teachers perceived their level of confidence at 51.7% on documenting students' progress using formative assessments, and perceived their level of

confidence at 44.8% to a great extent on using student academic progress to adjust their teaching. While the levels of confidence in activities perceived to a great extent were reflective of several differences among the teacher data sets, it was revealing to analyze the components teacher data sets perceived with the least confidence. Two of the components of the teacher evaluation system – understanding the Virginia Uniform Performance Standards for Teachers and setting SMART goals – were perceived by the highest percentage of teachers from all data sets at the not at all confidence level. It was also important to find that many of the activities teacher data sets perceived at a high confidence level were activities that are deeply embedded in the profession of teaching, such as communicating student performance with parents.

Principals were asked five of the questions teacher data sets were asked regarding the extent of confidence with activities required of the teacher evaluation system. School A principal perceived a great extent of confidence with all five of the question stems including understanding the Virginia Uniform Performance Standards for Teachers, assessing teacher performance on multiple measures, setting SMART goals, using student academic progress data to evaluate teacher performance, and communicating performance with the teacher. School B Principal perceived a great extent of confidence with assessing teacher performance on multiple measures and a moderate extent of confidence on the three other activities of teacher evaluation, and a great extent of confidence communicating performance with the teacher. School C principal perceived some extent of confidence understanding the Uniform Performance Standards for Teachers, a moderate extent of confidence assessing teacher performance on multiple measures, and a great extent of confidence on the other three activities required of the evaluation system. Identifying comparisons among principals and teachers on this survey instrument question was not possible because principals were not asked all of the question stems teachers were asked. However, it was important to note that only one principal perceived a great extent of confidence in setting SMART goals and understanding the Uniform Performance Standards for Teachers, both activities were the activities perceived by the highest percentage of teachers at the not at all level of confidence. Further, the principal perceiving a great extent of confidence in each of the five activities also attended more hours of professional development, which was provided by the VDOE.

Table 4.15

Level of Confidence in Activities of the Teacher Evaluation System Reported in Percentages⁹

Level of confidence in activities of evaluation system	District Z Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	265	28.3	40.4	22.3	9.1
Setting SMART goals	265	15.8	24.2	35.8	24.2
Using student academic progress data to adjust your teaching	267	1.1	12.7	32.2	53.9
Communicating student performance with parents (student achievement and growth)	267	.4	11.6	30.3	57.7
Level of confidence in activities of evaluation system	School A Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	15	26.7	40	33.3	0
Setting SMART goals	15	20	33.3	40	6.7
Communicating your teaching with parents	15	0	6.7	40	53.3
Communicating student performance with parents (student achievement and growth)	15	0	6.7	33.3	60
Level of confidence in activities of evaluation system	School B Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	36	22.2	36.1	36.1	5.6
Setting SMART goals	37	24.3	21.6	45.9	8.1

(table continued)

⁹ Table 4.15 was included in its entirety in Appendix C. Portions were removed from Chapter 4 for ease of reading based on information reported in findings and conclusions.

Table 4.15 (cont.)

Using student academic progress data to adjust your teaching	37	0	8.1	54.1	37.8
Using evaluation results to inform next year's teaching	37	2.7	8.1	51.4	37.8
Communicating your teaching with parents	37	2.7	8.1	51.4	37.8
Communicating student performance with parents (student achievement and growth)	37	0	8.1	43.2	48.6
Communicating your teaching with your principal	37	0	8.1	54.1	37.8
<hr/>					
Level of confidence in activities of evaluation system	School C Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	29	31	37.9	27.6	3.4
Setting SMART goals	29	27.6	31	17.2	24.1
Documenting students' progress using formative assessments	29	6.9	20.7	20.7	51.7
Using student academic progress data to adjust your teaching	29	6.9	10.3	37.9	44.8
<hr/>					
Level of confidence in activities of evaluation system	Principals				
	Not at all	Some extent	Moderate extent	Great extent	
Understanding the Virginia Uniform Performance Standards for Teachers		C	B	A	
Assessing teacher performance on multiple measures			C	A	
				B	
Setting SMART goals			B	A	
				C	

(table continued)

Table 4.15 (cont.)

Level of confidence in activities of evaluation system	Principals			
	Not at all	Some extent	Moderate extent	Great extent
Using student academic progress data to evaluate teacher performance			B	A C
Communicating performance with the teacher				A B C

Teachers and principals were asked to rate the extent of consistency of four belief/value statements to the performance standards, and the findings were reported in table 4.16. An analysis of teacher perceptions to the questions demonstrated comparisons between two teacher data sets. District Z teachers and School B teachers perceived the performance standards were consistent to a great extent with their own views of what teachers should do at 18.7% and 19.4% respectively. The same teacher data sets perceived the performance standards were consistent to a great extent with their schools' basic values at 24.1% and 27.8% respectively. District Z and School B teachers perceived the performance standards were consistent to a great extent with their job description at 24.9% and 22.2% respectively. Finally, these same teacher data sets perceived the performance standards were consistent to a great extent with how they were previously evaluated at 12.3% and 11.1% respectively. The majority of District Z and School B teachers perceived the extent of consistency with these statements at the some extent and moderate extent. School A teachers differed from District Z and School B teachers. School A teachers perceived to a great extent consistency with the performance standards and their own views of what teachers should do at 7.1%, with their school's basic values at 7.1%, with their job description at 7.1%, and none perceived a great extent of consistency with how they were previously evaluated. School C teachers perceived to a great extent consistency with the performance standards and their own views of what teachers should do at 7.1%, with their school's basic values at 17.2%, with their job description at 17.9%, and with how they were previously evaluated at 10.3%. Interestingly, School C teachers perceived not at all consistent at

the highest percentage of the four teacher data sets. Teacher perceptions across all data sets on the consistency of beliefs/values with the performance standards ranged greatly among the four data sets.

School A, B, and C Principals perceived to a great extent the new performance standards were consistent with their own job description; however, this was the only stem all three principals agreed to a great extent. School B Principals perceived some extent of consistency with the new performance standards and his or her own views of what teachers should do, and School A and C Principals perceived a moderate extent of consistency with this stem. School A Principal perceived to a moderate extent consistency with the performance standards and his or her own school's basic values, and School B and C Principals perceived this same stem to a great extent consistent. Each of the three principals perceived a different level of consistency with how teachers were previously evaluated; School A Principal perceived some extent consistent, School B Principal perceived moderate extent consistent, and School C Principal perceived a great extent consistent. Just as teacher perceptions ranged greatly with this survey instrument question, so too did the principal perceptions range. The majority of teachers at School A, B, and C perceived each of the four stems in the some extent to moderate extent consistent with performance standards. This perception aligned most with the School A Principal and least with School C Principal.

Table 4.16

Extent of Performance Standards Consistency Reported in Percentages

Extent of Consistency	N	District Z Teachers				
		Not at all	Some Extent	Moderate Extent	Great Extent	Not Applicable
With your own views of what teachers should do	252	8.7	24.6	43.3	18.7	4.8
With your school's basic values	253	4	24.1	43.9	24.1	4
With your job description	253	4.7	23.7	42.3	24.9	4.3
With how you were previously evaluated	253	7.1	28.9	41.9	12.3	9.9

(table continued)

Table 4.16 (cont.)

Extent of Consistency	School A Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Applicable
With your own views of what teachers should do	14	7.1	50	35.7	7.1	0
With your school's basic values	14	0	50	42.9	7.1	0
With your job description	14	14.3	42.9	35.7	7.1	0
With how you were previously evaluated	14	7.1	50	42.9	0	0

Extent of Consistency	School B Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Applicable
With your own views of what teachers should do	36	16.7	19.4	36.1	19.4	8.3
With your school's basic values	36	5.6	25	33.3	27.8	8.3
With your job description	36	8.3	33.3	27.8	22.2	8.3
With how you were previously evaluated	36	5.6	38.9	33.3	11.1	11.1

Extent of Consistency	School C Teachers					
	N	Not at all	Some Extent	Moderate Extent	Great Extent	Not Applicable
With your own views of what teachers should do	28	28.6	28.6	35.7	7.1	0
With your school's basic values	29	13.8	34.5	34.5	17.2	0
With your job description	28	17.9	39.3	25	17.9	0
With how you were previously evaluated	29	13.8	34.5	37.9	10.3	3.4

Extent of Consistency	Principals				
	Not at all	Some extent	Moderate extent	Great extent	Not applicable
With your own views of what teachers should do		B	A		
With your school's basic values			A	B	
				C	

(table continued)

Table 4.16 (cont.)

Extent of Consistency	Not at all	Some extent	Principals Moderate extent	Great extent	Not applicable
With your job description				A B C	
With how teachers were previously evaluated		A	B	C	

Summary

Data were collected from a waiver school district in the Commonwealth of Virginia to identify teacher perceptions and selected principal perceptions on the performance-based teacher evaluation system. The study examined teacher perceptions by utilizing the larger district teacher data set and comparing those perceptions with the perceptions of teachers at three subset elementary schools. Additionally, teacher perceptions at the subset elementary schools were compared with the perceptions of the principals at each of the subset elementary schools.

This chapter provided a review of the procedures used to collect the data and reviewed the criteria and selection process of the three-subset elementary schools. The subset elementary schools represented schools in the district with similar enrollment, which resulted in comparable staffing standards based on the district’s staffing allocations. The federal data reported annually to the VDOE was evaluated to identify additional similarities and differences among the subset elementary schools. Two of the subset elementary schools were identified as similar on many reporting categories and one subset elementary school differed on several reporting categories.

The study utilized the survey instrument that was developed by Professor Min Sun and R. Brock Mucheson and was part of a larger research project studying the implementation of teacher performance-based evaluation in the Commonwealth of Virginia. However, this study adopted a mixed-method design to answer the research questions, which included interviews of two district directors, the survey instrument for teachers, and interviews of the subset elementary school principals that used matched questions from the teacher survey instrument. Descriptive

statistics were presented in this chapter to explore the research questions. This chapter provided a detailed description of findings and an analysis of the findings.

Chapter Five

Introduction, Executive Summary, Conclusions, Limitations/Delimitations, and Recommendations

Introduction

The purpose of this study was to examine teacher perceptions of the new performance-based teacher evaluation system implemented in the Commonwealth of Virginia. District Z was selected as the district of study because the district implemented the new teacher evaluation system during the 2013-2014 school year, which was delayed one year because the school district received a VDOE approved waiver to prepare for implementation of the new requirements. This chapter contained the executive summary and the conclusions of the research study. Study limitations, delimitations, and recommendations for further study were also addressed in this chapter.

This study was a subset of a larger research project studying the implementation of the new performance-based teacher evaluation system in the Commonwealth of Virginia and utilized the survey instrument developed by Professor Min Sun and R. Brock Mucheson. While the survey data from the population of this study could be utilized in the larger research project, the findings and analysis reported in Chapter Four and the conclusions presented Chapter Five were based on the population described for this study only.

Executive Summary

Much of the research on policy implementation in education was grounded in production theories, which assumed the alignment of organizational and individual values and beliefs resulted in increased production (Akerloff & Kranton, 2005; Youngs, Frank, Thum, & Low, 2012). New research theories on policy implementation suggested how organizations prepared for implementation of new policies and the influence of the setting, formal and informal leaders, and the role of resources and support for schools influenced the success of implementation greatly (Spillane, et. al., 2009; Coburn, 2001, Little, 1982, and Sun, Frank, Penuel, & Kim, 2013). The findings and analysis of this research study supported this conclusion. District Z teachers and principals would have benefitted greatly from professional development on the new performance-based teacher evaluation system that included opportunities to engage in practical implementation strategies, as well as opportunities to gain an understanding of important aspects

of the evaluation tool including the *Uniform Performance Standards* and the use of SMART goals. District Z's investment in the development of human resources geared towards those affected by the new policy, including professional development, time, and materials, would likely increase teachers' and principals' perceptions of the new policy and result in improved practices. This investment in human resources would also allow District Z to capitalize on the collegial environment evident in many of its schools, which also would have the potential to increase the effectiveness of policy implementation. Selecting schools with differences in key reporting characteristics to compare with each other and with the District as a whole suggested principal leadership, specifically principal leadership that was collaborative and engaged teachers in the decision-making process, resulted in perceptions that nurtured a norm of collegiality and like-mindedness. Conclusions of this research study were explained in detail in the following section.

Conclusions

The purpose of this study was to explore perceptions of teachers and selected principals on the new performance-based teacher evaluation system implemented in a school district granted a one-year implementation waiver. A mixed-methodology was chosen to allow the researcher flexibility to interview district leaders regarding preparation and professional development and interview subset principals. The analysis of findings allowed the researcher to explore the research questions and present the following conclusions.

In response to research question one, District Z Directors reported during interviews the district organized a committee to review the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* during the 2012-2013 school year. This group met several times and directed the focus on Standard 7: Student Academic Progress, indicating a strong desire to comply with the law and DOE guidelines. It was evident from both director interviews the group also wanted implementation of the new performance-based standard to be reflective of student growth and achievement. District Z presented the document to the school board for approval and met the objective of the waiver, which was to comply fully with the law and DOE guidelines by July 1, 2013.

In response to research question two, the conclusion of this research study indicated very little professional development was provided to principals of School A, B, and C and District Z teachers, including teachers of School A, B, and C. The two directors reported during interviews

that District Z leadership provided support to principals on the implementation of teacher evaluation and reviewed District Z's board approved new performance-based evaluation system to all principals and assistant principals at the annual summer conference. This conclusion was also supported by the principal interviews and teacher survey instrument responses indicating the number of hours of professional development on teacher evaluation and using data to adjust instruction. Each of the data sets, principals and teachers, reported low totals for professional development. Further, teacher data sets did not perceive the hours of professional development on both topics as very useful. All three principals indicated a desire and need for more professional development on implementation of teacher evaluation during interviews. Policy implementation theories and research suggested resources and support and the ability to activate those resources and support would result in more effective implementation of new policies (Coburn, 2001; Little, 1982; Spillane et al., 2009; Sun, et al., 2013). Based on current research and the analysis provided from this study, District Z should provide professional development on the new evaluation system, and by doing so, perceptions of the new performance-based teacher evaluation system could improve for teacher and principals.

Analysis of findings in response to research questions three and four led to several conclusions among the teacher data sets. District Z teachers, including all data sets, nominated colleagues within their school sites as providing support with teacher evaluation at high percentages suggesting high levels of human resource support and collegiality at the school sites. Little (1982) found that schools with high levels of collegiality, along with a norm of continuous improvement, were schools with high levels of success. The findings of this study reflected high levels of collegiality at the school sites, and this finding could indicate a norm of continuous improvement at all District Z schools. There was a substantial difference in the number of other colleagues, a representative who was not assigned to his or her school site or was not included on the roster, nominated by all teacher data sets indicating an area for improvement for District Z. School C Teachers nominated the lowest percentage of other colleagues, which could be attributed to the academic performance of School C students on the previous SOL assessments. A possible explanation for this could be that District Z leadership was allocating human resources to schools based on student achievement results, which would account for a higher percentage of nomination at School B. However, this explanation does not account for the same percentage of nominated other colleagues at School A as the percentage nominated with the

District Z teacher data set. The findings of this study suggested more support from colleagues outside of the school site, classified as other colleagues in the data analysis, could also add to the perception of collegiality and nurture a culture of continuous improvement.

As mentioned previously, all teacher data sets reported low number of hours and low perceptions of usefulness for the professional development that was provided. Further, all teacher data sets reported perceptions of sufficient time and resources to complete the evaluation to a great extent at low levels. School A, B, and C teachers reported even lower levels of sufficient time and resources at a great extent percentage in comparison to the District Z teachers. District Z teachers, including all data sets, would benefit from access to collegial support from outside the school site, additional time to complete the evaluation, and additional resources to complete the evaluation. Further, policy implementation research (Spillane, et. al., 2009) suggested investing in resources to support teachers such as collegial support, time, and implementation resources could improve teachers' perceptions of the new evaluation system.

Teacher perceptions of School A, School B, and School C shared many common characteristics of the large teacher data set representing District Z teachers. Even though a key feature of this study selected three elementary schools based on similarities with staffing and enrollment and differences in the number of students receiving free/reduced lunch, the number of students receiving special education services, and the school's SOL student achievement results, the data analysis described for each of the four data sets shared many common perceptions. All teacher data sets perceived teachers' instructional knowledge and skills were focused upon to a great extent with the new evaluation system. All teacher data sets perceived informing improvement in classroom instruction was a very important purpose of teacher evaluation. This comparison suggested District Z teacher data sets perceived a high level of like-mindedness with regard to the purpose of teacher evaluation. This like-mindedness, as suggested by Youngs et al. (2012), could improve policy implementation for the schools and the school district.

While research suggested teacher evaluation systems should not focus solely on student achievement (Baker, et.a., 2010, Marzano & Toth, 2013, and Darling-Hammond, et al., 2011), all teacher data sets perceived formal observations by principals and students' academic growth (state standardized tests and/or SMART goal) were emphasized to a great extent with the new evaluation system. This finding demonstrated the shared perception about what was emphasized, but also demonstrated an opportunity to improve the overall evaluation by emphasizing other

aspects of the evaluation system such as artifacts or real examples of work, informal observations, evidence provided by colleagues, reports and feedback from parents, and feedback from students. All teacher data sets reported no involvement in the design and modification of elements of the evaluation system at exceedingly high percentages. All teacher data sets perceived at the highest percentage a low level of confidence understanding the Virginia Uniform Performance Standards for Teachers and setting SMART goals. Further, the level of involvement perceived by teachers in the design and modification strongly suggested the need to engage teachers in meaningful professional development on those topics and engage teachers in decision making at the school and district level on those elements, including knowledge of the uniform performance standards and setting SMART goals.

One distinct area of difference among District Z, School A, School B, and School C teachers was noted regarding the four belief statements. Interestingly, School B was most closely aligned with District Z, while School A and School C were most closely aligned. This alignment of perceptions with the belief statements of District Z and School B was less surprising than the alignment of the perceptions of teachers at School A and School C. Specifically, School A had the highest percentage of students classified as disadvantaged, the highest percentage of students with a disability, the highest minority student population, and the lower student achievement scores on the 2013 and 2014 SOL tests. School C was the elementary school with the lowest percentage of students classified as disadvantaged, the lowest percentage of students with a disability, the lowest percentage of minority students and the highest student achievement scores on the 2013 and 2014 SOL tests, excluding the History SOL averages. While School C teachers, perceived three of the extent of consistency with beliefs higher than School A, School C teacher perceptions were still lower than the two other teacher data sets. While this finding was reflective of teacher perceptions at the three schools and District Z, this perception did differ from what would be expected. It would be expected that teachers of students with higher achievement results would reflect a greater extent of consistency with the four belief statements included on the survey instrument and perceive performance-based teacher evaluation more favorably.

Research questions five and six were crafted to understand principal perceptions of the teacher evaluation system at the three schools selected for additional study and to compare those perceptions to the teachers at each of the targeted schools. The data analysis of matched

interview questions suggested principals' perceptions of the new teacher evaluation system were similar to each other and to the teachers at their schools. During the 2013-2014 year, principals reported hundreds of hours spent on teacher evaluation, indicating the principals' deep understanding of the importance of teacher evaluation and a commitment to implementing effectively and with fidelity.

The three principals perceived informing improvement in classroom instruction as a very important purpose of teacher evaluation, which aligned with teachers at each of the subset schools. This finding suggested principals were communicating, in both words and deeds, the purpose of teacher evaluation to such an extent that the teachers agreed. This normed perception of a common purpose for teacher evaluation could be attributed partly to the attention and resources expended in education since the enactment of NCLB. Schools and districts have been required to report student achievement data for years and a natural outcome of this accountability could be the perception among educators that teachers are responsible for this achievement and teacher evaluation is a method to ensure students continue to achieve. The three principals perceived to either a moderate extent or a great extent the rating teachers received on their evaluation would accurately portray their teaching, indicating an overall positive perception of the new evaluation system.

Principals of School B and School C and teachers at all three schools expressed a desire to increase the extent of confidence in components of the evaluation system. Interestingly, School A Principal expressed the most confidence in components of the evaluation, and School A Principal reported more professional development hours and attending a VDOE professional development. This finding suggested principal confidence in components should be increased and that principals should then share knowledge with the teachers at each school. Principals and teachers of School A, School B, and School C perceptions were aligned regarding the extent activities were emphasized with the new teacher evaluation and the extent of involvement in the design and modification of the new teacher evaluation system. This finding, supported by research, suggested principals of School A, School B, and School C were effectively demonstrating what they emphasized and valued as leaders with the new teacher evaluation system (Sun, et al., 2013). Additionally, this suggested the principals could provide effective leadership to expand what was emphasized in the evaluation system to include an emphasis on artifacts/actual work samples and evidence provided by colleagues, mentors, and other peers.

Additional support of this effective leadership was clearly evident in the number of colleagues teachers at the subset schools nominated as providing support with teacher evaluation, and this conclusion was also supported in research (Sun, et al., 2013, Spillane, et al., 2009, and Little 1982). Collegiality, as evidenced by the teacher participants of this study, occurs in an environment where leadership values and nurtures that norm. Just as the perception of extent of involvement in the design and modification of the teacher evaluation system was suggested as an opportunity to engage teachers, the principal participants of this study echoed the same opportunity. Principals expressed a desire to implement the evaluation system with fidelity and wanted to provide teachers with support.

Engaging principals in the design and modification at the district level would serve as a model for principals to discover ways to engage the teachers at their respective schools in design and modification of the evaluation system and would encourage like-mindedness of beliefs and values with the identity of the organization (Akerloff & Kranton, 2005, Youngs, et al. 2012, and Coburn, 2001). The findings of this study demonstrated principal leadership does influence policy implementation and teacher perceptions (Sun, et al., 2013). Further, despite differences in required reporting categories, students receiving free/reduced priced lunch, student with disabilities, student achievement on SOL tests, and very little professional development on teacher evaluation; School A, B, and C principals and teachers shared many common positive perceptions of teacher evaluation. The perceptions of the participants of this study demonstrated a desire to improve the quality of instruction in an effort to increase student achievement as an important purpose of teacher evaluation.

Limitations/Delimitations

The methodology of this study presented several limitations, which must be considered when evaluating the results. A primary purpose of this study was to explore teacher and principal perceptions in a waiver district in the Commonwealth of Virginia. This purpose presented a limitation to the study because only eight districts in Virginia applied for and received an implementation waiver from the VDOE. Another limitation to consider for this research study was the response rate of teachers within the school district. District Z employed 753 teachers during the 2013-2014 school year who were targeted as participants in the study. Of this sample population, 73 teachers were excluded from participation because the primary researcher served as the primary evaluator of those teachers. Therefore, 680 teachers were targeted as teacher

participants of this study. Three hundred fifty-seven District Z teachers participated in this study representing 53% of the targeted population. Another limitation of this study was because a mixed methodology was selected principals at subset schools were asked matched questions during interviews. Some of those matched questions were adapted to mirror the questions teachers were asked on the survey instrument. Some of the matched questions were not applicable to the principals at all, which is reflected in Chapter 4: Analysis of Findings.

Several delimitations were identified as a result of the methodology of this study. The researcher selected District Z as the school district to study. District Z represented only one of the eight school districts that were implementing the new teacher evaluation system during the 2013-2014 school year. A unique feature of this study included the selection of three subset elementary schools, which enabled the researcher to compare teacher perceptions at the subset schools to District Z teacher perceptions. This feature also enabled the researcher to compare principals' perceptions of teacher evaluation at the subset elementary school and compare those perceptions to the teachers at those subset elementary schools. This methodological decision, while providing deep insight at the subset schools, must be considered a delimitation. This design reduced the number of teacher participants by removing these subset teachers from the larger group. Response rates of teachers at School A represented 43% of the teacher participants, School B represented 95% of the teacher participants, and School C represented 94% of the teacher participants. While teacher response rates at School B and School C were extremely high, higher even than District Z teacher response rates, the teacher response rate at School A was lower than the overall response rate of District Z teachers. Results and findings presented in this study representing School A must be interpreted based on the limited response rate at the school. School A results were included in the analysis and findings despite the 43% response rate because School A represented one of two schools selected based on the percentage of students receiving free/reduced lunch, students receiving special education services, and student achievement results on the SOL tests. The researcher assumed teacher responses at School A would add depth to the study given the response rates at the subset school identified as similar to School A based on the reporting characteristics because School B teacher response rates were much higher. This assumption does not consider other site-based differences based on leadership styles and day-to-day operation of schools. A final delimitation of this study was the researcher

selected elementary schools as subset schools for additional study and comparison with District Z, as opposed to selecting middle or high schools as subset schools.

Recommendations for Further Research

New policy implementation is a costly endeavor. As this study highlighted, many factors can impact perception of new policies. The following recommendations for further study are proposed:

1. Conduct a study of teacher perceptions of the new performance-based teacher evaluation system at a school district or districts within the Commonwealth of Virginia who were not in the implementation year to determine how teacher perceptions at school districts in implementation years compare with school districts that have been implementing for multiple years.
2. Conduct a study of other school districts that are finishing the implementation year of a performance-based teacher evaluation system to compare with the findings of District Z teacher and principal perceptions.
3. Conduct a study of a school district implementing a performance-based teacher evaluation system that provided principal and teacher professional development prior to implementation to explore how professional development influences perceptions.
4. Conduct a quantitative study of principal perceptions of teacher evaluation that includes principals across the district, state, and/or nation. A focus of this study would be to identify the common practices of principals that improved perception of new policy.
5. Conduct a qualitative study that explores the causes of differences in teacher perceptions of performance-based evaluation systems. This study would interview groups of teachers with positive perceptions and groups of teachers with negative perceptions of performance-based evaluations systems in an effort to identify and specify those practices that can positively influence perception.
6. Conduct a study with the same methodology, selecting secondary schools as the subset focus schools to understand how secondary teacher and principal perceptions compare with elementary teacher and principal perceptions.

References

- Aaronson, D., Barrow, L. & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25(1), 95-135.
doi:10.1086/508733
- Akerlof, G., & Kranton, R. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1), 9-32.
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravitch, D., Rothstein, R., Shavelson, R., & Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute
- Battelle for Kids (2011). Selecting growth measures: A guide for education leaders. Retrieved from <http://BattelleforKids.org>.
- Bedford County Virginia Community Development (2007). *Bedford County 2025 Comprehensive Plan*. Retrieved on February 24 2014, from <http://www.co.bedford.va.us/Res/Planning/CompPlan/index.asp>.
- Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service.
- Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added: Report of a workshop*. Washington, DC: The National Academies Press. Retrieved on October 21, 2013 from http://www.nap.edu/catalog.php?record_id=12820.
- Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23, 145-170.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., and Rothstein, J. (2011). *Getting teacher evaluation right: A background paper for policy makers* (278). Washington, DC: American Educational Research Association. Retrieved on June 26, 2013, from www.epi.org.
- Ericsson, K., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406.

- Finnigan, K. & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low performing schools. *American Educational Research Association*, 44(3), 594-629.
- Fuchs, L., Deno, S., & Mirkin, P. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, 21(2), 449-460.
- Haladyna, T. (2011). Using student achievement tests to evaluate teachers – a very bad idea. *Nonpartisan Education Review/Essays*, 7(2). Retrieved June 26, 2013, from <http://npe.educationnews.org/Review/Essays/v7n2.pdf>
- Hanushek, E., Kain, J., O'Brien, D., & Rivkin, S. (2005). *The market for teacher quality*. Cambridge, MA: National Bureau of Economic Research. Retrieved September 28, 2013, from <http://www.nbeer.org/papers/w11154.pdf>.
- Hattie, J. (2003). Teachers make a difference: What is the research evidence? Retrieved September 28, 2013, from http://www.leadspace.govt.nz/leadership/pdf/john_hattie.pdf.
- Howell, W., Peterson, P., & West, M. (2011). Meeting of the minds. *Education Next*, 11(1), 20-31. Retrieved June 26, 2013, from <http://educationnext.org/winter-2011-correspondence/>
- Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26 (11), 101-136.
- Koretz, D. (2008). A measured approach value-added models are a promising improvement, but no one measure can evaluate teacher performance. *American Educator*, 32(3), 18-39.
- Little, J. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, 19, 325-340.
- Lockwood, J., McCaffrey, D., Hamilton, L., Stecher, B., Le, V., Martinez, J. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, 44(1), 47-67.
- Marzano, R., & Toth, M. (2013). *Deliberate practice for deliberate growth: Teacher evaluation systems for continuous instructional improvement*. Palm Beach Gardens, FL: Learning Sciences International Marzano Center Teacher & Leader Evaluation.
- McMillan, J., & Wergin, J. (2010). *Understanding and evaluating educational research* (4th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

- Measures of Effective Teaching Project. (January 2013). *Ensuring fair and reliable measures of effective teaching culminating findings from the MET project's three-year study* (Policy and Practice Brief). Retrieved on December 18, 2013, from <http://www.metproject.org/reports.php>.
- Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.
- Rossman, G. B., & Rallis, S. F. (2011). *Learning in the field: An introduction to qualitative research* (3rd ed.). Thousand Oak, CA : Sage.
- Rowan, B., Chiang, F., & Miller, R. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70(4), 256-284.
- Sanders, W., & River, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Spillane, J. Gomez, L., & Mesler, L. (2009). Notes on reframing the role of organizations in policy implementation. In G. Sykes, B. Schneider, & D. Plank (Eds.), *The AERA handbook of education policy research* (pp.409-425). New York: Routledge.
- Standard 5: Quality of classroom instruction and educational leadership, § 22.1-253.13.5 (2013).
- Stecker, P., Fuchs, L. & Fuchs, D. (2005). Using Curriculum-Based Measurement to Improve Student Achievement: Review of Research. *Psychology in the Schools*, 42(8), 795-819. doi:10.1002/pits.20113
- Stronge, J., Ward, T., Tucker, P., & Hindman, J. (2008). What is the Relationship Between Teacher Quality and Student Achievement? An Exploratory Study. *Journal of Personnel Evaluation in Education*, 20(3-4), 165-184. doi: 10.1007/s11092-008-9053-z
- Sun, M., Frank, K., Penuel, W., & Kim, C. (2013). How external institutions penetrate schools through formal and informal leaders? *Educational Administration Quarterly*, 49(4), 610-644.
- United States Department of Commerce (2010). *United States Census Bureau State and County Quickfacts*. Retrieved on February 25, 2014, from Retrieved from <http://quickfacts.census.gov/qfd/states/51/51019.html>.
- United States Department of Education (2009). *American Recovery and Reinvestment Act of 2009: Using ARRA funds to drive school reform and improvement*. Retrieved on

- December 16, 2013, from
http://www.doe.virginia.gov/home_files/leaving/redirect.cfm?url=http://www.ed.gov/policy/gen/leg/recovery/guidance/uses.doc.
- United States Department of Education (2013). *ESEA Flexibility*. Retrieved on October 27, 2013, from <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>
- United States Department of Education (2010) *U.S. Department of Education American Recovery and Reinvestment Act report: Summary of programs and state-by-state data*. Retrieved on December 18, 2013, from <http://www2.ed.gov/policy/gen/leg/recovery/spending/arra-program-summary2.pdf>.
- Virginia Department of Education. (2012). *ESEA flexibility request*. Retrieved June 26, 2013, from http://www.doe.virginia.gov/federal_programs/esea/flexibility/
- Virginia Department of Education. (2011). *Guidelines for uniform performance standards and evaluation criteria for teachers*. Retrieved June 26, 2013, from http://www.doe.virginia.gov/teaching/performance_evaluation/teacher/
- Virginia Department of Education. (2011). *The research base for the uniform performance standards for teachers*. Retrieved June 26, 2013, from http://www.doe.virginia.gov/teaching/performance_evaluation/teacher/.
- Virginia Department of Education Teacher and Principal Evaluation Collection (TPEC) (2013). Retrieved from <https://p1pe.doe.virginia.gov/ssws/login.page.do>.
- Virginia Department of Education Educational Information Management 2012-2013 Spring Student Record Collection 10 – Bedford County Public Schools (2013).
- Virginia Department of Education. (2009). *American Recovery and Reinvestment Act of 2009 use of funds summary*. Retrieved on December 18, 2013, from http://www.doe.virginia.gov/school_finance/arra/reporting/.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Retrieved November 5, 2013, from <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.
- White, B. (2004). *The relationship between teacher evaluation scores and student achievement: Evidence from Coventry, RI*. Madison, WI: Consortium for Policy Research in Education.
- Youngs, P., Frank, K., Thum, Y., & Low, M. (2012). The motivation of teachers to produce human capital and conform to their social contexts. In T. Smith, L. Desimone, & A.

Porter (Eds.), Yearbook of the National Society for the Study of Education: Vol. 110. *Organization and effectiveness of high-intensity induction programs for new teachers* (pp. 248-272). Malden, MA: Blackwell Publishing.

Appendix A

IRB Approval



Office of Research Compliance
Institutional Review Board
North End Center, Suite 4120, Virginia Tech
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website <http://www.irb.vt.edu>

MEMORANDUM

DATE: April 17, 2014
TO: M. David Alexander, Min Sun, Michelle Lee Morgan
FROM: Virginia Tech Institutional Review Board (FWA00000572, expires April 25, 2018)
PROTOCOL TITLE: Implementation of Reform with a Performance-Based Teacher Evaluation System:
A Case Study of One School District
IRB NUMBER: 14-367

Effective April 17, 2014, the Virginia Tech Institutional Review Board (IRB) Chair, David M Moore, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

<http://www.irb.vt.edu/pages/responsibilities.htm>

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: **Expedited, under 45 CFR 46.110 category(ies) 5,6,7**
Protocol Approval Date: **April 17, 2014**
Protocol Expiration Date: **April 16, 2015**
Continuing Review Due Date*: **April 2, 2015**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

MEMORANDUM

DATE: September 29, 2014
TO: M. David Alexander, Michelle Lee Morgan, Richard Salmon, John K Burton
FROM: Virginia Tech Institutional Review Board (FWA00000572, expires April 25, 2018)
PROTOCOL TITLE: Implementation of Reform with a Performance-Based Teacher Evaluation System:
A Case Study of One School District
IRB NUMBER: 14-367

Effective September 29, 2014, the Virginia Tech Institutional Review Board (IRB) Chair, David M Moore, approved the Amendment request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

<http://www.irb.vt.edu/pages/responsibilities.htm>

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: **Expedited, under 45 CFR 46.110 category(ies) 5,6,7**
Protocol Approval Date: **April 17, 2014**
Protocol Expiration Date: **April 16, 2015**
Continuing Review Due Date*: **April 2, 2015**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

Appendix B
School Approval Letter



Bedford County Public Schools

P.O. Box 748
310 S. Bridge Street
Bedford, VA 24523
Ph. 540-586-1045

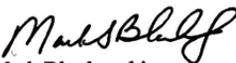
April 18, 2014

Dear Ms. Morgan,

The purpose of this letter is to provide written approval to conduct the research study entitled "Implementation of Reform with a Performance-Based Teacher Evaluation System: A Case Study of One School District. This approval grants permission to interview and survey participants and provides access to federal and state reporting information for Bedford County Public Schools, as outlined in the approved Virginia Tech IRB #14-367.

Best of luck to you as you begin this research project.

Sincerely,


Mark Blankenship
Supervisor of Assessment and Planning

Appendix C
Level of Confidences Table

Table C1

Level of Confidence in Activities of the Teacher Evaluation System Reported in Percentages

Level of confidence in activities of evaluation system	District Z Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	265	28.3	40.4	22.3	9.1
Knowing multiple measures of your teaching	267	6	27.7	38.6	27.7
Collecting sufficient evidence to demonstrate our effectiveness	266	1.5	26.7	39.5	32.3
Setting SMART goals	265	15.8	24.2	35.8	24.2
Documenting students' progress using formative assessments	266	.4	15.8	35	48.9
Using student academic progress data to adjust your teaching	267	1.1	12.7	32.2	53.9
Using evaluation results to inform next year's teaching	266	2.3	13.9	36.1	47.7
Communicating your teaching with parents	267	.7	14.6	32.2	52.4
Communicating student performance with parents (student achievement and growth)	267	.4	11.6	30.3	57.7
Communicating your teaching with your principal	266	.8	14.3	34.6	50.4

(table continued)

Table C1 (cont.)

Level of confidence in activities of evaluation system	School A Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	15	26.7	40	33.3	0
Knowing multiple measures of your teaching	15	0	40	53.3	6.7
Collecting sufficient evidence to demonstrate our effectiveness	15	0	33.3	40	26.7
Setting SMART goals	15	20	33.3	40	6.7
Documenting students' progress using formative assessments	15	0	13.3	53.3	33.3
Using student academic progress data to adjust your teaching	15	6.7	6.7	46.7	40
Using evaluation results to inform next year's teaching	15	0	20	33.3	46.7
Communicating your teaching with parents	15	0	6.7	40	53.3
Communicating student performance with parents (student achievement and growth)	15	0	6.7	33.3	60
Communicating your teaching with your principal	15	0	6.7	46.7	46.7

Level of confidence in activities of evaluation system	School B Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	36	22.2	36.1	36.1	5.6
Knowing multiple measures of your teaching	37	16.2	16.2	48.6	18.9
Collecting sufficient evidence to demonstrate our effectiveness	37	8.1	10.8	59.5	21.6
Setting SMART goals	37	24.3	21.6	45.9	8.1
Documenting students' progress using formative assessments	37	0	10.8	54.1	35.1

(table continued)

Table C1 (cont.)

Level of confidence in activities of evaluation system	School B Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Using student academic progress data to adjust your teaching	37	0	8.1	54.1	37.8
Using evaluation results to inform next year's teaching	37	2.7	8.1	51.4	37.8
Communicating your teaching with parents	37	2.7	8.1	51.4	37.8
Communicating student performance with parents (student achievement and growth)	37	0	8.1	43.2	48.6
Communicating your teaching with your principal	37	0	8.1	54.1	37.8
Level of confidence in activities of evaluation system	School C Teachers				
	N	Not at all	Some extent	Moderate extent	Great extent
Understanding the Virginia Uniform Performance Standards for Teachers	29	31	37.9	27.6	3.4
Knowing multiple measures of your teaching	29	13.8	37.9	27.6	20.7
Collecting sufficient evidence to demonstrate our effectiveness	29	13.8	34.5	24.1	27.6
Setting SMART goals	29	27.6	31	17.2	24.1
Documenting students' progress using formative assessments	29	6.9	20.7	20.7	51.7
Using student academic progress data to adjust your teaching	29	6.9	10.3	37.9	44.8
Using evaluation results to inform next year's teaching	29	6.9	17.2	41.4	34.5
Communicating your teaching with parents	29	6.9	20.7	34.5	37.9
Communicating student performance with parents (student achievement and growth)	29	6.9	17.2	34.5	41.4
Communicating your teaching with your principal	29	6.9	10.3	44.8	37.9