Spring 2016

Teacher and Principal Perceptions of A New Evaluation Program for Teachers

Ruth Shannon Finnegan
College of William and Mary, rsfinnegan@email.wm.edu

TEACHER AND PRINCIPAL PERCEPTIONS OF A NEW EVALUATION PROGRAM FOR TEACHERS

A Dissertation

Presented to

The Faculty of the School of Education

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

---

by

R. Shannon Finnegan

February 2016

TEACHER AND PRINCIPAL PERCEPTIONS OF A NEW EVALUATION PROGRAM FOR TEACHERS

by

R. Shannon Finnegan

Approved February 2016 by

Megan Tschannen-Moran, Ph.D.

Chairperson of Doctoral Committee

James H. Stronge, Ph.D.

Leslie W. Grant, Ph.D.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES vii

CHAPTER 1: INTRODUCTION 2
- Background 2
- Program Description 8
- Overview of the Evaluation Approach 15
- Summary 22

CHAPTER 2: REVIEW OF RELATED LITERATURE 24
- Policies Related to Teacher Evaluation 25
  - No Child Left Behind 25
  - Race to the Top 26
- Effect of Effective Teaching on Students 29
- Effective New Evaluation Programs for Teachers 31
  - Components of Effective Teacher Evaluation Programs 33
  - Student Growth 35
  - Description of the Joint Committee on Standards for Educational Evaluation 39
- Significance of Teacher and Principals Perceptions 47
- Perceptions of Teachers and Principals 53
- Summary 57

CHAPTER 3: METHODS 59
- Participants 59
- Data Sources 61
- Data Collection 61
- Data Analysis 62
- Limitations 63
- Ethical Considerations 64
- Summary 66

CHAPTER 4: RESULTS 67
- Response Rate 67
- Description of the Sample 67
- Reliability of the Scales in Study 70
- Analysis of the Research Questions 70
  - Research Question 1 71
  - Research Question 2 78
  - Research Question 3 82
- Summary 87

CHAPTER 5: DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS 89
Discussion of the Findings 90
Research Question 1 91
Research Question 2 103
Research Question 3 107
Conclusions 108
Implications 114
Recommendations for Practice 117
Recommendations for Policy 124
Recommendations for Further Research 125
Summary 126

APPENDIX A: PERMISSION TO USE SURVEYS 127

APPENDIX B: QUESTIONNAIRE 128

APPENDIX C. CORRESPONDENCE OF QUESTIONNAIRE ITEMS TO SURVEYS BY HOPKINS (2013) AND STRONGE (2013) 131

APPENDIX D: ALIGNMENT OF RESEARCH QUESTIONS, SURVEY QUESTIONS, AND JCSEE PERSONNEL EVALUATION STANDARDS 135

APPENDIX E: CONSENT FORM 142

REFERENCES 143

VITA 169

ACKNOWLEDGEMENTS

C. S. Lewis tells us, “You are never too old to set another goal or to dream a new dream.” As a life-long learner, I dreamed of studying at the doctoral level. For many years, this dream did not manifest itself due to a variety of circumstances. Years passed and the dream seemed further out of reach until I discovered the opportunity to pursue my dream through The College of William and Mary’s executive doctorate program. With the encouragement of my husband and daughters, I dared to pursue this dream. Nevertheless, the success of this endeavor did not lie with me alone; it was a direct result of many individuals.

This journey would have never been completed without my committee. These individuals not only helped me see that I could accomplish my dream, but also provided me the skills and competency to complete my dissertation. Dr. Megan Tschannen-Moran, chair of my dissertation committee, provided me direction, guidance, and encouragement. Dr. James Stronge’s patience, insights, and expertise kept my study moving forward. Dr. Grant provided feedback and direction. I also wish to acknowledge Dr. Margaret Constantino, my academic advisor, for her consistent encouragement that helped me overcome many of the obstacles set before me. I am also grateful to Drs. Michael DiPaola, Jennifer Parrish, Steven Staples, and Gene Rouche for their positive influence, both academically and professionally, during my studies at William and Mary.

During my time at William and Mary, I also had the pleasure to work with classmates I value and respect. These cohorts challenged and pushed me to become a better educator. I would like to thank Steven, Earl, Mike, Natalie, Cheryl, Jill, and Phil for their comradery, friendship, and support. I especially would like to thank Bloodine

Bob-Semple for helping me see, despite my doubts, that I could complete this program. I also want to thank my friends outside my academic world who were my constant cheerleaders and prayer partners–Christy, Andrea, Kathy, Ginny, Leslie–as well as family members and many other individuals who encouraged me.

Without the assistance of Dr. Kim Hill, this study would not have been completed. Because of her courage to investigate educators' perceptions of the new teacher evaluation program in the district, I was able to conduct the study. I would also like to acknowledge the help of the research and assessment department in the district and the teachers and principals who responded to my survey. Special thanks to principals Marvin Jones and Melissa Logan; they were supportive and understanding of my commitment during this journey. Thank you also to Dr. Lynda Idleman for her statistical and editorial support.

My dream of completing this program would never have been launched without the persistent and unfailing love and support of my wonderful husband, Michael, and two wonderful daughters, Olivia and Alexandria. Michael was my rock in this journey. Much as Sam Wise Gamgee was to Frodo, Mike would not let me give up, regardless of how overwhelmed I felt. He was a continuous source of encouragement. I am blessed that he is my husband and helpmate. Finally, in the words of J. S. Bach, "Soli Deo Gloria" (To the Glory of God Alone), for without the Lord's provision and grace none of this would have been accomplished.

LIST OF TABLES

Table		Page
1	Rating Scale for Qualitative (Professional Practice) Component of the Teachers Evaluation	12
2	Emerald County School District's Teacher Evaluation Domains	13
3	Attributes and Requirements of the JCSEE Standards	18
4	Research on Effects of Effective Teaching on Student Learning	31
5	Description of Propriety Standards	42
6	Description of Utility Standards	44
7	Description of Feasibility Standards	45
8	Description of Accuracy Standards	46
9	JCSEE Attributes and Items on Questionnaire	62
10	Data Analysis Plan	63
11	Description of the Sample	68
12	Age and Years of Experience of Teachers and Administrators in Study Sample	69
13	Reliability of the Scales Measuring Propriety, Utility, Feasibility, and Accuracy Attributes	70
14	Means and Standard Deviations of Teachers and Administrators' Ratings of the JCSEE Attributes	72
15	Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Propriety	73
16	Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Utility	74
17	Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Feasibility	76
18	Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Accuracy	77

19 Differences in Teachers and Administrators' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels 79

20 Differences in Principals' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels 80

21 Differences in Teachers' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels 81

22 Means and Standard Deviations of Teachers and Administrators' Ratings of JCSEE Attributes and Items Measuring Student Learning Objectives and Professional Practices 82

23 Differences in Teachers and Administrators' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System 83

24 Comparison of Ratings of JCSEE Attributes Within Teachers and Administrators 85

25 Comparison of Ratings of Student Learning Objectives Items Within Teachers and Administrators by JCSEE Attributes 86

26 Comparison of Ratings of Professional Practice Items Within Teachers and Administrators by JCSEE Attributes 87

ABSTRACT

The ability to evaluate teachers accurately is indispensable for both the development of effective teachers and for student achievement. In this era of accountability, it is important school districts develop evaluation systems that comply with the propriety, utility, feasibility, and accuracy standards of the Joint Committee on Standards for Educational Evaluations. This study focused on a process evaluation of a new teacher evaluation program. While previous studies have been conducted from the teachers or the evaluators' perspectives, this study examined both perspectives. More than 1,500 teachers and 41 principals were invited to complete an online instrument modified from surveys conducted by Hopkins and Stronge. Of concern to the teachers and principals was the accuracy of the new evaluation program, they did not see the value and validity of using SLOs to improve teaching practices to increase learning, and teachers slightly favored using the professional practices component more than the SLO component of the evaluation. Using SLO data in teacher evaluation is an unknown dynamic for teachers; therefore, school administrators need to understand how teachers perceive this change as it relates to teacher support of the new evaluation process. If districts are to safeguard the fidelity, implementation, and sustainability of new evaluation programs for teachers, districts must acknowledge the influence teacher perceptions have on endorsing implementation efforts toward change. Teachers' perceptions toward adjusting instructional practices to align with the standards and criteria of new evaluation programs can either hinder or ensure program implementation.

R. SHANNON FINNEGAN

SCHOOL OF EDUCATION

THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA

TEACHER AND PRINCIPAL PERCEPTIONS OF A NEW EVALUATION PROGRAM FOR TEACHERS

CHAPTER 1: INTRODUCTION

The ability to evaluate teachers accurately is indispensable, not only for the development of effective teachers but also for student achievement (Danielson, 2011; Darling-Hammond, 2014; Hanushek, 2011; W. L. Sanders & Rivers, 1996; Stronge & Tucker, 2003; Stronge, Ward, Tucker, & Hindman, 2008). In this era of educational reform and accountability, it is important that school districts develop evaluation systems that comply with the standards of the Joint Committee on Standards for Educational Evaluations ([JCSEE], 2009). Because of the high stakes involved, school systems must be diligent in constructing quality educator evaluation systems reflective of the JCSEE.

Background

The debate over school and teacher accountability is fueled by the public's concern over the gap between students who receive an effective teacher in a quality school and those who do not. Chenoweth (2010) reported that African American and Latino children by age 17 receive a level of education comparable to 13-year-old low-income White children. Chenoweth asserted that, "African American and Latino children are much less likely to graduate from high school or enter college, and once there, they are less likely to graduate" (p. 2). Overall, America's educational history has consisted of what Chenoweth referred to as sorting, "Instead of educating all students…schools sorted their children into different categories, each with their own educational opportunities" (p. 2). Due in large part to the findings of such research, the American public is beginning to call for legislation to remedy these inequities for students in U.S. public schools.

Seminal research studies (Aaronson, Barrow, & Sanders, 2007; Coleman et al., 1966; Nye, Konstantopulos, & Hedges, 2004; W. L. Sanders & Rivers, 1996; Stronge et al., 2008) reported that teacher effectiveness is a prevailing element in student achievement. Hanushek (1992, 2011) found that the achievement level of students learning under the most effective teachers out-measured peers learning under the least effective teachers by as much as one grade level. W. L. Sanders and Rivers (1996) found that students from low-income families benefit the most in learning from highly effective teachers. W. L. Sanders and Rivers's research also indicated that the consequences of learning under an ineffective teacher are indelible; students under the tutelage of ineffective teachers who were later assigned to effective teachers did not compensate for earlier gaps. Despite the research demonstrating teacher effectiveness on students, Weisberg et al. (2009) reported that less than 1% of teachers were rated unsatisfactory, although large percentages of their students were failing. Weisberg et al. reinforced the need for restructuring teacher evaluation systems to recognize the degrees of teacher effectiveness and distinctive strengths while providing resources for developing instructional practices.

One benefit of an accurate teacher evaluation system is improving teacher and administrative effectiveness, which ultimately results in improved student achievement. Danielson (2002) suggested that, "One of the significant influences on a school's culture is its system for teacher evaluation" (p. 35). While debate continues over which tools best measure teacher accountability for student performance, reformers agree it is essential that a fair, accurate, and legal evaluation system be created (DiPaola & Hoy, 2012). Stronge (2010b) contended that, "For evaluation to be fair and comprehensive it is

necessary to describe the performance standards of teachers with sufficient detail and accuracy so that both teachers and their supervisors can reasonably understand the job expectation" (p. 4). Through the evaluation process, school administrators can set measurable goals and objectives for teachers. This way, stakeholders can be assured that the curriculum is being taught in such a way as to help all students be successful.

Well-constructed teacher evaluation systems that include professional learning and development opportunities and measures of student growth can contribute to improving teacher effectiveness and, in turn, raise student achievement. Research by Hanushek (2010) and Danielson (2007) found that reliable and valid measurements for identifying teacher quality must be capable of distinguishing the performances of teachers with respect to the achievements of their students. Effective teacher evaluation systems consist of clear sets of standards and competencies integrated with broader assessments as part of an evaluation framework (DiPaola & Hoy, 2012). These evaluation systems are based on multiple measurements for providing timely feedback in order to give teachers opportunities to put into practice specific ideas for improving instruction (Darling-Hammond, 2013; DiPaola & Hoy, 2012; Drago-Severson & Blum-DeStefano, 2014; Stronge, 2010b). An effective teacher evaluation system provides professional development that aligns with the identified needs of teachers for developing communities of learners within and among schools. Finally, an effective evaluation system highlights improvement and aligns with a system of formative evaluations that influence the summative effect of the evaluation outcomes (Danielson & McGreal, 2000; Darling-Hammond, 2014; DiPaola & Hoy, 2012; Stronge, 2010b; Tomlinson, 2007).

Traditionally, the appraisal of a teacher's instructional practices rested on subjective summative observations made by school administrators and few teacher evaluation programs incorporated measurable outcomes of student achievement (DiPaola & Hoy, 2012; Peterson, 2000). However, in 2010, the Reauthorization of the Elementary and Secondary Education Act guidelines instructed states to use multiple measures to evaluate teacher effectiveness, including a strong emphasis on the growth in achievement of their students (U.S. Department of Education, 2010a). These guidelines require states to use the results of student achievement testing to measure teacher effectiveness in order to be eligible for federal funding. States also must implement rigorous teacher evaluation programs and use the results of teacher evaluations to improve teacher effectiveness and school performance.

Current educational policies are propelling researchers to investigate practices for increasing student outcomes by improving and evaluating effective teaching practices. Recent funding from the federal government has prompted a renewed focus on the implementation and evaluation of models of teaching effectiveness (Barry, 2010). These efforts to implement and evaluate methods of measuring teacher effectiveness have led state departments of education to submit statewide plans to address the issue.

The desired outcome of teacher evaluation programs is effective teachers who improve student achievement. However, in efforts to create effective evaluation programs, the perceptions of the teachers being evaluated and the administrators conducting the evaluations are often not taken into consideration (Behrstock-Sherratt, Rizzolo, Laine, & Friedman, 2013; DiPaola & Hoy, 2012; Johnson, 2012; Ovando, 2001; Stronge & Tucker, 1999). Muñoz, Scoskie, and French (2013) contended that teachers

"are the closet to the learning and learning action; incorporating teachers'…voices in the important debate around teacher effectiveness and evaluation" (p. 228) is important. Many of the debates in the field of evaluation are about what assumptions we make when we construct knowledge and about the nature of many fundamental concepts that we use in our work, like causation, generalization, and truth (Shadish, Campbell, & Cook, 2002). Policies governing the new evaluation programs require change to the status quo; this change potentially brings polarizing perceptions to both administrators and teachers involved in the new evaluation programs. Therefore, it is important to understand the difference in the perceptions of teacher and principals toward the new evaluation program for teachers.

Understanding how teachers construct meaning to educational reform such as new evaluation programs for teachers can provide valuable insight to those implementing new policies or programs. Additionally, educational leaders must understand the change process in order to implement and sustain the new evaluation program for teachers. Fullan (2011) suggested that for leadership to be effective in the change process it has to (a) have an explicit purpose that creates a sense of making a difference, (b) mobilize people to find solutions to difficult problems, (c) use indicators of success that are measurable, and (d) be assessed "to the extent to which it awakens people's intrinsic commitment" (p. 20). Educational leaders need to not only address the policies of the new evaluation systems, but also engage teachers in hearty discussion, debate, feedback, reflection, resolution, and ultimately ownership of the change in order to sustain the reform policies (Cousins & Leithwood, 1986; Fink & Stoll, 1996; Fullan, 2011; Roussin & Zimmerman, 2014). Fullan (2011) reported that when radical change is initiated, an

organization needs a leader who "welcomes differences, communicates the urgency of the challenges, talks about the broad possibilities in an inviting way, and creates mechanisms that motivate people to reach beyond themselves" (p. 47).

School cultures experiencing change produce highly charged emotions, especially from those affected by change. Emotions have a significant influence on teachers' reaction to the various educational reform efforts ranging from compliance to conflict or opposition (Fullan, 2011; Hargreaves, 2001; Roussin & Zimmerman, 2014; Schmidt & Datnow, 2005; van den Berg, 2002). Several researchers highlighted the significance of differences in interpretations of change or reform efforts developing from teachers' prior knowledge and experiences (Hill & Grossman, 2013; Spillane, Reiser, & Reimer, 2002; van den Berg, 2002). Spillane et al. (2002) found that the differences predicted more the level of implementation than of teachers' outright rejection of the reform.

Effective leaders are able to frame teachers' differences, resistance, and dissent into a "potential source of new ideas and breakthroughs" (Fullan, 2011, p. 74). Johnson (2012) provided insight for navigating through these dilemmas by understanding polarities, independent pairs of value sets that can support or undermine a common purpose. Polarities can reflect *both/and* rather than *either/or* thinking as teacher and administrators evolve in their perceptions of new evaluation programs for teachers. Johnson suggested that as an alternative to treating a polarity as a problem to be solved, the wisdom of each pole should be leveraged and the goal should be to find what both sides have in common.

Overhauling and implementing a new evaluation system for teachers requires executing a myriad details and actions. Creating an evaluation that encourages teachers'

growth and development, meets those requirements of the State, and garners buy in from the teachers and those affected is tedious. Therefore, an understanding of the perceptions of those most affected (teachers and principals) by teacher evaluation reforms is crucial for implementing and sustaining the success of the new evaluation program for teachers. The purpose of this quantitative study was to assess the degree to which teachers and evaluators concurred that the teacher evaluation program met the propriety, utility, feasibility, and accuracy attributes of the JCSEE.

Program Description

All students in all school systems across the United States deserve an effective teacher who is capable of providing students with sound instructional practices resulting in student and school improvement. Teacher quality and evaluating teacher quality is at the forefront of educational debate and policies. Toch (2008) reported that, "Teacher evaluations are at the very center of the education enterprise and can be catalysts for teacher and school improvement" (p. 32). For this to be achieved, school districts must provide a method for evaluating educators' capacity to improve learning and achievement for all students. Educators must be effective in closing the achievement gap for all students.

Context. The Education Reform Act of 2010 initiated a new requirement for educators: to be effective, teachers and principals must show they can successfully improve student learning. This law required that changes in student growth become a significant factor in the evaluation of teachers and principals. It also created the foundation for a new evaluation system for teachers that consistently and fairly identifies, supports, and rewards effective educators. Conversely, this new evaluation system for

teachers identifies, develops, or dismisses those who are ineffective (Behrstock-Sherratt et al., 2013; Darling-Hammond, 2014; Maryland Department of Education, 2012). In response to this legislation, the governor of a state on the eastern seaboard of the United States required the state board of education to establish general standards for performance evaluations for certified teachers. The resultant model included performance evaluation criteria consisting of student growth measures and professional practices. It also mandated that each county board have the flexibility to create its own evaluation criteria based on local needs and standards collaboratively agreed upon by the local school district and the local teachers association (Maryland Department of Education, 2012; U.S. Department of Education, 2010a).

The Race to the Top (RTT) legislation initiative necessitates that, for states to obtain federal money, they must reform the teacher evaluations to include evidence of student achievement and professional practices as a significant factor in determining teacher effectiveness (U.S. Department of Education, 2010a). Since 2010, educational leaders at both the state and local levels have studied various approaches to calculating student growth while attributing that growth to individual teachers and principals in educator evaluation programs (DiPaola & Hoy, 2012; Hanushek & Rivkin, 2010; Maryland Department of Education, 2012; Muñoz et al., 2013; Stronge, 2010b; U.S. Department of Education, 2010a).

In complying with the requirements for this initiative, Emerald County School District (a pseudonym), a suburban school district in the eastern seaboard state created a model for educational evaluation that includes two measures: one qualitative (professional practice) and the other quantitative (student growth). This model considers

the evaluation of teachers as a formative continuous cycle that promotes growth for both teachers and students. The developers asserted that this model for the new teacher evaluation system provides for fair, equitable, and continuous improvement of teaching practices by strengthening the knowledge, skills, and classroom practices of educators (Maryland Department of Education, 2012). By using this model for teacher evaluation and enhancing teacher effectiveness, it was hoped that student achievement would improve.

Emerald County School District is located in close proximity to a large urban area. Many of the county's residents work in the nearby urban area. The school district has 21 elementary schools, 7 middle schools, and 5 high schools. The student population of 26,000 is African American (80%), Caucasian (15%), Latino (5%), and other ethnicities (5%).

While some researchers in the field argue that teacher evaluation programs are ineffective and unsuccessful in changing teachers' behavior over time (Darling-Hammond, 2000; Donaldson, 2012), other researchers examining the effects of feedback from teacher observations provide essential suggestions for changing teachers' instructional practices and perceptions (Anast-May, Penick, Schroyer, & Howell, 2011; Pizzi, 2009; Sartain, Stoelinga, & Brown, 2011). The Emerald County School District developed an evaluation system model aligned with the state-mandated requirements of the Education Reform Act. The assumption behind Emerald County's new model of teacher evaluation was that, if there is a teacher evaluation system that is fair, equitable, and results in the continuous improvement of practice, there will be evidence of student growth (Maryland Department of Education, 2012). This model for educator evaluation

includes two measures: a qualitative measure (professional practice) and a quantitative measure (student growth), each comprising 50% of the overall evaluation.

Description of the program. The state requires all of the newly revised teacher evaluation models to provide qualitative measures for four domains of professional practice: (a) planning and preparation, (b) instructional delivery, (c) classroom environment and management, and (d) professional responsibilities. This qualitative measure of overall professional practice must account for half of the entire evaluation. Emerald County, along with the majority of other counties in this state, uses Danielson's (2002) framework for teaching as the protocol to measure the four required domains. The Danielson framework for teaching model is grounded in Shulman's (1987) research on pedagogical content knowledge and the Interstate New Teacher Assessment and Support Consortium (1992) standards. The constructs of both Vygotsky's (1978) social constructivism and zone of proximal development and Piaget's (1952) theories of development provide a constructivist lens by which learners are considered active participants not only in the classroom, but in the learning process. The Danielson Framework for Teaching Evaluation instrument (2013) contains 22 components in four domains: (a) planning and preparation, (b) classroom environment, (c) instruction, and (d) professional responsibilities. The four domains use a scoring rubric that articulates characteristics of the scoring criteria classified as *unsatisfactory*, *basic*, *proficient*, and *distinguished*.

Emerald County School District devised a system for calculating the rating for the professional practice components of the county's teacher evaluation program. Within each of the four domains of this framework are 22 components describing distinctive

aspects of the specific domain (Danielson, 2007). Emerald County School District teacher evaluators assign a point value to each component within the domains for rating teacher performance (Table 1). This system awards 1 point for *ineffective*, 2 points for *developing*, 3 points for *effective*, and 4 points for *highly effective*. The rating scale reflects the percentage of total possible points received in each domain (Table 1). The points received for the qualitative (professional practice) component of the county's teacher evaluation program are calculated in the final teacher evaluation that determines the teacher's overall performance rating (Table 2).

Table 1

Rating Scale for Qualitative (Professional Practice) Component of the Teacher Evaluation System

Rating category	Point value	Percentage of possible points received in each domain
Ineffective	1	0–30
Developing	2	31–49
Effective	3	50–81
Highly effective	4	82–100

In compliance with federal and state requirements, Emerald County School District incorporates student learning objectives (SLOs) to serve as the quantitative component for measuring student growth in the teacher evaluation. SLOs are defined as specific, rigorous, long-term goals for groups of students that educators identify to guide instructional and administrative efforts (Maryland Department of Education, 2012). Emerald County recognizes the complexities in establishing SLOs reflective of high, yet attainable, expectations for students and teachers in both tested and non-tested subjects.

Table 2

Emerald County School District's Teacher Evaluation Domains

Domain	Component	Weight
Planning and preparation	Demonstrating knowledge of content and pedagogy Demonstrating knowledge of students Setting instructional outcomes Demonstrating knowledge of resources Designing coherent instruction Designing student assessments	10%
Classroom environment and management	Creating an environment of respect and rapport Establishing a culture for learning Managing classroom procedures Managing student behavior Organizing physical space	15%
Instructional delivery	Communicating with students Using questioning and discussion techniques Engaging students in learning Using assessment in instruction Demonstrating flexibility and responsiveness	15%
Professional responsibilities	Reflecting on teaching Maintaining accurate records Communication with families Participating in a professional community Growing and developing as a professional Showing professionalism	10%

Emerald County School District takes a formative approach to implementing SLOs. This process provides training for teachers and administers in data analysis (both individually and collaboratively), identifying areas for student growth, and making data-driven instructional decisions that increase student achievement and close the achievement gap. This process begins at the start of the school year with the identification and setting of learning goals for students, as well as the determination of ways to measure the progress in achieving the learning goal. Any supports the teacher may need, such as professional development, are identified at this phase. Teachers then meet midyear with the evaluator

to determine any mid-course corrections. They meet again at the end of the year to discuss the outcome of student learning goals.

Emerald County's experience reveals a challenge in implementing and maintaining the professional practice portion of the teacher evaluation program (Danielson, 2007). While the student growth measure component of the evaluation is calculated annually, the professional practice requirement has a 3-year cycle option for tenured and effective teachers. Therefore, Emerald County School District established three groups of teachers for a continuous rolling evaluation.

Evaluation of nontenured teachers occurs annually, until the attainment of tenure, using both the professional practice and student growth measures. In the event a first-year teacher's test data are missing (data is used from the previous year), the composition of the evaluation is modified. Tenured teachers receive an evaluation containing both the professional practice and student growth measures once every 3 years. Tenured teachers in the second year and third year of the evaluation cycle use the current student growth measures coupled with a carryover score of the professional practice rating from Year 1 of their evaluation cycle. Any teacher rated *ineffective* or *developing* during the previous year's evaluation cycle receives an evaluation using both professional practices and student growth measures annually until an *effective* rating is achieved.

Emerald County School District, in conjunction with the State, created a professional development program to train a team of educational practitioners in each local school district. The aim of this local district cadre was to provide support and technical aid within the structure of each local district's timetable. A major emphasis of this professional development was to train both district and school evaluators to work

collaboratively with educators in developing SLOs that address student achievement gaps, instructional needs for all students, and supports for educators through professional development (Maryland Department of Education, 2012). Working together, both the evaluator and educator develop rigorous and achievable SLO targets, aligned with school and district improvement goals and with the state curriculum framework. Those SLO goals inform professional development programs that help practitioners meet their SLO goals.

Overview of the Evaluation Approach

Understanding the perceptions of teachers and principals is an essential element for engaging them in the discussion on educational policy reform, in particular new evaluation programs for teachers. The Recognizing Educational Success, Professional Excellence, and Collaborative Teaching (RESPECT) project launched in 2012, by the U.S. Department of Education signifies a juncture in elevating teachers' roles in shaping their own profession. Having teachers and principals weigh in on the discussions and decisions of evaluation will only strengthen these policies (Behrstock-Sherratt et al., 2013). The purpose of this quantitative study is to assess the degree to which teachers and evaluators in Emerald County School District concur that the teacher evaluation program meets the propriety, utility, feasibility, and accuracy attributes of the JCSEE (2014). Implementing a new evaluation for teachers will require systematic changes from the status quo. Bridges and Bridges (2009) suggested that the success or failure of change is predicated on how the stakeholders involved react to the change—if they do things differently. For change to be successful, stakeholders must change both their perspectives and actions (Kilgore & Reynolds, (2011).

For successful change to occur it must systemically begin within the organizational culture. In his theoretical framework for understanding organizational culture, Schein (2004) defined culture as:

> A pattern of shared basic assumptions that was learned by a group as it solved its problems of external adaptation and internal integration that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems. (p. 17)

Research highlights the effect of school principals in changing the status quo of school culture toward reform efforts. Principals can affect the professional community, organizational learning, and trust of the organization, as well as the effectiveness of the school toward growth and change, by providing meaningful opportunities for collaboration within the school context and by exhibiting a positive attitude and using an effective method for teacher improvement (Darling-Hammond, 2013; Maslow & Kelley, 2012). The importance of school administrators' actions should not be underestimated, particularly in the area of communicating effective feedback to teachers for developing professional growth (Halawah, 2005; Marshall, 2013; Ovando & Ramirez, 2007).

The focus of the study was on both the teachers and school administrators of the Emerald County School District. The study's evaluation approach is grounded in Alkin's (2004) use branch theory and followed the context, input, process, and product (CIPP) model developed by Stufflebeam (1968, 2007). Stufflebeam and Shinkfield (2007) outlined six components of evaluation: (a) coherence, (b) core concepts, (c) hypotheses on how evaluation procedures produce desired outcomes, (d) workable procedures, (e) ethical requirements, and (f) a framework for guiding program evaluation practices.

Distinct characteristics define the philosophical framework of the use branch theory for evaluation (Mertens & Wilson, 2012). These characteristics are (a) common sense and practical thinking are the basis of discovery, (b) the value of a program evaluation lies in the way the evaluation results are used, (c) evaluation discoveries should make a difference to the organization's stakeholders, (d) evaluators and stakeholders form relationships to the extent necessary to achieve the purpose of the evaluation, and (e) the methodology directly relates to the specific questions and purposes of the research. Mertens and Wilson (2012) defended using the use branch theory, "Because its assumptions align closely with the idea of use of evaluation findings as a priority" (p. 89). Therefore, the pragmatic paradigm and Alkin's (2004) use branch theory of program evaluation provided a foundation for this program evaluation.

Program evaluation model. The CIPP model contains the following: (a) the context of the program including an overview of background information outlining the process and components of new evaluation program for teachers at the school district level and how the program results are used; (b) the inputs of the program including the program's available resources; (c) the key program processes or activities of both teachers and administrators; and (d) the program's short-term, long-term, and intermediate outcomes for both teacher and student. One of the major purposes of a program evaluation is to determine areas in need of improvement or practices that need to change (Danielson, 2007; Darling-Hammond, 2014; DiPaola & Hoy, 2012; Drago-Severson & Blum-DeStefano, 2014; Stronge, 2010b). Using a process evaluation for this task informs the stakeholders about implementation of the process, materials, and other aspects of the program procedures. The process evaluation focuses on why the

anticipated results were or were not reached and what needs to be altered if the results are not effectively attained (Mertens & Wilson, 2012). The research activities in a process evaluation increase the likelihood of the program's success by providing indications of what happened and why.

Purpose of the evaluation. The study was designed to reflect the JCSEE requirements that personnel evaluations be ethical, fair, useful, feasible, accurate, and offer special attention to concerns of diversity (JCSEE, 2009). The intent of these standards is not to promote individual districts' specific evaluation programs, but rather to provide safeguards in order that personnel evaluation programs deliver a comprehensive method most likely to produce the desired outcomes. These standards are characterized by sound educational evaluation techniques and practices reflecting attributes of propriety, utility, feasibility, and accuracy (JCSEE, 2009). Table 3 contains an outline of the attributes and requirements of each standard.

Table 3

Attributes and Requirements of the JCSEE Standards

Attribute	Requirement of the attribute
Propriety	Are conducted legally, ethically, and with due regard for the welfare of evaluees and clients involved in the evaluation.
Utility	Guide evaluations so that they will be informative, timely, and influential.
Feasibility	Easy to implement as possible, efficient in their use of time and resources, adequately funded, and viable from a number of other standpoints.
Accuracy	Require that the obtained information be technically accurate and that conclusions be linked logically to the data.

The purpose of this quantitative study was to assess the degree to which teachers and evaluators concur that the teacher evaluation program meets the propriety, utility, feasibility, and accuracy attributes of the JCSEE. This study will add value to and complement the county's ability to monitor quality implementation of the teacher evaluation program by providing useful information to assist with ongoing program implementation and improvement. This study was initiated without any preconceived assumptions. However, if it is discovered that the perceptions of the evaluation process by both teachers and evaluators in this school district are aligned, and that the system is viewed positively, then the school system will likely continue implementation relatively unchanged. If, on the other hand, it is viewed negatively by either group of participants, or there are large differences in the perceptions of the two groups, then this study may precipitate further study to adapt the process in ways that are best suited for the district.

Mertens and Wilson (2012) reported that the CIPP model "has provided a new perspective for evaluators, moving them away from a way of thinking based on social science research and toward recognition of the need to consider stakeholders and their need for information" (p. 110). This research study addressed the process (sometimes called implementation) component of Stufflebeam's CIPP evaluation model. Focusing on this aspect of the CIPP model, the theoretical framework included a theory-based approach. While conventional evaluations have become synonymous with proving the effectiveness of a program, the CIPP program evaluation process also allows organizations to assess issues such as program implementation and improvement (Stufflebeam, 2007). Connell, Kubisch, Schorr, and Weiss (1995) found that traditional evaluations could actually have a negative influence on the broader and multifaceted

elements of a program as traditional evaluations do not take into consideration the untidy ways that the program's undertaking may affect change.

This research will be of interest to teachers, principals, and decision makers in the Emerald County School District. Decision makers who serve in research and assessment, instruction, and administration will find the outcomes of this research valuable to share with members of the school board. Moreover, decision makers and stakeholders at the state level in the area of educator evaluation may find this research noteworthy.

Focus of the evaluation. The Emerald County School District recently implemented a new evaluation program for teachers. Therefore, it was fitting to focus on a process evaluation designed for a new or changing program. Understanding why a program is or is not successful is critical to successfully maintaining that program—more critical than simply knowing that the program works. The objective of this evaluation was to provide program leaders with how the teachers and school administrators perceive the propriety, utility, feasibility, and accuracy of the evaluation program in order to improve the effectiveness of the program. This process evaluation focused on the appropriateness and quality of the program's implementation from the perspective of both the teachers and evaluators. While former studies have been conducted from either the teachers or the evaluators' perspectives, this study examined both perspectives.

Evaluation questions. It is vital for educational evaluations to be grounded in a research-based framework that adheres to the standards proposed by the JCSEE. JCSEE standards address areas of propriety, utility, feasibility, and accuracy. This study attempts to assess each of these areas by answering the following research questions:

1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date?
    a. What are the perceptions of the propriety of the evaluation system as implemented to date?
    b. What are the perceptions of the utility of the evaluation system as implemented to date?
    c. What are the perceptions of the feasibility of the evaluation system as implemented to date?
    d. What are the perceptions of the accuracy of the evaluation system as implemented to date?
2. Are there differences between teachers at different levels (middle, elementary, and high) in their perceptions of the Emerald County School District evaluation system as implemented to date?
    a. Are there differences in the perceptions of the propriety of the evaluation system as implemented to date?
    b. Are there differences in the perceptions of the utility of the evaluation system as implemented to date?
    c. Are there differences in the perceptions of the feasibility of the evaluation system as implemented to date?
    d. Are there differences in the perceptions of the accuracy of the evaluation system as implemented to date?

3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the evaluation system as implemented to date?
    a. Are there differences in the perceptions of the propriety of the evaluation system as implemented to date?
    b. Are there differences in the perceptions of the utility of the evaluation system as implemented to date?
    c. Are there differences in the perceptions of the feasibility of the evaluation system as implemented to date?
    d. Are there differences in the perceptions of the accuracy of the evaluation system as implemented to date?

Summary

The education system is experiencing the challenges of change in seeking to transform its method of evaluating teachers. Contemporary researchers have found inequity between achievement for students across ethnic, racial, and economic populations and economists and entrepreneurs raise concerns that today's students are not prepared to compete globally (Chenoweth, 2010; W. L. Sanders & Rivers, 1996). Greater emphases on causations of variances in student achievement have led researchers to a more discreet examination of the variables that significantly influence student learning (Hanushek & Rivkin, 2010; Stronge, 2010b). As an outcome, legislators created laws to remedy educational inequity. Thus, the creation of NCLB legislation that mandated states to impose high-stake standardized testing of all students and to require teachers to obtain requisite credentials in order to be consider highly qualified to teach.

Research regarding the influences on student achievement has identified the quality of the teacher as the most significant school factor in increasing student learning. Criticisms of NCLB's high-stake testing mandates have created new legislation, such as the Race to the Top federal grants, which provides incentives (as opposed to sanctions found in NCLB) to states wanting federal money. RTT's grant requirements require states to overhaul their teacher evaluation systems to include student achievement measures. Therefore, Emerald County School District's evaluation system for teachers incorporates multiple measures, both quantitative and qualitative, in its evaluation tool. This system aligns with the requirements set forth by its governing educational body and RTT.

CHAPTER 2: REVIEW OF RELATED LITERATURE

The ability to evaluate teachers accurately is indispensable not only for the development of effective teachers, but also for student achievement (Danielson, 2011; Darling-Hammond, 2014; Hanushek, 2011; Hanushek & Lindseth, 2009; W. L. Sanders & Rivers, 1996; Stronge & Tucker, 2003; Stronge et al., 2007). The public, business leaders, politicians, and economists worry that our current student body will not be equipped to compete in the 21st century global economy. As a result, evaluating teachers' effectiveness on student achievement has become a priority for the nation's education system. New reform policies are creating and implementing high-stakes evaluation programs to measure teacher performance. In this era of educational reform and accountability, it is important for school districts to develop evaluation systems that comply with the standards set forth by the JCSEE (2009). These new evaluation systems bring both intended and unintended consequences; therefore, school systems must be diligent in constructing quality educator evaluation systems reflective of the JCSEE.

This review of literature contains four sections to provide readers with an examination of principal elements for reforming the current teacher evaluation system. The first section discusses the national polices leading to the current educational reform movement. The second section contains a description of effective teaching as well as the effect that effective teaching has on students. The third section contains a review of the literature regarding the purpose and components of an effective teacher evaluation programs. Additionally, this section provides a discussion of the JCSEE (2009, 2014b)

standards and personnel evaluation standards. These standards are pertinent to the literaure review as they frame the program evaluation questions. A review of research on teacher perceptions of the new evaluation programs for teachers is in the final section.

Policies Related to Teacher Evaluation

Coleman et al.'s (1966) seiminal research informed the American public that teacher characteristics explained more variance in student achievement than any other school factor. Wechsler and Shields (2008) reported that, "The quality of a student's teacher is the most important determinant of learning after family background" (p. 1). Student learning and academic growth, or the lack thereof, are the result of teachers' instructional practices. Teaching and learning are at the core of educational practice, and teacher quality is the most important school-level factor affecting student achievement (Darling-Hammond, 2000; Hanushek & Lindseth, 2009). Hanushek (1992) found that the students learning under the most effective teachers outperformed those of their peers learning under the least effective teachers by as much as one grade level.

No Child Left Behind

The need for educational reform became apparent with the continued widening of the achievement gap between minorities and White students. Federal initiatives and policies began to stress the significance of teacher effectiveness and student achievement. Thus, the largest federal education program, the Elementary and Secondary Education Act (more commonly known as No Child Left Behind of 2001), was created. No Child Left Behind (NCLB) spotlighted the gap in achievement and mandated educational reforms to address the gap. The NCLB mandate changed the educational paradigm, requiring schools to cultivate talent and ability in all children. NCLB explicitly stated that

schools would be held accountable for producing evidence that all students were learning. Boykin (2011) reported that, "Despite its flaws, NCLB constituted a radical break with the idea that the relationship between race and intellect was immutable" (p. 7).

Research also points to the affect that quality teaching has on the achievement gap. Kovach and Manning (2003) noted that, "Increasingly, evidence shows that the spread and sustainably of new and improved approaches to teaching and learning require new professional and social norms and normative structures that are foreign to many schools" (p. 40). Quality teaching was identified in an analysis of the National Assessment for Education and Progress as the most powerful influence on academic achievement (Wenglinsky, 2002). NCLB highlighted the importance of teacher quality by listing explicit requirements for highly qualified teachers (U.S. Department of Education, 2004). Additionally, NCLB's standard-based reforms demanded that all students succeed, hence requiring all state educational agencies to provide state accountability testing to all students (U.S. Department of Education, 2004).

NCLB depended on federal mandates requiring compliance at the state level (McGuinn, 2006), thereby forcing states to change many of their educational practices. Nevertheless, these changes were less substantive due to gaps both in ability and in political conflict at the state level. These conflicts resulted in a law that did not produce significant levels of school improvement or progress in closing the student achievement gaps (Mintrop & Sunderman, 2009; Sunderman, 2010).

Race to the Top

Even after the enactment of NCLB, resolving the education dilemma was no easy task, and the political debate on reforming the U.S. education system continued

(McGuinn, 2012). A new legislative initiative known as Race to the Top (RTT) emerged in 2009. The policymakers who designed RTT not only contemplated the effect of legislation such as NCLB, but also,

> The enormously difficult task of driving systemic change in a fragmented and decentralized education system. The newness of and the political opposition to federal efforts to push systemic education reform on the states, and the weakness of state and federal administrative capacity in education. (McGuinn, 2012, p. 138)

RTT created a competitive grant process to provide states with incentives for driving educational improvements for students and schools; instead of the state sanctions found in NCLB (McGuinn, 2012). RTT funds only those states that show robust trajectories and strategies for educational reform innovations. Furthermore, states must demonstrate the commitment of stakeholders to the reform efforts that the states outline.

While RTT provides for funding incentives that drive reform, it still faces a "difficult institutional situation, the limited capacity of federal and state education agencies to push reform down to the school level" (McGuinn, 2012, p. 138). While there is much diversity among the states in their education systems, there also exist remarkable differences in school quality within the states. Despite the fact that RTT explicitly expresses educational goals at the national level, there is no centralized system for following those objectives; thereby, leaving the federal government with only the capacity to drive reform indirectly through the grant-in-aid system (Cavanagh, 2011).

Traditionally, the appraisal of a teacher's instructional practices rests on subjective summative observations from school administrators (DiPaola & Hoy, 2012; Popham, 2013). Legislative policies began to question the worth of these evaluations of

teaching effectiveness (Darling-Hammond, 2013; Popham, 2013; U.S. Department of Education, 2010a). The RTT initiative requires schools to use multiple measures to gauge teacher effectiveness, with a substantial portion coming from student growth data (U.S. Department of Education, 2010a). The RTT initiative also provides school districts the capacity for using teacher evaluation results "to inform human capital decisions such as professional development, compensation, promotion, retention, tenure, and removal" (U.S. Department of Education, 2010a, p. 34). The Reauthorization of the Elementary and Secondary Education Act guidelines require states to use the results of student achievement to measure teacher effectiveness in order to be eligible for federal funding. States should use multiple measures to evaluate teacher effectiveness with a strong emphasis on the growth in achievement of their students (U.S. Department of Education, 2010a). States also must implement rigorous teacher evaluation programs and use the results of teacher evaluations to improve teacher effectiveness and school performance. Furthermore, states receiving federal funding must incorporate a substantial portion of student achievement data into teacher evaluation. In the past, few teacher evaluation programs incorporated measurable outcomes of student achievement (Peterson, 2000).

Using student growth data as a component of teacher evaluation has created controversy in the education community (Darling-Hammond, 2013; DiPaola & Hoy, 2012; Muñoz, Prather, & Stronge, 2011). In response, researchers are investigating how to improve student outcomes by evaluating effective teaching (Danielson, 2013; Darling-Hammond, 2013; DiPaola & Hoy, 2012; Hanushek & Lindseth, 2009; Muñoz et al., 2011). States are creating a variety of models for evaluating the instructional practices of teachers. Many states are currently developing statewide goals and accountability

systems that target high levels of achievement for all students. However, for such actions to be successful, it appears that teachers, rather than legislators, need to adopt a goal of high achievement for all students (Danielson, 2013; Muñoz et al., 2011). Teachers' readiness to adopt such goals is related to teachers' outcomes and expectations (Tollefson, 2000). The inclusion of student achievement in the assessment of a teacher's ability to educate is central to school reform. As school districts investigate different models of evaluation, they should not ignore the effect these models might have on teacher and administrators' perceptions of the evaluation systems.

Effect of Effective Teaching on Students

Recent educational policies, such as NCLB, emphasized the need for highly qualified teachers. Previously teacher characteristics such as experience, advanced degrees, and credentials were identified as substitutions for teacher quality (Ahn, 2013). Clotfelter, Ladd, and Vigdor (2007) found these variables to have weak to moderate positive correlation with higher student achievement. While these characteristics explain part of the effect of teachers, they fail to account for all of the observed variation in achievement (Goldhaber & Anthony, 2007). However, Goldhaber and Brewer (2001) demonstrated significant student gains in math were related to the teacher's major and level of higher education. More recently, researchers using hierarchical linear modeling have found statistically significant positive results for teacher experience; a valuable predictor of student learning at the classroom level (Muñoz et al., 2011).

The setting for current educational reform recognizes that students must be taught by effective teachers. Years of research show a direct relationship between teacher quality/effectiveness and student learning (Creemers & Kyriakides, 2008; DiPaola &

Hoy, 2012; Goldhaber & Anthony 2007; Hanushek, 2010; Hattie, 2009). Identifiable characteristics of teachers that were predictive of their success in the classroom have been found (Darling-Hammond & Youngs, 2002; Wayne & Youngs, 2003).

Teachers affect how students learn, what students learn, and how much students learn (Stronge, 2007, 2010a). Teacher effectiveness is a strong predictor of student achievement (Darling-Hammond, 1996, 2000; Hanushek & Lindseth, 2009; Sanders & Rivers, 1996; Stronge et al., 2008). The work of Sanders and Rivers (1996) established that teacher effectiveness is the central component of student growth. Early research studies by Sanders and Rivers found that heterogeneity among classroom contexts did not contribute to student achievement and that students of various ethnic groups responded equally to effective teachers. Rivkin, Hanusheck, and Kain (2005) concluded that differences in teacher quality showed a difference of 7.5% in student achievement.

Research studies confirm that teacher effectiveness or lack thereof have a residual effect on students; the quality of teachers matters when it comes to how much students learn and the affect of their teachers' effectiveness remains with students for many years (Stronge, Ward, & Grant, 2011; Tucker & Stronge, 2005). Moreover, the effect of this instruction becomes stable over time, not influenced by student/teacher assignments or by use of selected test scores (Aaronson et al., 2007; Mendro, Jordan, Gomez, Anderson, & Bemby, 1997). Thererfore, for student performance to improve, the emphasis must be directed toward improving teacher performance and capacity to provide effective instructional practices (DiPaola & Hoy, 2012; Hanushek, Kain, & Rivkin, 1998; Sanders & Rivers, 1996**;** Stronge, 2010b). Table 4 contains research on the effect of effective teaching on student learning.

Table 4

Research on Effects of Effective Teaching on Student Learning

Authors	Results
W. L. Sanders & Rivers (1996); Wright, Sanders, & Horn (1997); Sanders & Horn (1998)	As teacher effectiveness increased, student gains increased beginning with lowest achievers first and average achievers next. Heterogeneity among classroom contexts did not contribute to student achievement and that students of various ethnic groups responded equally to effective teachers. The residual effect of teachers are cummulative, even after 2 years.
Mendro, Jordan, Gomez, Anderson, & Bemby (1997)	Teacher effectivness on student achivement becomes stable over time, not influenced by student/teacher assignments or by use of selected test scores.
Rivkin, Hanusheck, & Kain (2005)	Differences in teacher quality showed a difference of 7.5% in student achievement

Effective New Evaluation Programs for Teachers

The purpose of evaluating teachers is dependent on the perspective of who is asked. The legislators creating policy may view the evaluation of teachers as a means to remove inept teachers, while a school administrator may view the evaluation system as a means to determine the teacher's skills for instructing a specific group of students. Parents may agree that the evaluation process offers information about the quality of the teacher in teaching a particular content, and finally, a teacher may perceive the evaluations system as a means to provide support for professional growth and improvement in the value of instruction in the classroom. There is debate even among leading researchers in the field of education regarding the purpose for teacher evaluation. Ellett and Teddlie (2003) stated that,

> During the past three to four decades, the question about appropriate means and ends for education in the USA has been strongly reflected in concerns about (a)

> producing, selecting, and assessing effective teachers and (b) understanding linkages between effective teaching, teacher evaluation, school effectiveness and ultimately effective schools. (p. 102)

Hanushek and Rivkin (2010) believed that districts using evaluations for removing the lowest 5% to 10% of ineffective teachers each year will cause an increase in student achievement. While other researchers in the field of education view the purpose of teacher evaluation as improving the quality of teaching and cultivating an excellent supply of good teachers through strong professional and career development (Danielson 2007; Darling-Hammond, 2014; Popham, 2013; Stronge, 2010b). More than improving individual teacher development, a need exists to create and sustain collegial working conditions; allowing teachers to work collectively in a supportive environment that sustains learning for them and their students. Darling-Hammond (2014) stated that,

> The country needs a teacher evaluation as part of a teaching and learning system that supports continuous improvement, both for individual teachers and the profession as a whole. We should not adopt an individualistic, competitive approach to ranking and sorting teachers that undermines the growth of the learning communities. (p. 5)

Many researchers advocate and support using teacher evaluation as a means of improving teacher quality and link teacher evaluation to student achievement (Stronge, 2007, 2010b; Stronge, Gareis, & Little, 2006; Stronge & Tucker, 1999, 2003; Stronge, Tucker, & Hindman, 2004; Tucker & Stronge, 2005; Tucker, Stronge, & Gareis, 2002). Teacher evaluations can function as either a summative or a formative appraisal. The purpose of the summative evaluation is for quality assurance through credentialing,

promoting, providing tenure, and demoting or dismissing teachers. The purpose of the formative design is to promote the professional growth and development of teachers; informing teacher performance through feedback, building capacity for new instructional practices, and modifying current instructional practices (Danielson & McGreal, 2000; Namaghi, 2010).

Components of Effective Teacher Evaluation Programs

The most significant school factor in student performance is a teacher's instructional practice; therefore, it is critical to examine the process by which teachers' performance and contributions to student performance is evaluated (Danielson, 2007). Danielson and McGreal (2000) found that evaluations based on standards comprised of clearly defined performance indicators can have a positive affect on teacher effectiveness. The performance indicators incorporated multiple measures, such as classroom observations and student achievement data. Tucker and Stronge (2005) asserted that communication, collaboration, and commitment are essential elements of any teacher evaluation model. Tucker and Stronge further suggested that for the model to have value for both the teacher and the district, the model must (a) align its goals to the goals of the district, (b) base the evaluation on clearly defined job duties, (c) differentiate between achievement levels for each duty, (d) use multiple sources of data, (e) use a rubric for clear dialogue, and (f) maintain a clear focus on teacher growth and accountability.

Evaluations based on single event data points such as yearly observations are limited. Stronge (2007) contended that using observations as the only evidence of a teacher's work provides a flawed and misleading picture of instructional performance. A comprehensive picture of the teaching process needs to be reflected in teacher

evaluations. According to Stronge, an accurate and trustworthy evaluation will provide numerous measures of performance evaluation through organized and rigorous effort.

With respect to teacher effectiveness, the ultimate evidence is student results and measurable outcomes (Stronge, 2007). The U.S. Department of Education (2010b) allowed the states to develop their definition of teacher effectiveness, which must be based in part on student growth. Therefore, combining state and federal guidelines, teacher effectiveness can be defined as the ability of a teacher to increase student learning based on measurable outcomes (Stronge, 2007). Measurable outcomes of teacher effectiveness may include teacher evaluations linked to student growth based on assessments aligned to common sets of standards (U.S. Department of Education, 2010b). Both Stronge (2007) and Danielson and McGreal (2000) advocated for differentiated evaluations that recognize teachers' varying levels of expertise. The primary purpose of this differentiated evaluation system is to generate usable and reliable data to inform administrators how best to provide supports that are focused on teachers' needs and expertise (Danielson & McGreal, 2000; Stronge, 2007).

Using a comprehensive evaluation system to evaluate teachers also highlights teachers' professional needs and potential instructional improvement. A number of studies have been conducted regarding the relationship between teacher evaluation and student achievement. For example, Milanowski (2004) suggested that standards-based teacher evaluation systems based on the framework for teaching (Danielson, 2007) appear to have the potential to provide measurements of teacher effectiveness that may be strongly related to student achievement. Milanowski (2004) conducted a large-scale study comparing teacher evaluation scores based on Danielson's framework for teaching and

student achievement on standardized state exams. The results indicated that teacher evaluation scores are positively related to higher than expected levels of achievement. Studies conducted by Kimball, White, Milanowski, and Borman (2004) and Borman and Kimball (2005) reported similar findings.

Frameworks that incorporate standards-based teacher evaluations in measuring teacher performance can effectively link teaching behaviors to student achievement and define a competency model for effective teaching (Heneman, Kimball, & Milanoskwi, 2006). Darling-Hammond (2013) found few evaluations offered opportunities for teachers to set goals, receive useful feedback, and have a system that could support learning and timely effective personnel decisions. Stronge and Tucker (2003) asserted that, "Communication, collaboration, and commitment are essential elements of any teacher evaluation model" (p. 65). Stronge and Tucker further suggested that the model must have six components to have value for both the teacher and the school teacher: (a) the evaluation model must align its goals to the goals of the district, (b) the evaluation is based on clearly defined job duties, (c) achievement levels for each duty are differentiated, (d) multiple sources of data are used, (e) a rubric is used for clear dialogue, and (f) a clear focus on teacher growth and accountability is maintained.

Student Growth

Using evaluation measures that incorporate data on the achievement of the teacher's students is a central component of the new reform (e.g., Danielson, 2007; Gates Foundation, 2013; Milanowski 2004). Student growth models could be defined as "a collection of definitions, calculations, or rules that summarizes student performance over two or more time points and supports interpretations about students, their classrooms,

their educators, or their schools" (Castellano & Ho, 2013, p. 16). The National Governors Association (2011) categorized the growth models used by various states into five types: (a) categorical, (b) gain score, (c) regression, (d) value-added, and (e) normative. Castellano and Ho (2013) also identified three primary growth model interpretations: (a) growth description, (b) growth prediction, and (c) value-added. Growth description offers a growth metric related to the extent of growth for an individual or group (Auty et al., 2008). Growth prediction provides information about the future scores of students given current and past achievements (Bonk et al., 2012). Value-added measures offer information about what causes growth, for example specific teachers and schools (Ryser & Rambo-Hernandez, 2014).

Some evaluation systems include value-added measures in their teacher evaluation models as a means to determine student gains in learning for that year. Value-added measures represent each teacher's effect on student learning during the prior year and describe one important dimension of the teacher's effectiveness (Danielson, 2007; Gates Foundation, 2013; McCaffrey, Lockwood, Koretz, & Hamilton, 2004; Milanowski, 2004.). However, this value-added approach presents several factors to be considered in evaluating teachers. For instance, most student outcomes tend to focus on a narrow set of educational goals, such as math and reading achievement (Rothstein, Jacobsen, & Wilder, 2008). Furthermore, estimates of effectiveness based on such measures may be biased as a result of students' demographic characteristics, nonrandom assignment of students to teachers, student mobility, alignment of assessments to criteria such as Common Core State Standards and potential incomparability of gains across grades (Amrein-Beardsley, 2008; Davis, Chopin, Drake, & McDuffie, 2014; McCaffrey et al., 2004).

Finding sound frameworks for measuring growth beyond existing growth or value-added models is complex. One option to using the growth models that rely on sophisticated statistical methods for attributing student achievement growth to teachers is student learning objectives (SLO). The SLO is an academic goal that educators establish for each individual or subgroup of students (Marion, DePascale, Domaleski, Gong, & Diaz-Biello, 2012). SLOs use a teacher and principal's awareness and expectations of individual student growth during a school year; therefore, the SLO target is collaboratively determined by the teacher and the principal. Measures such as SLOs also have the capacity for accurately assessing teachers in non-tested grades and subjects by allowing for a more extensive assessment of the influences of all teachers (Gill, Bruch, & Booker, 2013). Teacher effectiveness is then determined using the SLO data to determine student academic growth to measure the degree to which the goals have been achieved; thereby, concluding the degree of teacher effectivness rests on the ideas that high performing teachers equip students with the skils necessary to make larger learning gains.

SLOs are becoming more popular with states and districts looking to satisfy the requirement to include growth measures in teacher evaluation metrics. An additional advantage for using SLOs is that it encourages teacher engagement in the evalution process. The SLO process is indicative of a "particpatory method where teachers set measurable goals or objectives based on a teacher's particular students, subject, and grade, and of determining possible ways to measure growth in light of these" (Behrstock-Sherrat et al., 2013, p. 78).

At the center of each SLO are pre and post measures used to determine student learning toward the student learning objective. Gareis and Grant (2008) reported that,

"The attribute of validity is arguably the most important quality of an assessment" (p. 35). Validity refers to the ability of the assessment to measure what it is meant to assess. Therefore, a quality measure connects teacher, student, and course standards. SLO measures are more authentic than standaredized metrics as they are reflective of classroom assessments that are designed to give timely feedback about individual student content knowledge and skill attainment. Validity and relaiblity of SLO data is an element of the discussion relating to new evaluation programs for teachers.

Using SLOs can be problematic, as the objective must be set at a level providing for rigorous yet obtainable standards that encourage teachers to "reach maximmum potential with their students" (Behrstock-Sherrat et al., 2013, p. 78). While SLOs have the prospective to better discriminate teachers based on performance than traditional evaluation metrics, research has not looked at the reliabillty of the SLO data due to limited studies focused only on teachers achieving their SLO objectives (Community Training and Assistance Center, 2013; Gill et al., 2013; Goldhaber & Walch, 2011; Proctor, Walters, Reichardt, Goldhaber, & Walch, 2011). These and other issues complicate the use of student growth data as a basis for evaluating teachers.

Student growth and achievment is not haphazard, but occurs by design and strategic accountablity. Accountability is important for the growth and achievement of students in the United States (Raymond & Hanushek, 2003). Despite design flaws in most existing systems, Raymond and Hanushek (2003) found that holding schools accountable for student achievement has a positive effect on achievement. However, the affect holds true only for states attaching consequences to performance. States that simply provide information through report cards without attaching consequences to performance did not

get significantly larger affects than those with no accountability. The results were clear that a measurable connection exists between teacher effectiveness and student learning. Using student achievement information can provide a valuable tool for examining the classroom practices of teachers who improve student learning above expected levels of accomplishment. Student achievement, in turn, is an important source of feedback on the effectiveness of schools, administrators, and teachers.

Closely examining the effects of quality instruction on student growth is of paramount importance to the new educational reform policies. As teachers face high-stakes evaluations, policymakers must be cautious in creating evaluation tools that provide fair, reliable, and valid measures for examining multiple components for teaching standards. Evaluations that provide growth for teachers and improve the craft of teaching potentially can support teachers in enhancing student growth. Given the central role teachers have always played in successful schools, connecting teacher performance and student performance is a natural extension of the educational reform agenda.

Description of the JCSEE

In 1975, the JCSEE was created. The standards JCSEE created provided the education profession a common evaluation language, a conceptual framework, and guidelines to foster collaborative evaluation work (Reineke, Willeke, Walsh, & Sawin, 1988; J. B. Sanders, 1999; Stufflebeam, 2004). Later in 1988 (with revisions in 2009), the JCSEE developed personnel evaluation standards to guide the evaluations of education professionals, including teachers and principals. The JCSEE (2009) defined personnel evaluations "as the systematic assessment of a person's performance and/or qualifications in relation to a professional role and some specified and defensible institutional purpose"

(p. 3). The intent of the personnel evaluation standards is to address the concerns and practices leading to valid, fair, and useful evaluations of teachers and other educators. JCSEE used six assumptions as a guide in developing the personnel evaluation standards:

1. The primary use of evaluations is to provide effective services to students.
2. The evaluation practices must be free of needless threatening or demoralizing characteristics.
3. The use of the personnel evaluations must adhere to culturally competent practices.
4. Sound professional development and training experiences must result from the personnel evaluations.
5. Although disagreements may arise about what constitutes good teaching, good administration, or good research, these disagreements are necessary.
6. Evaluations will vary in complexity and importance.

Within the JCSEE framework are found four essential attributes of sound evaluation practice: propriety, utility, feasibility, and accuracy (PUFA). The JCSEE further supplemented the PUFA attributes by developing another set of standards around these four attributes as a deeper guide for the evaluation of educational programs, personnel, and students. These standards recognize pertinent concerns of propriety, utility, feasibility, and accuracy (JCSEE, 1988). This section articulates each of the PUFA attributes.

Propriety. Within the propriety attribute are seven standards for safeguarding that educational administrators provide legal and ethical evaluations, ensure not only the welfare of those being evaluated, but also the welfare of others involved in the

evaluation. As such, interpretation of propriety standards focuses primarily on matters of legality and form. While it may seem obvious that this requirement be included in the evaluation for teachers, it is important that all parties involved in the evaluation process have a clear and shared understanding of both the process and purpose of the evaluation (Stufflebeam & Sanders, 1990). Evaluators should use evaluation policies and practices that are consistent, equitable, and fair.

One way for evaluators to promote human dignity and professionalism is by providing a balanced evaluation that identifies both strengths and weaknesses. In a study conducted by Hill and Grossman (2013), three fourths of teachers surveyed "reported that their most recent evaluation failed to identify areas for improvement" (p. 373). The remaining teachers who reported that their evaluations did identify areas for improvement said they did not get any support for those improvements.

In light of current policies requiring multiple sources of input in the evaluation process, this standard provides for a formalization of those formative elements of teacher evaluation. All parties involved in the evaluation process can articulate strategies that allow teachers to individualize their professional development efforts within the scope of common protocols for documentation and assessment. These strategies may include the option for teachers to develop individual plans to describe the professional development activities for which they can be held accountable (Holland & Adams, 2002), and teacher portfolios that document teachers' work and learning, and also may include evidence of their students' performance (Holland, 2005). Using similar strategies can enhance the view of those involved in the evaluation of teaching from a bureaucratic procedure that is done to teachers, to a professional process where teachers as well as administrators can

work together to determine various ways to develop their practice, and participate in joint deliberation with supervisors in judging its worth (Holland, 2005).

Conflict of interest must be made transparent while addressed cooperatively and honestly for the evaluation to be of worth. Finally, evaluations must be legally defensible by adhering to all federal, state, and local laws. For stakeholders to trust in the evaluation system, the standard of propriety must be upheld. Table 5 provides a description of the propriety standards.

Table 5

Description of Propriety Standards

Propriety standard	Description of standard
Service orientation	Personnel evaluations should promote sound education, fulfillment of institutional missions, and effective performance of job responsibilities, so that the educational needs of students, community, and society are met.
Appropriate policies and procedures	Guidelines for personnel evaluations should be recorded and provided to the evaluatee in policy statements, negotiated agreements, and/or personnel evaluation manuals, so that evaluations are consistent, equitable, and fair.
Access to evaluation information	Access to evaluation information should be limited to persons with established legitimate permission to review and use the information, so that confidentiality is maintained and privacy protected.
Interactions with evaluatees	The evaluator should respect human dignity and act in a professional, considerate, and courteous manner, so that the evaluatee's self-esteem, motivation, professional reputations, performance, and attitude toward personnel evaluation are enhanced or, at least, not needlessly damaged.
Balanced evaluation	Personnel evaluations should provide information that identifies both strengths and weaknesses, so that strengths can be built upon and weaknesses addressed.
Conflict of interest	Existing and potential conflicts of interest should be identified and dealt with openly and honestly, so that they do not compromise the evaluation process and results.
Legal viability	Personnel evaluations should meet the requirements of all federal, state, and local laws, as well as case law, contracts, collective bargaining agreements, affirmative action policies, and local board policies and regulations or institutional statutes or bylaws, so that evaluators can successfully conduct fair, efficient, and responsible personnel evaluations.

Source. JCSEE (2014b, para. 1)

Utility. Personnel evaluations should be timely, informative, and influential. The objective is that effective evaluations support educators and administrators in their professional growth. Not only should evaluators identify, at the onset, (a) those who will use the evaluation system and (b) how stakeholders will use the evaluation result, but also possess the qualifications, skills, training, and authority to conduct personnel evaluations. As districts develop evaluation systems, attention should be directed toward actions that decrease the likelihood of confusions about performance expectations. Confusion can be avoided by clearly outlining and justifying the criteria used for personnel evaluation. Timely feedback from evaluators should be presented and documented following any evaluation undertakings. Personnel evaluation results are to be accurate, thereby providing educational professionals with identified areas for improvement in instructional practices and achieving the missions and goals of the organization (JCSEE, 2009, 2014b). Table 6 provides a description of the utility standards.

Feasibility. Feasibility demonstrates attributes of effectiveness and efficiency. The standard of feasibility provides a guide for education professionals to develop evaluation systems that are easy to implement, efficient, adequately funded, and politically viable. Increasing the feasibility can add value to an evaluation; therefore, designers of teacher evaluations should focus on the management of the logistical and administrative requirements of the evaluation process. Designers should also plan a program that is flexible, as the process and procedures of programs change from initial design to final product. Therefore, the process is iterative. What is considered feasible at the onset of the program may no longer be considered feasible at another juncture in the development of the evaluation (Yarbrough, Shulha, Hopson, & Caruthers, 2011).

Table 6

Description of Utility Standards

Utility standard	Description of standard
Constructive orientation	Personnel evaluations should be constructive, so that they not only help institutions develop human resources but encourage and assist those evaluated to provide excellent services in accordance with the institution's mission statements and goals.
Defined uses	Both the users and intended uses of a personnel evaluation should be identified at the beginning of the evaluation so that the evaluation can address appropriate questions and issues.
Evaluator qualifications	The evaluation system should be developed, implemented, and managed by persons with the necessary qualifications, skills, training, and authority, so that evaluation reports are properly conducted, respected, and used.
Explicit criteria	Evaluators should identify and justify the criteria used to interpret and judge evaluatee performance, so that the basis for interpretation and judgment provide a clear and defensible rationale for results.
Functional reporting	Reports should be clear, timely, accurate, and germane, so that they are of practical value to the evaluatee and other appropriate audiences.
Professional development	Personnel evaluations should inform users and evaluatees of areas in need of professional development, so that all educational personnel can better address the institution's missions and goals, fulfill their roles and responsibilities, and meet the needs of students.

Source. JCSEE (2014b, para. 2)

Evaluation procedures that are feasible will align with and not interrupt the normal program activities of the school. Personnel evaluations use procedures that do not interrupt the daily functioning of the organization and are responsive to cultural and background influences (JCSEE, 2014a; Yarbrough et al., 2011). Sartain et al. (2011) considered the implementation of a new teacher evaluation system that included a higher numbers of teacher observations. The participants in the study found that the feasibility of the program was deficient due to time constraints, resulting in little time for follow-up conversations with teachers, causing teachers to withdraw support for the evaluation.

Furthermore, developers must acknowledge how the contextual viability affects a program's feasibility. Two major factors influencing a program's contextual viability are

political interests and the values individuals and groups bring to the context for the evaluation (JCSEE, 2014a; McNeil, Hood, Kurtz, Thousand, & Nevin, 2006; Yarbrough et al., 2011). Evaluators can employ strategies to improve feasibility in context viability by not only examining and studying the political and cultural influences, but also by engaging in vigorous discourse with individuals and groups who influence the program (Stufflebeam & Shinkfield, 2007). Fiscal support must be present for any well-developed initiatives to be effective. Table 7 provides a description of the feasibility standards.

Table 7

Description of Feasibility Standards

Feasibility standard	Description of standard
Practical procedures	Personnel evaluation procedures should be practical, so that they produce the needed information in efficient, non-disruptive ways.
Political viability	Personnel evaluations should be planned and conducted with the anticipation of questions from evaluatees and others with a legitimate right to know, so that their questions can be addressed and their cooperation obtained.
Fiscal viability	Adequate time and resources should be provided for personnel evaluation activities, so that evaluation can be effectively implemented, the results fully communicated, and appropriate follow-up activities identified.

Source. JCSEE (2014b, para. 3)

Accuracy. Accuracy refers to the truthfulness of an evaluation; the truthfulness of representations, propositions, and findings that specifically support judgments. The goal of the accuracy standard is to increase the truthfulness of evaluation findings and conclusions; therefore, it is important to communicate how an evaluation creates accuracy in each program and evaluation context (JCSEE, 2014b; Miller, Linn, & Gronlund, 2009; Yarbrough et al., 2011). Table 8 provides a description of the accuracy standards.

Table 8

Description of Accuracy Standards

Accuracy standard	Description of standard
Validity orientation	The selection, development, and implementation of personnel evaluations should ensure that the interpretations made about the performance of the evaluatee are valid and not open to misinterpretation.
Expectations	The qualifications, role, and performance expectations of the evaluatee should be clearly defined, so that the evaluator can determine the evaluation data and information needed to ensure validity.
Analysis of context	Contextual variables that influence performance should be identified, described, and recorded, so that they can be considered when interpreting an evaluatee's performance.
Documented purposes and procedures	The evaluation purposes and procedures, both planned and actual, should be documented, so that they can be clearly explained and justified.
Information	The information collected for personnel evaluations should be defensible, so that the information can be reliably and validly interpreted.
Reliable information	Personnel evaluation procedures should be chosen or developed and implemented to assure reliability, so that the information obtained will provide consistent indications of the evaluatee's performance.
Systematic data control	The information collected, processed, and reported about evaluatees should be systematically reviewed, corrected as appropriate, and kept secure, so that accurate judgments about the evaluatee's performance can be made and appropriate levels of confidentiality maintained.
Bias identification and management	Personnel evaluations should be free of bias, so that interpretations of the evaluatee's qualifications or performance are valid.
Analysis of information	The information collected for personnel evaluations should be systematically and accurately analyzed, so that the purposes of the evaluation are effectively achieved.
Justified conclusions	The evaluative conclusions about the evaluatee's performance should be explicitly justified, so that evaluatees and others with a legitimate right to know can have confidence in them.
Metaevaluation	Personnel evaluation systems should be examined periodically using these and other appropriate standards, so that mistakes are prevented or detected and promptly corrected, and sound personnel evaluation practices are developed and maintained over time.

Source. JCSEE (2014b, para. 3)

To ensure accuracy, both the evaluator and the evaluatee should understand and adhere to the identified expectations for job performance. Without this shared understanding of the characteristics of effective job performance, "teachers [and other educational professionals] won't know how their performance will be evaluated and observers won't know what to look for" (Danielson, 2012, p. 34). Evaluators should clearly explain, justify, and document the evaluation purposes and procedures. Evaluators should also be able to defend the results of an evaluation. Inaccurate scores may cause administrators to focus professional development in the wrong direction, wasting time and resources on efforts that may possibly harm, rather than help, teachers and students (Hill & Grossman, 2013). Evaluators should collect, process, store, and analyze evaluation data in a systematic fashion to ensure results are defensible and lead to evidence-based judgments. Furthermore, school districts should periodically assess the personnel evaluation system to ensure sound evaluation practice (JCSEE, 2014b).

Significance of Teacher and Principals Perceptions

In response to the public demand for improved teaching and learning in public schools, policymakers focus on accountability measures for evaluating teacher effectiveness. Student achievement and academic progress have prompted education reformers to emphasis effective teaching by including student test scores in determining levels of teacher performance. Teacher evaluation is "a pressing issue in education and educational reform" (Pearlman & Tannenbaum, 2003, p. 633). Past and present teacher evaluation systems, while well intentioned, proved to be taxing and unsupportive (Danielson & McGreal, 2000). Even though research and practice over the past two decades has advanced in areas such as teacher effectiveness, school improvement, student

engagement in instruction, teacher and learner practices, teaching for understanding, and cognitive learning theory, evaluation systems for teacher have not (Danielson & McGreal, 2000). Understanding how teachers and principals perceive the coupling of student performance data and professional practices into the new evaluation systems will contribute to advancing teacher and principal buy-in, thus promoting implementation and sustaining fundamental reforms in schools (Bascia & Hargreaves, 2000; Turnbull, 2002). Research illuminates how teachers' attitudes about the evaluation process influences their perceptions toward the benefits derived from the evaluation process (Bransford & Donavan, 2005; Rogers, 1995; Tuytens & Devos, 2009).

According to Rogers's (1995) theory of perceived attributes, individuals are more likely to adopt an initiative when they perceive it as a positive. Rogers (2003) took into account the concept of relative advantage, defined as, "the degree to which an innovation is perceived as being better than the idea that it supersedes" (p. 212). Additionally, Bandura (2001) asserted that change is dependent on one's perceived belief about his or her ability to exercise control and make these changes. Bandura's social cognitive theory informs us that perceptions can develop as a function of feedback from the broader school social environment that is comprised of other teachers and school leaders.

While the teacher evaluation landscape of the past was presented with vague and subjective performance criteria, today the landscape requires multiple data points and detailed performance measures, along with measures of student growth. Principals in this age of reform and accountability encounter multiple, and often, simultaneous demands for greater involvement in (a) improving instructional practices, (b) observations that provide rigor and relevant feedback aligned with professional development opportunities,

(c) reallocating time and prioritizing commitments, and (d) improving student achievement while interfacing and building relationships with all stakeholders (Derrington & Campbell, 2013). Hall and Hord (2015) suggested that the extent to which an organization changes hinges on the changes occurring with each individual inside the organization. Moreover, Fullan (2005, 2008) contended that the school leader is a central part of the organization as well as part of the contextual flow of events influencing the perspectives both inside and outside schools. Principals provide a central role in the implementation of new teacher evaluation programs in promoting important variables toward successful change that include policy interpretation, capacity for implementation, adaptation, and management of the organization.

New reforms and changes in policies are contingent on school principals' ability and motivation to implement changes both in the school culture and procedures for the new evaluation programs (Fowler, 2009; Glickman, Gordon, & Ross-Gordon, 2010; Hall, 2013; Hallinger & Heck, 2011; Schmidt & Datnow, 2005). Therefore, districts desiring successful implementation of new evaluation programs and policies will benefit from understanding principals' perceptions and concerns with changes in teacher evaluation practices in order to provide principals with strategies to intervene or enhance change efforts (Derrington & Campbell, 2013; Glickman et al., 2010; Honig & Hatch, 2004; Leithwood, Strass, & Anderson, 2007).

Several researchers (Knight, 2008; Tuytens & Devos, 2009; Zimmerman & Deckert-Pelton, 2003) contended that teachers, those most affected by accountability policies in NCLB and RTT, have not been heard, nor been given the opportunity to contribute to the discussion regarding the construction and implementation of their

evaluation program. Perceptions of internal stakeholders relating to job maintenance and security in light of accountability policies in new evaluation programs can generate an environment of fear (Conley & Glasman, 2008). Conley and Glasman (2008) investigated how fear contributed to the perceptions of both teachers and administrators toward new evaluation programs. Conley and Glasman showed evidence for considering the thinking and feeling aspects of teachers and administrators and their link to improving evaluations and improving desired skills in teaching. For example, Conley and Glassman noted that principals' negative perceptions of the effect of new evaluation programs can create a fear relating to the "politics of maintenance" (p. 66). In response to their fear and perceptions, principals seek to "minimize loss associated with a lower quality of instruction, contributing directly to the overall performance of the school" (Conley & Glassman, 2008, p. 66). Accordingly, teachers' negative perceptions of the new evaluation program propel them toward self-preservation against losses from unfavorable evaluations, job security, and autonomy in applying their skills in teaching (Hackman & Oldham, 1980). Tuytens and Devos (2010) studied Dutch-speaking teachers in Belgium and found that teachers' perceptions of leadership variables influence the perceived utility of feedback and professional learning of teachers. Tuytens and Devos concluded that despite some research doubting the value of teacher evaluations in promoting improvement in teacher's instructional practices (Frase, 2001; Kleinhenz & Ingvarson, 2004); teachers in their study did in fact engage in professional development activities after receiving feedback from their evaluations.

Acheson and Gall (2003) found that teachers held a strong position that evaluations were potentially helpful. Acheson and Gall also noted that opposition to

evaluations was presented largely as a reaction to how the evaluation was implemented rather than the concept of an evaluation. Reeves (2004) found that criticism of the evaluation program stemmed from teachers' feelings of futility and disengagement from the process. Richardson and Placier (2001) found that teacher perceptions of teacher evaluations at the organizational level have been largely unexplored.

Researchers have found that teachers' perceptions varied according to the system used, the purposes of the evaluation, specific components within the system, teachers' own experiences, as well development and implementation. Policymakers for educational reform and school districts can use the new evaluation programs to either promote or obstruct teacher development. Conley, Muncey, and You (2005); Milanowski and Heneman (2001); and Pizzi (2009) found that teachers want to know what standards or indicators they will be evaluated against and how the evaluation will be conducted. For an evaluation system to be well crafted, school districts must provide a shared understanding of the process, expectations, and goals for the evaluation program (Stronge, 2006). Teachers desire more transparent evaluation programs (Feeney, 2007) informing them specifically of how they will be evaluated (Pizzi, 2009). Legally defensible evaluation programs must provide teachers both procedural and substantive due process (JCSEE, 2009).

JCSEE's framework is based on four standards: propriety, utility, feasibility, and accuracy. Research over the last two decades describes perceptions of educators toward the JCSEE standards as related to the changing evaluation programs in this era of accountability. The first component of the propriety standard recognizes that evaluations should reflect the goal of sound education for all students (JCSEE, 2009). However,

Engram (2007) and Marks (2005) found that most teachers did not believe that their teacher evaluation systems would increase student achievement or that student achievement data would produce improvement in teacher effectiveness. Educational leaders have a responsibility to reflect on the evaluation results, attend to the trajectory of trends in student achievement, and adjust the curriculum, instruction, and assessment as necessary (Marzano, Waters, & McNulty, 2005).

Among the components of the utility standard is constructive orientation. Institutions should develop resources to assist evaluatees in their performance as it is related to the school's mission, goals, and objectives (JCSEE, 2009). Researchers have noted that teachers agree that an effective teacher evaluation system must relate directly to the organization's mission, goals, and objectives (Castillo, 2005; Stronge & Tucker, 1999; Tuytens & Devos, 2009).

Educational leaders creating evaluation programs for teachers would benefit from adhering to the feasibility standard to promote ease in implementing efficient, adequately funded, and politically viable evaluations. Evaluation systems must run smoothly within an organization (JCSEE, 2014b). Pizzi (2009), Zimmerman and Deckert-Pelton (2003), and Sartain et al. (2011) found that both teachers and principals reported inadequate time spent during the evaluation process, including too little time providing feedback to teachers. Tuytens and Devos (2009) found that when teachers perceive that elements of the evaluation program are ambiguous and unfeasible, developers should provide clarity regarding the usefulness and value of the evaluation policy. The political viability component of the feasibility standard asserts that evaluations should engage and be responsive to stakeholders. Marzano et al. (2005) asserted that professional relationships

enhance the application of leadership responsibilities. Marzano et al. noted that educational leaders should demonstrate an awareness of the personal lives of teachers and staff as well as nurture the teachers' empowerment.

JCSEE (2009) calls for evaluations to provide sound data. However, teachers perceive the results of evaluation to be neither valid nor reliable, claiming a lack of training and bias on the part of the evaluator (Castillo, 2005; Flores, 2012; Hopkins, 2013; Marks, 2005; Pizzi, 2009; Wacha, 2013; Zimmerman & Deckert-Pelton, 2003). Moreover, Hopkins (2013) found that teachers did not credit the use of student growth data for providing increases with evaluator objectivity or decreases in evaluator bias.

Perceptions of Teachers and Principals

Educational reform polices relating to new evaluation programs for teachers induce change. These changes elicit concerns for individuals affected by the new process. Hall (2013) explained that these concerns manifest in people as feelings, thoughts, reactions, and perceptions to change affecting their lives therefore; change elicits concern. Hall and Hord (2015) provided various stages of the affective or personal feeling side of change. The initial consideration of individuals in the primary stage of change is the effect the change has on themselves. Individuals may experience self-doubt in their capacity to manage the change. The second stage of concern manifests after several years as individuals experience success as the change becomes familiar and individuals become more proficient with the change. The third stage of concerns is demonstrated as individuals' focus turns from learning to be competent with the change process to investigating the benefits and satisfaction, and then, finally mastering the change process.

Successful change requires that those involved in the process respond positively to the change; therefore, educational reformers need to consider the perceptions of teachers, who are often a neglected component for implementing new policies. Nias (1999) found that (a) emotional and perceptual reactions of teachers are deeply rooted to the view they have of themselves and of others, (b) teachers' thoughtful actions reflect emotional involvement and moral judgment, and (c) neither perceptions nor feelings can be separated from the cultural and social influences that form and shape them.

It is important to understand the effect that teacher's affectivity plays in light of changes in educational policy. Those seeking to change educational policy need to consider the teachers' responses to and perceptions of change, as well as perceptions of the implementation process (Hargreaves, 2004; Schmidt & Datnow, 2005; Tuytens & Devos, 2009). Changes in school policies cause teachers to respond emotionally to the potential affects these changes may bring. These emotional responses influence how teachers perceive, interpret, and evaluate the changing environment (Troman & Woods, 2001).

A natural inclination toward changes in policies is either support or resistance. Van Veen, Sleegers, and Van de Ven (2005) and Lasky (2005) found that while some teachers were pleased to support and sustain educational reform, others experienced anxiety, fear, defeat, or frustration and, thereby, resisted reform efforts. Understanding how teachers make sense of change is often understudied and overlooked (Schmidt & Datnow, 2005; Spillane et al., 2002). Educational researchers acknowledge the importance of teachers' emotions in the change process and school climate (Hargreaves, 2004; Hargreaves & Dawe, 1990; Nias, 1999; Sergiovanni, 1992; van Veen & Lasky,

2005). Within the landscape of the current reform policies of school accountability, tensions among teachers are increasing. Conley and Glasman (2008) indicated that,

> Individual teachers fearing a summative evaluation may be less than forthcoming about their performance shortcomings and/or goals, and supervisors may hesitate to give teachers detailed feedback. The result is that teachers may fear that evaluation is less about personal improvement involving professional growth and more of a political hurdle. (p. 68)

Understanding teachers' emotional geography as well as their perceptions of change can assist policymakers in navigating educational reform. While resistance to change may be seen as an impediment to school reform, Zembylas and Barker (2007) acknowledged that resistance toward change is "part of the process, in fact, it has a modifying influence and that ambivalence and confusion that teachers have toward change can be understood on the basis of how individuals respond to change and why they change" (p. 240).

Understanding how teachers make meaning of change provides crucial insight for implementing school reform. Zembylas and Barker (2007) conducted a 2-year ethnographic study of 14 elementary teachers involved in the pilot of a new science literacy program. Using a grounded theory approach, Zembylas and Barker examined three aspects of teachers coping with change in schools: (a) time and space as sources of emotional and support, (b) teacher collegiality and trust, and (c) teachers' moral values and concerns. Their findings suggested that teachers created spaces for coping with change efforts by adapting the reforms in ways that were consistent with their values, while limiting their own frustrations in the proposed change. Zembylas and Barker

(2007) also noted that two distinctive components of teachers' collegiality and trust persuaded reform efforts. Teacher collegiality based on friendship and trust may prove to be powerful in determining successful reform efforts. Alternatively, the researchers found that teacher collegiality based on politeness and avoidance of conflict may subvert reform efforts, because the real issues relating to change are not addressed.

Finally, Zembylas and Barker (2007) found that working conditions, social relations, and moral/personal values and concerns involve teachers emotionally with respect to what is at stake for them apart from whether they support or refute the school reform initiative. These findings are comparable to other studies by Hargreaves (2001) and Nias (1999), suggesting that the personal, social, and emotional aspects of change have wide-ranging effects on classroom practices and reform efforts.

A major implication of previous research is the necessity for school reformers to acknowledge opportunities for dealing with the emerging perceptions and feelings of teachers toward school change. Acknowledging teachers' perceptions is not about assuaging their feelings toward reform, but rather using the feelings as a valuable vehicle in finding ways to integrate and/or reconcile opposing feelings about the change efforts (Sarason, 1996; Schmidt & Datnow, 2005; Zembylas & Barker, 2007). Additionally, teachers need the emotional and social supports to understand and reasonable cope with deeply embedded perceptions or feelings of conflict, tension, and disruption that are part of the processes of school change. Change is not about forcing teachers to conform to the new policies, but about allowing teachers to individually and collaboratively reflect, build trust, share visions, promote openness with risk taking, and make sense of the change so they can adopt the changes into their professional practices (Price, 2012: Sarason, 1996).

In a related study, Schmidt and Datnow (2005) attempted to link research on emotions and sense making. This qualitative study explored teachers who were involved in implementing a comprehensive school reform model. The framework for their study was built on Blumer's (1969) concept of symbolic interaction. This concept asserts that individuals act toward things based on the meanings they have constructed for them. Schmidt and Datnow stated that, "The meaning of events arises from social interactions leading to unforeseen and often unpredicted emotions that can frustrate or enhance policy implementation" (p. 950). Individuals use an interpretive process to modify meaning of events. Mehan (2000) added that sense making or meanings can be contested or affected by power relationships in a given interaction. Therefore, emotions, sense making, and perceptions that are created, play an important role in teachers' behavior toward school reform policy. Making sense of the reforms is "emotionally laden as teachers sort through feelings of anxiety and the unknown, frustration of the ambiguous, joy, and recognition of shared ideologies (i.e., reform and self), and guilt in constructing modifications despite possible professional repercussions" (Schmidt & Datnow, 2005, p. 960). Successful implementation of reform depends to a great degree on how the teachers perceive the value and worth of the reform. The teachers' responses can either enhance or sabotage the efficacious implementation of reform initiatives.

Summary

Front and center in today's educational reform movement is the debate on school and teacher accountability. This debate has evolved over many decades and has resulted in federal legislation such as NCLB and RTT. NCLB signaled a substantial drive toward accountability for student achievement. NCLB outlined accountability measures such as

established standards in state mandated curriculum as well as, requiring districts to use standards-based evaluations to ensure a quality education for all students. More recently, legislation such as RTT provides states with federal funding through competitive grants requiring new evaluation systems for teachers. These new evaluation systems for teachers require multiple measures of evidence, including student growth, in documenting teacher effectiveness.

Teacher effectiveness and its effect on students is a key component of the policy and legislation. Researchers agree there is a direct relationship between teacher effectiveness and student learning. In this educational and political climate of school accountability, policymakers have determined that a new system of high-stakes teacher evaluations serves as the vehicle to ensure teachers' responsibility for student growth. While developing an effective evaluation system for measuring teacher effectiveness is a complex issue, the rationale behind the evaluation process is less complex. DiPaola and Hoy (2012) suggested, "There is common agreement that the overall purposes of personnel evaluation are accountability and professional growth leading to student achievement" (p. 147). For evaluations to be effective JCSEE has developed four standards commonly known as PUFA. These standards provide that evaluations will address standards of propriety, utility, feasibility, and accuracy. States and districts too often neglect the voice of the teacher, who is at the center of the evaluation process. Understanding the effect of school reform through the perceptions of teachers is beneficial for both successful implementation and sustained changed.

CHAPTER 3: METHODS

The purpose of this quantitative study was to assess the degree to which teachers and evaluators concur or differ in their perceptions of whether the county's new teacher evaluation program meets the propriety, utility, feasibility, and accuracy attributes of the JCSEE. This chapter contains details of the methodology used to collect data to answer three research questions. The online survey is described and the procedures used in the study are outlined.

Participants

Approximately 1,562 teachers, all members of Cohort 1, as well as 41 principals from elementary, middle, and high schools, were asked to participate in the study. Even though all teachers in Emerald County School District receive an annual evaluation, the student growth measures are only included annually in the new evaluation program. Of the two components used in the new evaluation program for teachers, the professional practice (qualitative) component provides greater challenges to both implement and maintain (Danielson, 2007). While the student growth measure (quantitative) component is required to be included in the evaluation annually, Emerald County School District has determined that the professional practice component of the evaluation be performed on a 3-year cycle for tenured teachers who are rated effective. Emerald County School District has determined a 3-year evaluation cycle where both the student growth measures and professional practices are included in the evaluation. On the two alternate years, when the professional practices component is not included, either previous professional practices

scores or previous satisfactory scores from the former evaluation are included in the new evaluation program for teachers; thus, allowing for a continuous, rolling evaluation plan.

Emerald County School District established Cohort 1 during the 2013–2014 school year. The cohort is comprised of (a) tenured teachers who were scheduled for evaluation during the 2013–2014 school year using the former evaluation program, (b) any non-tenured teachers or teachers previously rated ineffective under the former evaluation program, and (c) one third of remaining tenured teachers. Cohort 1 consists of teachers across the K–12 spectrum who (a) teach both content and elective subjects; (b) teach courses that are tested by standardized national, state, and local measures; or (c) teach in a general, inclusion, or self-contained setting, including students with and without an IEP or 504 plan. Included in this cohort are teachers who teach in Title I schools, schools with regional programs, and non-Title I schools.

Teachers in Cohort 1 are at schools where the student populations include both Black or Caucasian students in the majority. No schools in Emerald County School District have an ethnic majority of Latino or Asian. Tenured teachers in Cohort 1 hold advance professional to professional eligibility certifications with undergraduate to graduate degrees in education and noneducation majors. Nontenured teachers hold graduate to undergraduate degrees as well as provisional to advanced professional certificates.

Five hundred educators completed the questionnaire and were included in the analysis. Of the 470 teachers included in the dataset, 76% were female, 85% were Caucasian, and 67% held masters' degrees. Of the 30 principals included in the dataset, 70% were female, 67% were Caucasian, and 80% held masters' degrees.

Data Sources

This study used survey research, a method of collecting information by asking questions typically on questionnaires. The broad area of survey research encompasses any measurement procedures that involve asking questions of respondents. The data collection instrument was an online survey modified from research surveys conducted by Hopkins (2013) and Stronge (2013). Both researchers granted permission for modifying their surveys for the current study (Appendix A). The online questionnaire (Appendix B) consists of 27 items and are answered using a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). The 27 items were chosen to measure the teachers and administrators' perceptions of the propriety, utility, feasibility, and accuracy of the new evaluation system in the Emerald County School District. Table 9 contains each JCSEE attribute and the items on the questionnaire that measured it.

A demographic section asked respondents to provide their gender, ethnicity, age, years of experience, and grade level (pre-K through elementary, middle, or high). This information was used to describe the sample. The correspondence of items on the study's questionnaire to those from Hopkins and Stronge's surveys and to the JCSEE standards is presented in Appendix C and Appendix D.

Data Collection

The members of Cohort 1 received an email through Emerald County School District's school server inviting them to participate in the survey. The email contained a link to the online questionnaire. The first page of the questionnaire was a consent form detailing the purpose of the study, the confidentiality of their responses, and the risks and benefits of the study (Appendix E). If the teachers clicked *YES* at the bottom of the

Table 9

JCSEE Attributes and Items on Questionnaire

Attribute	Description of attribute	Item
Propriety	Safeguard that educational administrators provide legal and ethical evaluations; ensure not only the welfare of those being evaluated, but also the welfare of others involved in the evaluation	1–6
Utility	Personnel evaluations should be found useful in that they are timely, informative, and influential. The objective is that effective evaluations support educators and administrators in their professional growth.	7–14
Feasibility	Feasibility demonstrates attributes of sufficient effectiveness and efficiency. The standard of feasibility provides a guide for education professionals to develop evaluation systems that are easy to implement, efficient, adequately funded, and politically viable.	15–19
Accuracy	Accuracy refers to the truthfulness of an evaluation; the truthfulness of representations, propositions, and findings that specifically support judgments. The goal of the accuracy standard is to increase the truthfulness of evaluation findings and conclusions.	20–27

consent form, they provided their implicit consent to participate in the survey, and they were directed to the first page of the questionnaire. If they choose to click *NO*, they were logged out of the survey. Three reminder emails were sent at intervals of 4–5 days. The educators' responses were stored on the online server and were downloaded at the end of the survey process. The data collection process took approximately 2 weeks.

Data Analysis

The items measuring each JCSEE subscale were averaged to produce a scale score for propriety, utility, feasibility, and accuracy. The scale scores ranged from 1 to 4, where a high score represented positive perceptions of the new evaluation system. Reliability of the four scales was measured using Cronbach's alpha coefficient.
The sample sizes of the two groups of participants (teachers and administrators) were disproportionate. The administrators' responses were analyzed separately to determine the means of their responses to each scale. Their responses were used as specific values

in four one-sample *t* tests. The null hypothesis of each one-sample *t* test was that the mean of the teachers was equal to the specific mean of the administrators. The one-sample *t* tests were run to determine if the perceptions of the teachers and school administrators were different. Analysis of variance was used to determine if differences among teachers and principals at each grade level were different. Table 10 contains research questions, data source, and analyses used to answer the questions.

Table 10

Data Analysis Plan

Question	Source of data	test
1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date?	Questionnaire items 1–27	Descriptive statistics
2. Are there differences among teachers at different levels (middle, elementary, and high) in their perceptions of the Emerald County School District evaluation system as implemented to date?	Questionnaire items 1–27	One-sample *t* tests ANOVA
3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the evaluation system as implemented to date?	Questionnaire items 1–27	One-sample *t* tests

Limitations

Limitations of a study are characteristics of the design that may affect the interpretation of the data collected in the study. Several limitations were present in the current study. First, it is unknown how representative the sample was to the population of teachers surveyed. Second, the instrument designed for this study is comprised of items adapted from two other surveys. They were chosen to measure the four JSCEE attributes of good evaluation practice. The adequacy or inadequacy of the items in the current study to measure these attributes is unknown. Finally, the data collected were self-reported. The

teachers and principals' responses could not be independently verified. Therefore, the researcher had to rely on respondents' truthfulness when responding to the questionnaire.

Ethical Considerations

As a researcher, I adhered to the JCSEE (2009) standards of program evaluation. I made myself available to the potential participants, allowing for open and responsive communication during the research process. Additionally, the school district's application and approval process provided a formal agreement between the evaluator and stakeholders that took "into account the context, needs, and expectations of clients and other parties" (Mertens & Wilson, 2012, p. 25). The evaluation program plan, methodology, and data collection instruments were presented for review and approval by the assigned dissertation committee, The College of William and Mary's School of Education Internal Review Committee, and the school district's department of research and assessment, ensuring a full measure of protection to participants. If unforeseen conflicts of interest arose that may have compromised the process and results of the study, participants and/or program leaders were directed to the study's faculty advisor. If participants had ethical concerns with the conduct of this study, they were directed to contact the chair of the Protection of Human Subjects Committee at The College of William and Mary (Appendix E).

My educational background and my professional experience speak to my qualifications for conducting an effective and credible evaluation. I have an understanding of the evaluation process from the perspective of both a teacher and administrator. That dual understanding may help create a trusting environment for

distributing the surveys for the purpose of data collection, and for communicating the practical use of results (Mertens & Wilson, 2012).

Stakeholders from the school district were included throughout the study. The plan promoted teacher ownership of findings. The data collection activities may have encouraged sharing of individual judgments as well as collegial conversation following participation. In this regard, the evaluation may foster an ongoing dialog that may benefit teachers and students for years to come.

This evaluation used procedures and resources familiar to participants. Teachers in the district use computers daily; therefore, the web-based format of the survey was a convenient delivery format. The familiar format of the survey should have given teachers an outlet for expressing their perceptions about the new teacher evaluation process.

I am committed to clearly documenting "findings, interpretations, conclusions, and judgments…without omissions or flaws" (Mertens & Wilson, 2012, p. 26). Multiple communication approaches strengthen valid reporting practices. Full disclosure of findings and reports to the school district also provided a safeguard against invalid reporting and communication. Protection of participants from potential harm came through adherence to prescribed federal, university, and school district requirements for research. The research study was reviewed and approved by The College of William and Mary's School of Education Internal Review Committee per the requirements established by the university and according to regulations of the U.S. Department of Health and Human Services before proceeding with the study. Moreover, the research was submitted to the school district's department of research and assessment for review and approval before proceeding with the study.

Summary

Almost 500 tenured teachers and 30 school principals participated in an online survey to assess the degree to which teachers and evaluators concurred that the new teacher evaluation program met the propriety, utility, feasibility, and accuracy attributes of the JCSEE. Prospective participants received an email through Emerald County School District's school server inviting them to participate in the survey. Reminder emails were sent over a 2-week period. Descriptive statistics were used to determine the perceptions of the teachers and administrators regarding the evaluation system as implemented to date. Additionally, one-sample *t* tests and analysis of variance were used to determine differences between and among the teachers and administrators.

CHAPTER 4: RESULTS

The purpose of this quantitative study was to assess the degree to which teachers and administrators concurred that the new teacher evaluation program met the propriety, utility, feasibility, and accuracy attributes of the JCSEE. The data were used to answer three research questions. This chapter contains the results of the analyses of those data.

Response Rate

Email invitations were sent to 1,562 teachers. Almost 700 (n = 679) teachers opened the link to the survey; however, 488 continued to the end of the questionnaire, creating a final response rate of 31.2%. Eighteen of these respondents did not indicate in which grade level they taught. Therefore, a sample of 470 was used to answer the research questions. Email invitations were also sent to 41 administrators. Thirty-four opened the link to the survey; however, 30 continued to the end of the questionnaire, creating a final response rate of 73.2%. The data collected from these 470 teachers and 30 administrators were used to answer the research questions. The data include responses from 206 elementary, 111 middle, and 153 high school teachers. Sixteen elementary, 11 middle, and 3 high school administrators were included in the administrator dataset.

Description of the Sample

Tables 11 and 12 contain a description of the teachers and administrators in the two datasets. Female teachers and administrators were a majority at every grade level, with more at the elementary level (90% teachers and 81% administrators) than at the middle school level. More than 80% of the teachers and two thirds of the administrators

Table 11

*Description of the Sample**

	Teachers								Administrators							
	Elementary *n* = 206		Middle *n* = 111		High *n* = 153		Total *n* = 470		Elementary *n* = 16		Middle *n* = 11		High *n* = 3		Total *n* = 30	
Characteristic	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Gender																
Female	183	90.1	77	71.3	89	59.3	349	75.7	13	81.3	6	54.5	2	66.7	21	70.0
Male	20	9.9	31	28.7	61	40.7	112	24.3	3	18.8	5	45.5	1	33.3	9	30.0
Race																
African American	17	8.6	17	16.3	20	13.8	54	12.1	6	37.5	2	18.2	2	66.7	10	33.3
Hispanic	1	0.5	4	3.8	2	1.4	7	1.6	0	0.0	0	0.0	0	0.0	0	0.0
Caucasian	176	88.9	83	79.8	120	82.8	379	84.8	10	62.5	9	81.8	1	33.3	20	66.7
Other	4	2.0	0	0.0	3	2.1	7	1.6	0	0.0	0	0.0	0	0.0	0	0.0
Degree																
Bachelor's	61	29.6	29	26.1	47	31.5	137	29.4	1	6.3	1	9.1	0	0.0	2	6.7
Master's	137	66.5	78	70.3	96	64.4	311	66.7	13	81.3	8	72.7	3	100.0	24	80.0
Specialist	7	3.4	3	2.7	2	1.3	12	2.6	2	12.5	2	18.2	0	0.0	4	13.3
Doctorate	1	0.5	1	0.9	4	2.7	6	1.3	0	0.0	0	0.0	0	0.0	0	0.0

* Some respondents did not provide demographic information. Therefore, the categories may not total the number for each group of respondents.

Table 12

Age and Years of Experience of Teachers and Administrators in Study Sample

	Teachers								Administrators							
	Elementary *n* = 206		Middle *n* = 111		High *n* = 153		Total *n* = 470		Elementary *n* = 16		Middle *n* = 11		High *n* = 3		Total *n* = 30	
Characteristic	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	39.24	11.49	37.45	10.05	39.45	11.58	38.89	11.20	47.75	7.59	46.20	5.25	51.00	7.07	47.43	6.68
Years of experience	13.10	10.28	11.05	7.86	11.96	9.48	12.24	9.51	25.06	7.97	20.70	5.83	27.00	7.00	23.76	7.34

were White and the majority of teachers and administrators held master's degrees. The average age of the teachers was less than 40 (M = 38.9, SD = 11.2), while the average age of administrators was 47. Administrators also had more years of educational experience (M = 23.8, SD = 7.3) than did teachers (M = 12.2, SD = 9.5).

Reliability of the Scales in Study

The reliability of the scales used to measure the participating teachers and administrators' perceptions of the propriety, utility, feasibility, and accuracy attributes of the new evaluation system in the district were measured using Cronbach's alpha coefficient (Table 13). The alpha coefficients in this study for each of the subscales ranged from .70 to .91. With all scales at or above .70, the values indicated an acceptable reliability (Nunnally, 1978).

Table 13

Reliability of the Scales Measuring Propriety, Utility, Feasibility, and Accuracy Attributes

		Cronbach's alpha coefficient	
Scale	# of items	Teachers	Administrators
Propriety	6	.86	.70
Utility	8	.91	.81
Feasibility	5	.85	.77
Accuracy	8	.91	.85

Analysis of the Research Questions

Three research questions guided this study. This section contains the results of the analyses used to answer those research questions. In each instance, the research question is presented, the data are described in tables, and the results of the analyses are presented. The sample sizes of the groups of respondents (teachers and administrators) were

disproportionate. The administrators' responses (n = 30) at each school level were analyzed separately to determine the means of their responses to each scale.

Research Question 1

What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date?

Means were calculated across the items in each of the four scales measuring the propriety, utility, feasibility, and accuracy of the district's teacher evaluation system. The responses to the items were measured on a scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*), with a mean of 2.5 the midpoint at which agreement of a group shifts from tending to disagree to tending to agree. Table 14 contains the means and standard deviations of the teachers and administrators' ratings of the JCSEE attributes. The scale items measuring propriety garnered the highest levels of agreement for both teachers (M = 2.28) and administrators (M = 2.93). The scale measuring accuracy garnered the lowest levels of agreement for both teachers (M = 2.21) and administrators (M = 2.62).

The propriety standard was the only standard that was above the midpoint of 2.5 for teachers, meaning that teachers tended to agree more than disagree. The means for the utility, feasibility, and accuracy all fell below the midpoint for the sample of teachers, but all were above 2.0, the anchor indicating *disagree.* For the administrators, all of the means were above the midpoint of 2.5, meaning that as a group they tended to agree more than disagree, but none of the means were above 3.0, the anchor that signified *agree.* The standard deviations for teachers were wider than for administrators, with a range of .60 to .66 for teachers and .37 to .47 for administrators. This indicated a wider diversity of perspectives among the teachers, although this was also a much larger sample.

Table 14

Means and Standard Deviations of Teachers and Administrators' Ratings of the JCSEE Attributes

	Teachers (n = 470)		Administrators (n = 30)	
	M	*SD*	*M*	*SD*
Propriety	2.58	.60	2.93	.41
Utility	2.37	.64	2.71	.37
Feasibility	2.35	.63	2.81	.38
Accuracy	2.21	.66	2.62	.47

Tables 15–18 contain the percentages of teachers and administrators who responded to the Likert scale for each item in the questionnaire. Each table contains the items for one of the JCSEE attributes. In all these analyses, administrators were more likely to agree with each item than were teachers.

Table 15 contains the teachers and administrators' responses to items measuring propriety. Almost all of the administrators (96%) agreed that there was a clear understanding of the expectations of the teacher's job performance, however, only three quarters (76%) of the teachers agreed that these expectations were clear. Nearly two thirds of the teachers (62%) and administrators (63%) agreed that the county provides clear and concise documentation of the procedures and guidelines outlining the policies and procedures of the evaluation system.

Half of the teachers (52%) and 70% of the administrators agreed that using student learning objectives data encourages professional discussion during follow-up conferences. Two thirds of the teachers (66%) and nearly all of the administrators (93%) agreed that using professional practices data encourages professional discussion during

follow-up conferences. In a point of divergence, almost two thirds of the teachers (64%) disagreed that using student learning objectives data documents teachers' areas of strength, while 60% of the administrators agreed that it did. Teachers (64%) and administrators (87%) agreed, however, that using professional practices data documents teachers' areas of strength.

Table 15

Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Propriety

		Percentage of respondents†					
#	Item	SD*	D	A	SA	*M*	*SD*
1	Clear understanding of expectations						
	Teachers	3	20	59	17	2.90	.71
	Administrators	0	7	46	50	3.43	.63
2	County provides clear and concise documentation						
	Teachers	7	32	50	12	2.67	.77
	Administrators	0	37	43	20	2.83	.75
3	Using student learning objectives data encourages professional discussion						
	Teachers	15	34	45	7	2.44	.82
	Administrators	7	23	57	13	2.77	.77
4	Using professional practices data encourages professional discussion						
	Teachers	9	26	56	10	2.67	.77
	Administrators	0	7	73	20	3.13	.51
5	Using student learning objectives data documents teachers' areas of strengths						
	Teachers	25	39	30	6	2.17	.87
	Administrators	13	27	53	7	2.53	.82
6	Using professional practices data documents teachers' areas of strengths						
	Teachers	9	26	54	10	2.66	.79
	Administrators	3	10	80	7	2.90	.55

† Percentages may not total 100 due to rounding.
* SD = *strongly disagree*, D = *disagree*, A = *agree*, SA = *strongly agree*

Table 16 contains the teachers and administrators' responses to items measuring utility. Fewer teachers (36%) and administrators (60%) agreed using student learning objectives data improves teaching than those who agreed that using professional practices

Table 16

Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Utility

#	Item	Percentage of respondents†				*M*	*SD*
		SD*	D	A	SA		
7	Using student learning objectives data improves teaching						
	Teachers	25	39	30	6	2.17	.88
	Administrators	7	33	57	3	2.57	.68
8	Using professional practices data improves teaching						
	Teachers	12	32	48	8	2.52	.81
	Administrators	0	20	67	13	2.93	.59
9	Administrators are qualified to evaluate student learning objectives data						
	Teachers	13	34	47	6	2.46	.79
	Administrators	0	17	73	10	2.93	.52
10	Administrators are qualified to evaluate professional practices data						
	Teachers	9	26	57	9	2.66	.76
	Administrators	0	0	83	17	3.17	.38
11	Criteria for using student learning objectives is clear and accurate						
	Teachers	27	40	30	4	2.11	.84
	Administrators	13	57	23	7	2.23	.77
12	Criteria for using professional practices is clear and accurate						
	Teachers	14	39	42	6	2.38	.80
	Administrators	0	47	47	7	2.60	.62
13	Using student learning objectives data informs professional development activities						
	Teachers	25	38	32	5	2.17	.87
	Administrators	17	17	53	3	2.43	.82
14	Using professional practices data informs professional development activities						
	Teachers	13	29	50	7	2.51	.81
	Administrators	0	27	63	10	2.83	.59

† Percentages may not total 100 due to rounding.
* SD = *strongly disagree*, D = *disagree*, A = *agree*, SA = *strongly agree*

data improves teaching (56% and 80%, respectively). Teachers and administrators varied in their agreement as to whether administrators are qualified to evaluate the student learning objectives and professional practices components of the system. Slightly more than half (53%) of the teachers, but 83% of the administrators agreed that administrators are qualified to use the student learning objectives component. All of the administrators

(100%) and two thirds of the teachers (66%) agreed administrators are qualified to use the professional practices component of the system.

Over two thirds of both the teachers (67%) and administrators (70%) disagreed that the criteria for using student learning objectives in rating teacher performance is clear and accurate, while approximately half of the teachers (48%) and administrators (54%) agreed that the criteria for using professional practices data is clear and accurate. Again, fewer teachers (37%) than administrators (56%) agreed that using SLO data informs professional development activities. However, more teachers (57%) and administrators (73%) agreed that professional practices data does inform professional development activities.

Table 17 contains the teachers and administrators' responses to items measuring the feasibility standard. Almost three fourths of administrators agreed that using student learning objectives data (70%) is a responsible use of assessment data and provides two-way communication between administrators and teachers (72%). However, fewer than half of teachers agreed that student learning objectives data (36%) is a responsible use of assessment data or that it provides two-way communication between administrators and teachers (47%). Both teachers (71%) and administrators (100%) agreed that using professional practices data provides opportunity for two-way communication between administrators and teachers. Again, teachers (51%) and administrators (80%) are more in agreement that the teacher time required for employing professional practices data in the new evaluation system is feasible than using the student learning objectives data (35% and 60%, respectively).

Table 17

Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Feasibility

#	Item	Percentage of respondents†				*M*	*SD*
		SD*	D	A	SA		
15	Using student learning objectives data is a responsible use of assessment data						
	Teachers	22	42	32	4	2.19	.82
	Administrators	13	17	63	7	2.63	.81
16	Using student learning objectives data provides communication between administrators and teachers						
	Teachers	17	35	42	5	2.36	.83
	Administrators	3	24	62	10	2.79	.68
17	Using professional practices data provides communication between administrators and teachers						
	Teachers	9	21	63	8	2.70	.73
	Administrators	0	0	76	24	3.24	.44
18	Teacher time required for employing student learning objectives data is feasible						
	Teachers	23	42	32	3	2.15	.81
	Administrators	3	36	60	0	2.57	.57
19	Teacher time required for employing professional practices data is feasible						
	Teachers	19	31	47	4	2.36	.83
	Administrators	3	16	73	7	2.83	.59

† Percentages may not total 100 due to rounding.
* SD = *strongly disagree*, D = *disagree*, A = *agree*, SA = *strongly agree*

Table 18 contains the teachers and administrators' responses to items measuring accuracy. The items measuring accuracy contained statements about how teachers and administrators agree that the use of student learning objectives and professional practices accurately contributes to evaluating teaching, making the evaluations more objective, and helping administrators identify low-performing/ineffective teachers. In each case, teachers were in less agreement than were administrators. However, both teachers and administrators were in more agreement about the use of professional practices data than the use of student learning objectives. Over three quarters of teachers (77%) and a majority administrators (53%) did not agree that using student learning objectives data accurately contributes to evaluating teaching, while more teachers (51%) and

Table 18

Teachers (n = 470) and Administrators' (n = 30) Ratings of Items Measuring Accuracy

#	Item	Percentage of respondents†				*M*	*SD*
		SD*	D	A	SA		
20	Using student learning objectives data evaluates my teaching						
	Teachers	37	40	20	7	1.88	.81
	Administrators	20	33	40	7	2.33	.88
21	Using professional practices data evaluates my teaching						
	Teachers	21	28	45	6	2.38	.88
	Administrators	3	13	70	13	2.93	.64
22	Using student learning objectives data in will make my evaluation more objective						
	Teachers	30	36	30	7	2.07	.86
	Administrators	10	33	50	7	2.53	.78
23	Using professional practices data in will make my evaluation more objective						
	Teachers	21	29	46	5	2.35	.86
	Administrators	3	20	63	13	2.87	.68
24	Using student learning objectives data directs attention to achievement gaps in classrooms						
	Teachers	18	31	47	4	2.37	.83
	Administrators	7	27	60	7	2.67	.71
25	Using professional practices data directs attention to achievement gaps in classrooms						
	Teachers	18	39	39	5	2.31	.81
	Administrators	0	48	45	7	2.59	.63
26	Using student learning objectives data helps identify low-performing teachers						
	Teachers	31	45	20	4	1.97	.81
	Administrators	20	37	40	3	2.27	.83
27	Using professional practices data helps identify low-performing teachers						
	Teachers	20	33	42	5	2.32	.85
	Administrators	7	20	60	13	2.80	.76

† Percentages may not total 100 due to rounding.
* SD = *strongly disagree*, D = *disagree*, A = *agree*, SA = *strongly agree*

administrators (83%) agreed that using professional practices data accurately contributes to evaluating teaching.

In a point of divergence, only 40% of teachers agreed that using student learning objectives data makes evaluations more objective, while 57% administrators agreed. Nonetheless, that left 43% of administrators who disagreed that the student learning

objectives data made evaluations more objective. In contrast, half of the teachers (51%) and three fourths of the administrators (76%) agreed that using professional practices data makes the evaluations more objective. Both teachers (51%) and administrators (67%) tended to agree that using student learning objectives directs attention to potential achievement gaps for students in individual classrooms. However, fewer teachers (44%) and administrators (52%) agreed that using professional practices data does the same. Three quarters of teachers (76%) and a majority of administrators (57%) did not agree that using student learning objectives data help administrators identify low-performing/ ineffective teachers, but more teachers (47%) and administrators (73%) agreed using professional practices data helps identify low-performing/ineffective teachers.

Research Question 2

Are there differences between teachers and administrators at different levels (elementary, middle, and high) in their perceptions of the evaluation system as implemented to date?

The responses of administrators were used as test values in one-sample *t* tests using the teachers' responses at each school level. The null hypothesis of each one-sample *t* test was that the mean of the teachers would equal to the mean of the administrators. In each case, the teachers' perceptions of the propriety, utility, feasibility, and accuracy of the district's new teacher evaluation system were significantly lower than the perceptions of the administrators (Table 19). Therefore, it can be concluded that there are statistically significant differences between the perceptions of teachers and administrators at all school levels.

Table 19

Differences in Teachers and Administrators' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels

	Administrators (n = 30)			Teachers (n = 470)						
	n	*M*	*SD*	*n*	*M*	*SD*	*Mean diff*	95% CI of difference	*t*	*p*
Propriety										
Elementary	16	2.90	.38	206	2.63	.61	-.27	-.35 – -.18	-6.22	< .001
Middle	11	3.03	.46	111	2.54	.59	-.60	-.60 – -.38	-8.88	< .001
High	3	2.78	.51	153	2.55	.60	-.23	-.33 – -.14	-4.82	< .001
Utility										
Elementary	16	2.67	.36	206	2.45	.66	-.22	-.31 – -.13	-4.78	< .001
Middle	11	2.74	.42	111	2.27	.61	-.47	-.59 – -.36	-8.14	< .001
High	3	2.83	.29	153	2.35	.61	-.48	-.58 – -.39	-9.83	< .001
Feasibility										
Elementary	16	2.83	.37	206	2.42	.63	-.41	-.49 – -.32	-9.32	< .001
Middle	11	2.83	.44	111	2.25	.62	-.58	-.70 – -.46	-9.82	< .001
High	3	2.67	.31	153	2.33	.65	-.34	-.44 – -.24	-6.49	< .001
Accuracy										
Elementary	16	2.56	.50	206	2.29	.65	-.27	-.36 – -.18	-5.98	< .001
Middle	11	2.71	.49	111	2.07	.66	-.64	-.76 – -.52	-10.26	< .001
High	3	2.63	.33	153	2.19	.65	-.44	-.54 – -.33	-8.30	< .001

Research Question 2 was analyzed by comparing the means of teachers and administrators across the three school levels. No significant differences in perceptions of the JCSEE attributes were found among the administrators across school levels (Table 20). This may have been in part due to the small sample size of administrators, with a total number of administrators of 30, and as few as 3 at the high school level. Therefore,

Table 20

Differences in Principals' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels

Source	*SS*	*df*	*MS*	*F*	*p*
Propriety					
Between groups	0.20	2	.10	.56	.58
Within groups	4.78	27	.18		
Total	4.98	29			
Utility					
Between groups	0.08	2	.04	.27	.76
Within groups	3.83	27	.14		
Total	3.91	29			
Feasibility					
Between groups	0.07	2	.03	.22	.81
Within groups	4.22	27	.16		
Total	4.29	29			
Accuracy					
Between groups	0.15	2	.07	.31	.73
Within groups	6.31	27	.23		
Total	6.46	29			

the statistical analysis lacked power. Among the teachers, significant differences were found, however, in utility and accuracy across different school levels (See Table 21).

Post hoc Bonferroni comparisons determined that elementary teachers (M = 2.45) had significantly higher levels of agreement about the utility of the teacher evaluation system than did middle teachers (M = 2.27), although mean scores of both groups were below the midpoint. Thus, elementary teachers disagreed less strongly than middle school teachers did on this standard. Moreover, elementary teachers (M = 2.29) also had

Table 21

Differences in Teachers' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System at the Elementary, Middle, and High School Levels

	SS	*df*	*MS*	*F*	*p*
Propriety					
Between groups	0.99	2	.50	1.37	.26
Within groups	168.66	467	.36		
Total	169.65	469			
Utility					
Between groups	2.63	2	1.32	3.30	.04
Within groups	186.35	467	.40		
Total	188.99	469			
Feasibility					
Between groups	2.26	2	1.23	2.83	.06
Within groups	186.30	467	.40		
Total	188.55	469			
Accuracy					
Between groups	3.39	2	1.70	3.98	.02
Within groups	199.11	467	.43		
Total	202.50	469			

significantly less disagreement in the accuracy of the teacher evaluation system than did middle teachers (M = 2.07). For both of these standards, the levels of agreement of high school teachers about the utility (M = 2.35) and accuracy (M = 2.19) of the evaluation system fell between the elementary and middle teachers and were not significantly different from either elementary or middle school teachers. No significant differences were found among the teachers at different school levels concerning the propriety of the teacher evaluation system.

Research Question 3

Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the evaluation system as implemented to date?

To answer Research Question 3, the teachers and administrators' responses were analyzed using three lenses: mean rating of each JCSEE attribute scale, mean rating of only statements asking about student learning objectives, and mean rating of professional practices statements. The responses were also analyzed between groups (teachers and administrators) and within groups. Table 22 contains the means and standard deviations for each JCSEE attribute by total scale (all items in scale), student learning objectives items, and professional practices items.

Table 22

Means and Standard Deviations of Teachers and Administrators' Ratings of JCSEE Attributes and Items Measuring Student Learning Objectives and Professional Practices

	Teachers (n = 470)		Administrators (n = 30)	
Standard	*M*	*SD*	*M*	*SD*
Propriety	2.58	.60	2.93	.41
Student learning objectives	2.31	.76	2.65	.70
Professional practices	2.66	.72	3.02	.43
Utility	2.37	.64	2.71	.37
Student learning objectives	2.23	.70	2.54	.55
Professional practices	2.52	.67	2.88	.39
Feasibility	2.35	.63	2.81	.38
Student learning objectives	2.24	.69	2.66	.56
Professional practices	2.53	.68	3.03	.39
Accuracy	2.21	.66	2.62	.47
Student learning objectives	2.07	.70	2.45	.67
Professional practices	2.34	.73	2.80	.54

Mean ratings for each JCSEE attribute scale (between groups). The aggregated data for all teachers and all administrators were tested in a series of one-

sample *t* tests (Table 23). In every measure of the JCSEE attributes, the administrators' mean levels of agreement were significantly higher than the teachers' mean levels of agreement on the total attribute scales, those items measuring student learning objectives, and those items measuring professional practices.

Table 23

Differences in Teachers and Administrators' Perceptions of the Propriety, Utility, Feasibility, and Accuracy of the District's Teacher Evaluation System

	Administrators (n = 30)		Teachers (n = 470)					
	M	*SD*	*M*	*SD*	*Mean diff*	95% CI of difference	*t*	*p*
Total								
Propriety	2.93	.41	2.58	.60	-.35	-.40 – -.29	-12.50	< .001
Utility	2.71	.37	2.37	.64	-.34	-.39 – -.28	-11.50	< .001
Feasibility	2.81	.38	2.35	.63	-.46	-.51 – -.40	-15.54	< .001
Accuracy	2.62	.47	2.21	.66	-.41	-.47 – -.35	-13.66	< .001
Student learning objectives								
Propriety	2.65	.70	2.31	.76	-.34	-.41 – -.27	-9.78	< .001
Utility	2.54	.55	2.23	.70	-.31	-.37 – -.25	-9.64	< .001
Feasibility	2.66	.56	2.24	.69	-.42	-.49 – -.36	-13.29	< .001
Accuracy	2.45	.67	2.07	.70	-.38	-.44 – -.31	-11.65	< .001
Professional practices								
Propriety	3.02	.43	2.66	.72	-.36	-.42 – -.29	-10.77	< .001
Utility	2.88	.39	2.52	.67	-.36	-.42 – -.30	-11.74	< .001
Feasibility	3.03	.39	2.53	.68	-.50	-.56 – -.44	-15.96	< .001
Accuracy	2.80	.54	2.34	.73	-.46	-.53 – -.39	-13.67	< .001

Mean ratings for each JCSEE attribute scale (within groups). A series of paired samples *t* tests was used to determine if the rating of the scales were statistically different within each group (teachers and administrators; see Table 24). From a statistical

point of view, teachers rated utility (M = 2.37) and feasibility (M = 2.35) the same (p = .31), but rated propriety (M = 2.58) higher than the other three scales (p < .001), while rating the accuracy scale (M = 2.21) the lowest of the four scales (p < .001). The delineation of the scales is less clear in the administrators' ratings. The mean of propriety was 2.93, while feasibility had a mean of 2.81. Administrators were less in agreement with the utility items (M = 2.71) and items in the accuracy scale (M = 2.62). Statistically, administrators rated propriety and feasibility the same (p = .04), while rating propriety statistically higher (p < .001) than utility and accuracy. Feasibility was rated statistically higher (p = .02) than accuracy, but the same as utility (p = .12).

Table 24

Comparison of Ratings of JCSEE Attributes Within Teachers and Administrators

	Paired differences					
Comparison	*M*	*SD*	95% CI of difference	*t*	*df*	*p*
Teachers (*n* = 470)						
Propriety – Utility	.21	.36	.18 – .24	12.78	469	< .001
Propriety – Feasibility	.23	.46	.19 – .27	10.81	469	< .001
Propriety – Accuracy	.38	.47	.33 – .42	17.34	469	< .001
Utility – Feasibility	.02	.44	-.02 – .06	1.02	469	.31
Utility – Accuracy	.17	.40	.13 – .20	9.17	469	< .001
Feasibility – Accuracy	.15	.40	.11 – .18	7.87	469	< .001
Administrators (*n* = 30)						
Propriety – Utility	.22	.32	.10 – .34	3.74	29	< .001
Propriety – Feasibility	.12	.32	.00 – .24	2.10	29	.04
Propriety – Accuracy	.31	.39	.16 – .46	4.33	29	< .001
Utility – Feasibility	-.10	.33	-.22 – .03	-1.63	29	.12
Utility – Accuracy	.09	.37	-.05 – .23	1.32	29	.20
Feasibility – Accuracy	.19	.40	.04 – .34	2.57	29	.02

Mean ratings for each JCSEE attribute scale of only SLO statements (within groups). Using the items rating *student learning objectives* (Table 25), the analyses found that teachers rated utility (M = 2.23) and feasibility (M = 2.24) the same (p = .78), but rated propriety (M = 2.31) higher than the other three scales (p < .001), while rating the accuracy scale (M = 2.07) the lowest of the four scales (p < .001). Administrators rated propriety (M = 2.65) and feasibility (M = 2.66) the same (p > .05), while feasibility and propriety were statistically higher (p < .001) than utility (M = 2.54).

Table 25

Comparison of Ratings of Student Learning Objectives Items Within Teachers and Administrators by JCSEE Attributes

	Paired differences					
Comparison by group	*Mean diff*	*SD*	95% CI of difference	*t*	*df*	*p*
Teachers (*n* = 470)						
Propriety – Utility	.08	.46	.04 – .12	3.65	469	< .01
Propriety – Feasibility	.07	.57	.02 – .12	2.68	469	< .01
Propriety – Accuracy	.23	.58	.18 – .29	8.78	469	< .01
Utility – Feasibility	-.01	.52	-.05 – .04	-.28	469	.78
Utility – Accuracy	.16	.47	.11 – .20	7.26	469	< .01
Feasibility – Accuracy	.16	.46	.12 – .20	7.66	469	< .01
Administrators (*n* = 30)						
Propriety – Utility	.11	.43	-.05 – .27	1.38	29	.18
Propriety – Feasibility	-.01	.56	-.22 – .20	-.11	29	.91
Propriety – Accuracy	.20	.53	.00 – .40	2.06	29	.05
Utility – Feasibility	-.12	.43	-.28 – .04	-1.54	29	.14
Utility – Accuracy	.09	.51	-.10 – .28	.98	29	.33
Feasibility – Accuracy	.21	.46	.04 – .38	2.53	29	.02

Using the items rating *professional practices* (Table 26), the analyses found that teachers rated utility (M = 2.52) and feasibility (M = 2.53) the same (p = .59), but rated propriety (M = 2.66) statistically higher than the other three scales (p < .001), while rating the accuracy scale (M = 2.34) the lowest of the four scales (p < .001). Statistically, administrators rated propriety (M = 3.02) and feasibility (M = 3.03) the same (p = .75). Utility (M = 2.88) and accuracy (M = 2.80) were also rated the same (p = .21). Finally, propriety and feasibility were statistically higher (p < .05) than utility and accuracy.

Table 26

Comparison of Ratings of Professional Practice Items Within Teachers and Administrators by JCSEE Attributes

	Paired differences					
Comparison	*Mean diff*	*SD*	95% CI of difference	*t*	*df*	*p*
Teachers (*n* = 470)						
Propriety – Utility	.14	.42	.11 – .18	7.41	469	< .001
Propriety – Feasibility	.13	.55	.08 – .18	5.24	469	< .001
Propriety – Accuracy	.32	.56	.27 – .37	12.42	469	< .001
Utility – Feasibility	-.01	.57	-.05 – .03	-.54	469	.59
Utility – Accuracy	.18	.45	.14 – .22	8.66	469	< .001
Feasibility – Accuracy	.19	.47	.15 – .23	8.79	469	< .001
Administrators (*n* = 30)						
Propriety – Utility	.13	.28	.03 – .24	2.64	29	.01
Propriety – Feasibility	-.02	.28	-.12 – .09	-0.33	29	.75
Propriety – Accuracy	.22	.35	.09 – .35	3.43	29	< .01
Utility – Feasibility	-.15	.33	-.27 – .03	-2.52	29	.02
Utility – Accuracy	.08	.36	-.05 – .22	1.28	29	.21
Feasibility – Accuracy	.23	.44	.07 – .40	2.94	29	.01

Summary

Responses from teachers and administrators were used to assess the degree to which teachers and administrators concurred that the new teacher evaluation program met the propriety, utility, feasibility, and accuracy attributes of the JCSEE. Agreement scores ranged from 1 (*strongly disagree*) to 4 (*strongly agree*). Three research questions were answered:

- While teachers tended to *disagree* that the evaluation system met the attributes of the JCSEE and administrators tended to *agree*, both teachers and

administrators' levels of agreement were highest on the propriety and lowest on the accuracy of the evaluation system. Teachers were less likely than were administrators to agree to all items measuring the JCSEE attributes. However, there was more agreement from both groups that the use of professional practices data was more in line with the JCSEE standards than using student learning objectives data.

- Statistically significant differences were found between teachers and administrators at each school level across all four attributes. No significant differences were found among the three levels of administrators; however, elementary teachers disagreed less strongly about the utility and accuracy attributes of the evaluation system than middle school teachers.
- Analyses found that administrators' levels of agreement on all of the JCSEE attributes were statistically higher than the teachers' mean levels of agreement.

A discussion of those results, conclusions drawn from the analyses, implications of the results, and recommendations for further research are found in Chapter 5.

CHAPTER 5: DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Effective teaching is at the forefront of the educational debate. The effect a teacher's instructional practices have on students is well-established (Aaronson et al., 2007; Coleman et al., 1966; Hanushek, 1992, 2011; Nye et al., 2004; W. L. Sanders & Rivers, 1996; Stronge et al., 2008). While the research is plentiful regarding the necessity for every child to be taught by an effective teacher, those individuals who evaluate teachers must be judicious in the development and implementation of tools that accurately and fairly evaluate teacher performance (Danielson, 2002; DiPaola & Hoy, 2012; Stronge, 2010b; Weisberg et al., 2009). However, policymakers developing new evaluation programs to measure effective teaching and learning often fail to solicit the input of the teachers and principals who are most affected (Behrstock-Sherratt et al., 2013; DiPaola & Hoy, 2012; Johnson, 2012; Ovando, 2001; Stronge & Tucker, 1999). Moreover, it is important to understand how teachers' emotions and perceptions influence their sense making and coping with the changes in new evaluation policies (Hargreaves, 2000; Zembylas & Barker, 2007).

Policymakers as well as school leaders must be adept in understanding not only the change process, but also the effect on teachers and principals' perceptions of what the change will cost them in terms of working conditions, their values, and relationships (van Veen et al., 2005). Therefore, an understanding of the perceptions of those most affected by the new evaluations programs for teachers is crucial for implementing and sustaining a successful evaluation program. The purpose of this quantitative study was to assess the

degree to which both teachers and school administrators in the Emerald County School District concurred that the teacher evaluation program met the propriety, utility, feasibility, and accuracy attributes of the JCSEE (2014a). The study's evaluation approach was grounded in Alkin's (2004) use branch theory and followed the context, input, process, and product (CIPP) model developed by Stufflebeam (1968, 2007). Additionally, this study examined only the perceptions of teachers and principals in Emerald County and not the merit of the new evaluation program for teachers. Hall (1976) stated that, "Information about concerns can be of great help determining the kinds of implementation and supporting actions that users will see as personally relevant and will also be effective in reducing problems and advancing the Level of Use of the innovation'' (p. 23).

Discussion of the Findings

Literature abounds regarding the need to develop evaluation programs for teachers that conform to the mandates of ESEA and RTT. The quest for school districts is not only to develop effective and quality programs, but also to ensure that the evaluation programs conform to standards that are fair and equitable. Furthermore, it is important for districts to understand how those most affected by the evaluation program perceive how the evaluation process is changing and how those changes affect them personally; otherwise, the misunderstanding of the purpose of teacher evaluation could hinder teacher growth and the program itself (Popham, 2013). Three research questions guided this study. The findings of each are discussed in relation to the comparative research literature for this study.

Research Question 1

What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date?

While teachers tended to disagree that the evaluation system met the attributes of the JCSEE and administrators tended to agree, both teachers and administrators' levels of agreement were highest on the propriety and lowest on the accuracy of the evaluation system. However, there was more agreement from both groups that the use of professional practices data was more in line with the JCSEE standards than using student learning objectives data.

Propriety. Propriety standards protect the rights of the persons affected by the evaluation. The standards require evaluators to understand and obey the laws concerning areas such as privacy, access to information, diversity, and the protection of human subjects. Slightly more than half of administrators and teachers agreed that the county provided clear and concise documentation of the procedures and guidelines outlining the procedures of the evaluation program. A majority of both teachers and administrators agreed there was a clear understanding of the expectations for teacher job performance. These results are consistent with the research evidence that teachers favor having an understanding of the standards or indicators by which they are evaluated as well as knowing how the evaluation will be conducted (Conley et al., 2005; Giliya, 2006; Milanowski & Heneman, 2001; Pizzi, 2009; Seyfarth, 2001). Studies also show that teachers would like more transparency in the evaluation programs (Castillo, 2005; Feeney, 2007).

When responding to items regarding professional discussions and documenting teachers' areas of strengths, a majority of teachers and administrators agreed that using the professional practices data encouraged professional discussion during follow-up conferences. However, fewer teachers and administrators agreed that SLO data encouraged professional discussion during follow-up conferences. In a point of divergence, nearly two thirds of teachers *disagreed* that using SLO data document teachers' strengths, while a majority of administrators *agreed*. Teachers have regularly argued against the use of student performance data because it fails to recognize the inherent differences in every classroom and every school (Kelsey, 2009). Teachers are fearful of what harm or consequences could come to them if test results are interpreted incorrectly by principals or district officials, creating uncertainty for teachers and administrators in the face of new evaluation programs (Conley & Glasman, 2008; Emery & Ohanian, 2004). The teachers and administrators responding to this study tended to agree about the use of the professional practice component of the evaluation, while tending to disagree about the use of the SLO component of the evaluation. Requiring SLO data in the new evaluation program for teachers may elicit various levels of alarm.

Utility. The objective of the utility standard is that effective evaluations support educators and administrators in their professional growth, thereby providing educational professionals with identified areas for improvement in instructional practices to achieve the mission and goals of the organization (JCSEE, 2009, 2014b). Researchers agree that incorporating measures that use data on the achievement of the teachers' students is a central component of the new reform (e.g., Danielson, 2007; Gates Foundation, 2013; Milanowski, 2004). Although all of the responses fell between 2.0 (disagree) and 3.0

(agree), the current results indicate that teachers tended more toward *agree* with the professional practice component of the evaluation system. These results depart from previous research suggesting that most evaluation programs did little to improve practice or instruction, produced minimal results with changes in teaching and learning, and had little influence in improving teaching (Colby, Bradshaw, & Joyner, 2002; Peterson, 2000; Weisberg et al., 2009).

A little more than half of the teachers (56%) responding in the Emerald County School District indicated that the professional practices component of the evaluation program was useful in providing feedback addressing their areas of strength and areas needing improvement. Milanowski (2004) suggested that standards-based teacher evaluation systems based on the Danielson framework for teaching appear to have the potential to provide measurements of teacher effectiveness that may be strongly related to student achievement. Milanowski reported that teacher evaluation scores were positively related to higher than expected levels of achievement.

Evaluation systems that have utility regularly apply the evaluation results to improve staff performance (JCSEE, 2009). However, teachers tended toward *disagree* that neither the professional practices (44%) component, nor the SLO (36%) component improved teaching. Teachers' perceptions may in part be due to a misconception of how the new evaluation system uses the evaluation results to not only recognize teachers' strengths and weaknesses, but also to promote individual professional development. The teachers' perceptions may be due to a lack of understanding of how the professional practices and student learning objects relate to each other, as well as how the two components help improve instructional practices. The district should seek to align its

practices and standards for professional development with federal legislation, peer-reviewed research, and professional development organizations dedicated to standards of practice based on sound research.

Moreover, fewer teachers and administrators agreed that using SLO data improves teaching than those who agreed that using professional practices data improves teaching. The districts' principals were provided SLO professional development by personnel from the district's research and assessment team. However, teachers' understanding of both the SLO and professional practices components was dependent on training opportunities created in each school building by the principal. Differences in how the training was delivered posed questions about fidelity, quality, and consistency of the teachers' professional development across schools in the district. Instead of focusing on how to calculate SLOs, perhaps there is a need for a more pragmatic, richer, and greater understanding of how SLOs can provide data that are useful in improving achievement through improved instructional practices. Emerald County School District should consider using a professional development model other than train-the-trainer to ensure consistently high quality training in the evaluation program. Another consideration is that the school district consider the JCSEE (2009) guidelines for improving accuracy (A1-Valid Judgment) by "ensuring evaluators are well trained…and avoid systematic bias such as the 'halo-effect' in which a general impression or previous rating influences the present rating" (p. 118).

Professional development is critical when implementing changes to evaluation programs, particularly when incorporating new measures of evaluation. Using the SLO data to measure student growth is complex and requires districts to consider multiple

factors for quality and successful implementation. Without a coherent theory of action depicting how SLOs are intended to promote and support instructional practice, student learning, and the district vision and mission, teachers may not perceive the benefits of how using SLOs in teacher evaluations informs instruction.

Provisions for professional development should be included in the theory of action to provide stakeholders the understanding and skills to use SLOs to reflect on improving instructional practices and how to align best practices of instruction with the college and career readiness standards, and the district, school, and grade-level goals. Understanding how the assessments of the SLOs are developed and administered is of utmost priority for professional development. Not only do teachers and principals need professional development to learn how to identify and develop quality assessments for measuring student progress and to link specific objectives with specific assessments, but they also need training in data and assessment literacy. SLO assessments in the school district are developed by a team of teachers and district office personnel; therefore, improving assessment and data literacy for teachers and principals is essential.

It is also important to clarify the difference between SLOs and the SLO assessment. The SLO is a specific long-term goal for student learning that is customized to a teacher's particular students. SLOs are designed to both support instruction and measure student growth for teacher evaluation. In contrast, the SLO assessment is the instrument used to measure SLOs. Creating a SLO assessment varies by states and districts. The three central roles of SLO assessment are (a) pre-assessment, assessing student learning prior to teaching; (b) formative assessment, assessing how student learning is incorporated into instructional practices; and (c) summative assessment,

assessing student learning at the instructional period (Gareis & Grant, 2015). Providing teachers and principals the scope and breadth of assessment literacy is crucial in improving perceptions in an evaluation program's assessments tools. Developing assessment literacy can enhance perceptions of the reliability and validity of the new evaluation program.

Sustaining the quality of an SLO process is dependent on the quality of the measurement used to define students' beginning and end-of-year performance level. In response, various states recognize a list of pre-approved assessments measuring students' performance for use by schools and districts. However, in some states individuals or groups of teachers develop assessments to measure student growth. In these instances, districts must assess the quality of the measurements. Gareis and Grant (2015) offered a ranking of assessment types based on how the assessments align to rigorous, valid, and reliable standards. Ranking from highest to lowest are (a) assessments created by the state and containing items proportionate to the content specified in the SLO; (b) commercially available assessments; (c) assessments created by teams in school districts, provided they meet the criteria for assessments and are administered in classrooms throughout the district in order to increase comparability across classrooms; and (d) teacher-created assessments used by teachers other than the designer.

The utility standard provides that evaluators should not only identify those who will use the evaluation system and how stakeholders will use the evaluation result, but also who possess the qualifications, skills, training, and authority to conduct personnel evaluations (JCSEE, 2009). Teachers and administrators in Emerald County School District differed in their agreement as to whether administrators are qualified to evaluate

the SLO and professional practices components of the system. Slightly more than half of the teachers (53%), but almost all of the administrators agreed that administrators are qualified to use the SLO component. All of the administrators and a majority of the teachers (66%) agreed that administrators are qualified to use the professional practices component.

A consideration for this finding may be related to the time and quality of professional development allotted to principals for both the SLO and professional practices component of the new evaluation program. Personnel from the district's research and assessment team provided limited SLO training for principals. In addition, principals received several days of intense training from The Danielson Group on how to use the Danielson framework. School administrators were also required to complete a self-paced online training program developed to help observers increase their reliability and accuracy in identifying, categorizing, and scoring evidence of teaching practice. The modules of the course covered all four domains of the Danielson framework for teaching. After completing the online course, administrators were required to pass a two-part assessment. With their SLO and Danielson framework training, principals were charged with planning and implementing training for the teachers. Upon request from individual principals, district personnel provided support to train teachers on developing SLOs.

Educational reform such as new evaluation programs for teachers produces elements of change and conflict due to political interest and power. These elements of change and conflicting perspectives may produce resistance by those who perceive the evaluation puts them at a disadvantage (Taut & Alkin, 2003). Teachers experiencing fear and a lack of trust with the evaluation may demonstrate several forms of resistance

toward the new evaluation program, such as questioning the competency and qualifications of the evaluator and having misconceptions about the purpose and objectives of the evaluation (Youngcourt, Leiva, & Jones, 2007). Emotions have a significant influence on teachers' reaction to educational reform efforts, ranging from compliance to conflict or opposition (Fullan, 2011; Hargreaves, 2001; Roussin & Zimmerman, 2014; Schmidt & Datnow, 2005; van den Berg, 2002).

Approximately half of the teachers and administrators agreed that the criteria for using professional practices data are clear and accurate and even more agreed that using professional practices data does inform professional development activities. However, a majority of both teachers and administrators disagreed that the criteria for using SLO data in rating teacher performance are clear and accurate, and even fewer agreed that using SLO data informs professional development activities. These findings are in contrast to Proctor et al. (2011), who found that 50% of teachers reported that using SLOs affected professional growth. The Tennessee Department of Education (2012) studied teachers' perceptions of SLOs and found that using SLOs in the evaluation program provided more intentional use of student data, more schoolwide collaboration, and new kinds of conversations around instruction and outcomes. However, the findings related to the professional practices components in this study concur with Tuytens and Devos (2010), who found a trend toward teachers engaging in professional development after receiving feedback and the influence of active leadership supervision on teachers' perceptions of both feedback and utility.

Feasibility. The standard of feasibility operates on the premise that personnel evaluations occur in a real world context influenced by multifaceted dynamics, such as

evaluation procedures and approaches, political pressures, and potential limitations of resources. Incorporating this standard into an evaluation program for teachers can increase the likelihood that the evaluation program is efficiently implemented, user friendly, and viable regardless of political constraints, as well as adequately funded (JCSEE, 2009).

In the current study, a majority of administrators agreed that using SLO data is a responsible use of assessment data and provides two-way communication between administrators and teachers. The current results also indicated that almost half of the teachers (47%) agreed that using SLO data provides two-way communication between administrators and teachers; however, these results are not consistent with the Austin Independent School District (2012) study that found elementary teachers who used student learning objective data were more likely than teachers not using SLO data to (a) discuss professional development needs and goals, (b) communicate assessment data for individual students, (c) set student learning objective goals for groups of students, and (d) group students based on learning needs. Teachers responding to the Emerald County School District Study were almost evenly divided in their perceptions of the feasibility of the SLO data to improve communication relating to professional development needs and the responsible use of SLO assessment data. Behrstock-Sherrat et al. (2013) reported that while using SLO data encourages teacher engagement in the evaluation process, setting objectives at a level that is obtainable yet allows students to obtain their maximum performance is complicated.

Teachers need guidance, professional development, resources, and appropriate tools to implement SLOs successfully. The school district used the train-the-trainer model

to provide teachers with SLO professional development. This model provided the principals with professional development from the district's research and assessment team. Principals were charged with creating and implementing a SLO professional development plan for teachers in their schools. Concerns of consistency and fidelity to the training program across schools and for all teachers call into question the train-the-trainer model. Those affected by new evaluation systems must be provided professional development that is consistent, reliable, accurate, practical, and efficient. The JCSEE standards require that personnel evaluations reinforce positive behaviors, improve evaluatee understanding of skills, and promote personnel evaluations leading to professional development (JCSEE, 2009).

Research studies, including the Austin Independent School District (2012) revealed that teachers wanted support in setting and implementing SLOs. The Austin Independent School District (2012) study found that teachers requested enhanced direction on the SLO assessment process and that some respondents were unfamiliar with the measures in use. A study of SLOs in Denver found that teachers originally considered the SLO setting procedure to be difficult and needed greater support and feedback (Community Training and Assistance Center, 2004). Consistent with the Austin (2012) results are the findings in the current study that teachers and administrators tended to disagree that using the SLO data is feasible. The current study revealed that a majority of administrators agreed that the teacher time required for employing professional practices data in the new evaluation system is feasible. However, teachers were equally divided between *disagree* and *agree* that the teacher time required for employing professional practices data in the new evaluation system is feasible.

Accuracy. To meet the accuracy standard, the evaluation must serve its intended purpose and the results must be correct, defensible, and based on a sound system of evaluation. Evaluation decisions must be based on the explicit criteria of the evaluation program where the evaluator followed the procedures and accurately analyzed the data leading to the outcomes of the evaluation so that the validity of the results is protected (JCSEE, 2009). The items measuring accuracy contained statements about the extent to which teachers and administrators agreed that the use of SLO data and professional practices accurately contributes to evaluating teaching, making the evaluations more objective, and helping administrators identify low-performing/ineffective teachers. In each case, teachers were in less agreement than were administrators. However, both teachers and administrators were in more agreement about the use of professional practices data than the use of SLO data.

Research relating to the reliability and validity of SLO data suggests limited data for their statistical properties. The relationships were more specific with value-added measures and year-to-year reliability (Proctor et al., 2011; Schmitt & Ibanez, n.d.; Tennessee Department of Education, 2012; Terry, 2008). Most of these studies focused on the evidence consisting of implementation lessons with the SLO data. In a review of research on SLO data, Gill et al. (2013) identified fundamental areas for implementing SLO data, such as provisions for teachers to obtain suitable training, the use of appropriate tools for creating SLOs, as well as acquiring data, and finally considering validity concerns that may arise when teachers set SLO targets.

In other research regarding the fairness of SLO data, implementation findings from the Austin Independent School District (2012) suggested participants were

frustrated that variables such as student mobility, and dropout and attendance rates affect teachers' ability to meet SLO targets and yet were not taken into consideration in the evaluation system. Burns, Gardner, and Meeuwsen (2009) found that two thirds of teachers in another study in Austin reported positive perceptions of instructional purposes for SLOs; however, two thirds of the responding teachers also disagreed that SLOs provide a positive measure of effective teaching. The Tennessee Department of Education (2012) determined that teachers perceived the SLO component of the evaluation program to be the least effective, as groups of teachers were inconsistent in selecting the same measures due to teacher and principal speculation on which assessments would yield the greater scores. In a similar study conducted by Proctor et al. (2011), responding teachers expressed concerns about the consistency of the implementation of SLOs.

A majority of teachers and administrators in the current study did not agree that using SLO data helped administrators identify low-performing/ineffective teachers, but more teachers and administrators agreed that using professional practices data helps identify low-performing/ineffective teachers. In a point of divergence, a smaller proportion of teachers agreed that using SLO data makes evaluations more objective than did administrators. In contrast, a majority of teachers and administrators agreed that using professional practices data makes the evaluations more objective. Although teachers and administrators do not agree that using SLO data helps to identify low-performing/ ineffective teachers, they tended to agree that using SLO data directs attention to potential achievement gaps for students in individual classrooms. At the same time, they acknowledged that using professional practices data does not.

Research Question 2

Are there differences between teachers and administrators at different levels (elementary, middle, and high) in their perceptions of the evaluation system as implemented to date?

The results indicated in every case that the teachers' perceptions of the JCSEE standards of propriety, utility, feasibility, and accuracy of the district's new teacher evaluation system were significantly lower than the perceptions of the administrators. In all areas, teachers tended toward *disagree*. Interestingly neither teachers nor administrators overall ratings of *disagree* or *agree* reached either the 2.0 mark of *disagree* or the 3.0 mark of *agreement,* suggesting ambivalence with the program.

One area of focus imposed under Race to the Top policies in teacher evaluation emphasizes the significance of the principal's supervisory responsibilities of the new evaluation system for teachers. Principal leadership is vital to successful implementation of high-accountability, state-mandated teacher evaluation systems. Therefore, understanding the perceptions of principals toward implementing these complex changes in teacher evaluation programs is crucial for effective change. Moreover, it is important that researchers consider not only principals' concerns but also their perceptions of the implementation support in the practice of leading change during this time of accountability and reform.

Principals play a substantial and significant role in the implementation of new teacher evaluation programs. The ability and motivation of the principal to enact change is critical for successful implementation of any school reform, particularly in the reform of teacher evaluations (Fowler, 2009). Current reforms requiring the use of evidence-

based data aligned with improved student performance places higher demands and expectations on the supervisory roles and responsibilities of principals for school improvement (Anderson et al., 2010). Furthermore, if districts are to provide principals the needed support, interventions, and resources to successfully implement and sustain successful reform in teacher evaluation programs, then districts must understand principals' perceptions and concerns regarding the changes in teacher evaluation as well as concerns with implementation of the new evaluation. Principals, through their roles of instructional leader, must not only be committed to achievement, but also be willing and able to initiate and facilitate the essential changes regardless of the complexity in the new evaluation program for teachers.

Hallinger and Heck (2011) reported that the effect of school leaders on student achievement is noteworthy. For new evaluation programs to be successful, districts should reconsider the role of the principal during the change process. If successful change is to occur with the implementation of the new evaluation program, districts must encourage strong leadership that makes, encourages, and enhances teaching as a way of life in every school (Donaldson, 2012). Strong school leadership propels teachers toward commitment to their craft by committing to high expectations for instruction, building a culture of trust and risk taking, and promoting reflection and professional growth. Principals must help teachers examine their emotions and concerns regarding the new evaluation system by providing an environment where the evaluation results are seen positively and useful for improving instructional practices (Zepeda, 2011).

Sullivan and Glanz (2005) equated the principal's role with the attributes found in the role of the classroom teacher, as "a mentor, inspirer, and a facilitator of learning" (p.

162). Trusting relationships between teachers and principal are limited when teachers perceive the evaluation process as "an empty process or as retribution or manipulation" (Zepeda, 2011, p. 53). Therefore, school leaders must not only endorse the virtues of the teacher evaluation program, school leaders must also hold an obvious commitment for their own growth (Duke & Stiggins, 1986). Principals also need support and resources to build and promote teacher commitment for the new evaluation program by advocating aggressively for resources, funding, and time for teachers to reflect and gather evidence (Derrington & Campbell, 2015).

When researching the effect that effective communication of high school principals has on school climate, Halawah (2005) suggested, "Effective principals recognize the unique styles and needs of teachers and help them achieve their own performance goals" (p. 336). The type of feedback provided during the evaluation process must lend itself toward improving and enhancing both professional practices and the individual growth of the teacher (Marshall, 2013). Subsequently, principals must consider a change in attitudes and actions toward the quality and quantity of supervision of the new evaluation program to include multiple mini-observations followed by one-on-one conversations to help improve teaching practices (Marshall, 2013).

Overall, teachers participating in the Emerald County School District study tended to *disagree* that the new evaluation program met the standards of utility and accuracy. Elementary and middle school teachers differed in their perceptions of both the utility and accuracy standards. The elementary school teachers indicated significantly higher levels of agreement about the utility standard than did middle school teachers. Moreover, elementary teachers indicated significantly less disagreement in the accuracy

of the teacher evaluation system than did middle teachers. The differences in teachers' perceptions across levels concerning the feasibility standard were very close to being statistically significant. Analysis that compared administrator responses across the three grade levels (elementary, middle, high) revealed no significant differences in perceptions of the JCSEE attributes among the administrators. However, these findings lacked statistical power due to the small sample size of administrators in each grade level.

The responding teachers in Emerald County School District perceived that the county's new evaluation program was limited in the JCSEE standards of utility, feasibility, and accuracy. These standards provide that evaluations not only contain measures for accountability, but also provide for professional development leading to student learning. DiPaola and Hoy (2014) reported that, "Evaluation needs are basic; the need for thoughtful, thorough, and fair evaluation based on performance and designed to encourage improvement in both the person being evaluated and the school" (p. 159).

Evaluations having limited alignment with the JCSEE standards may produce negative perceptions from the evaluatees. Vekeman, Devos, and Tuytens (2015) found that most teachers initially feared that new evaluation programs would be solely summative in rating teacher job performance. Tuytens and Devos (2009) found that teachers expressed concerns regarding how schools implemented the new teacher evaluation policy even though the teachers' perceptions were positive toward the new teacher evaluation policy. Other research found that teachers had summative expectations regarding the new evaluation program, resulting in teachers' increased fear that the new evaluation systems would result in greater teacher control and sanctions, thus giving teachers a negative perception of the new evaluation (Flores, 2012; Morgado & Sousa,

2010; Stronge & Tucker, 1999). However, evaluations that have a greater alignment with the JCSEE standards may improve teacher perceptions toward the evaluations. Vekeman et al. (2015) reported teachers' perceptions were more favorable in schools where expectations of the evaluation were both formative and summative regarding the implementation of the new teacher evaluation policy in their school.

Results from the Emerald County School District study had similarities with an Austin Independent School District (2012) study of elementary teachers who participated in a study using the SLO process. The results showed that the elementary teachers were more likely than were comparison teachers to engage in discussion concerning their needs and objectives for professional development as well as dialog about individual and group assessment data for students derived from the SLOs. However, the findings of the current study diverged from the findings of the Austin Independent School District study, as well as from Hopkins (2013). Results from the Austin Independent School District study suggested no significant differences between participants at the middle or high school levels. Moreover, Hopkins concluded that the level of school where a teacher taught did not account for any significant differences among responses on any of the four evaluation standards.

Research Question 3

Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the evaluation system as implemented to date?

Findings indicated that in every measure of the JCSEE attributes, the administrators' mean levels of agreement were significantly higher than the teachers'

mean levels of agreement on the total attribute scales, those items measuring student learning objectives, and those items measuring professional practices. Those findings support Hopkins's (2013) research, which found that teachers did not perceive the use of student performance data as positively affecting the propriety standard in their evaluation.

However, the Emerald County School District study is, contrary to other research, suggesting teachers and principals not only have positive perceptions of new evaluations programs, but also that principals and teachers find the new evaluation processes enhance conversations around instruction and reflection on practice (Sartain et al., 2011; Tuytens & Devos, 2009). Winslow's (2015) study of Illinois schools using the Danielson framework for teaching found that both teachers and principals had high levels of agreement toward improvement in methods of the new evaluation system as compared to the methods of the former system. In Winslow's study, responders also indicated levels of agreement toward more meaningful and timely feedback identifying instructional practices.

Conclusions

It is important to note that this current study, conducted in the second year of the new teacher evaluation implementation, is a perceptual study, examining the perceptions of the teachers and administrators regarding the evaluation program to date. This perceptual study may contribute to district leaders understanding the principal and teacher perceptions toward new evaluation programs for teachers. The intent of this study was not to evaluate or measure the merit of new evaluation program but rather provide insight into the perceptions of those most affected by the new evaluation program. Similar studies of perceptions found teachers feared implementation of solely summative

evaluations, feared that new evaluation systems would contribute to control and sanctions, and were opposed to the implementation of evaluation programs (Flores, 2012; Morgado & Sousa 2010; Stronge & Tucker, 1999; Tuytens & Devos, 2009). Considering the teachers and principals' perceptions of the new evaluation program can help district leaders in the current study avoid or diminish resistance and promote and enhance acceptance and compliance of the new evaluation program.

Understanding teacher and principal perceptions regarding the new evaluation program for teachers provides relevant information for designing, implementing, and maintaining effective teacher evaluation practices. Moreover, understanding how both teachers and principals perceive an evaluation process that incorporates both the Danielson framework and SLOs aligns with the JCSEE standards can help district leaders achieve greater success with evaluation practices. Through this understanding, the district can develop its greatest assets, the teachers, to improve learning.

Successful implementation of teacher evaluation programs using the Danielson framework and SLOs entails more than professional development and resources. More often than not school districts fail to consider human dynamics when implementing new programs (Bransford & Donovan, 2005; Mielke & Frontier, 2012). Teachers respond to the demands of new evaluation programs with unique emotions, attitudes, and perceptions. A teacher evaluation system supported by the teachers has the greatest potential to improve teacher practices and, ultimately, to improve student learning.

Of concern to the teachers and principals in the present study was the accuracy of the new evaluation program. Concerns regarding high-stakes evaluations are certain to promote fear and anxiety, especially from those most affected by the new evaluation

program. While administrators consistently tended more toward *agree* in every item than did the teachers, no group met the 3.0 of *agree* on the scale. Policy change often polarizes stakeholders (Behrstock-Sherrat et al., 2013). Too often changes in policies and reform are placed on the implementation fast track and neglect the process of authentic engagement. Authentic engagement requires a:

> thoughtfully designed and well facilitated process which ensures that teachers have a seat at the table when decisions are being discussed…it involves much more than a few public hearings, an occasional survey, and cursory lip service to the attitudes of teachers…it is not a one-time event…throughout the process there are opportunities for all teachers to have a forum to share their ideas and concerns; and to know that they were genuinely considered when decisions were made. (Behrstock-Sherrat et al., 2013, p. 83)

The process of authentic engagement is time consuming. Unexpected criticism may be viewed as venting or counterproductive. However, overlooking the apprehensions of teachers regarding new evaluation programs for teachers and changes in policy may cause innovations to fail (van den Berg, 2002); therefore, it is important to not only understand, but acknowledge teachers' perception of educational policy (Datnow & Castellano, 2000). Rolling out new evaluation policies must not only be well designed and developed, but also must provide adequate time, resources, and ongoing opportunities for all parties affected by the new evaluation to authentically engage, grapple, debate, negotiate, review, correct, and modify, the process and the tool for successfully implementing and sustaining the program. Even when the new evaluation policies are in process, districts such as Emerald County can benefit from using authentic

engagement for mid-course reflections to improve the teachers and administrators' perceptions of the evaluation program.

Other reasons that responders in this study were concerned with accuracy may be due to the district deviating from Danielson's (1996) original intent of her framework for teaching as a formative tool to improve instruction. Even though adaptations of Danielson's framework show a relationship to outcomes such as student achievement in various research studies, the effects are modest and varied across settings (Kimball et al., 2006; Milanowski, 2004). Little, Goe, and Bell (2009) considered those differences a result of various modifications of Danielson's original framework for teaching. Moreover, research by Sartain et al. (2011) of principal observations of teaching practice conducted twice a year using the Danielson framework for teaching, found differences in principal and observer ratings influenced by a teacher's previous evaluation rating. A quantitative analysis revealed that the ratings principals assigned to teachers on previous evaluations contributed to current evaluation results, suggesting that principals may have considered previous evaluation ratings when assigning new ratings. Additionally, teachers and principals may not see the value and validity of using SLOs to improve teaching practices to increase learning. Emerald County School District may benefit from incorporating the JCSEE standards more clearly into their current evaluative procedures.

Without a well-developed strategic plan, the assumption cannot be made that stakeholders will acquiesce when a new policy for teacher evaluation is introduced. The strategic plan must include training principals and teachers how to use data effectively to achieve stated outcomes. Enhancing the teachers and principals' understanding of the value and usefulness of data can eliminate misconceptions about its use in teacher

evaluation programs. Positive perceptions from stakeholders can be achieved when purposeful collaborating on the development of goals and objectives for using data occurs (Behrstock-Sherrat et al., 2013).

The levels of accountability and responsibility between teachers and building level principals are varied; therefore, teachers and principals use school data in different ways. For principals to be effective, they need to use data to inform themselves of both student learning and teacher effectiveness, to navigate their course for leadership. Teachers need to understand and use school data to inform their instructional practices and to improve student learning. Both teachers and principals can benefit by acquiring skills to improve their data literacy. Teachers and principals must be able to not only understand and select what data are needed, but also have in-depth knowledge and understanding of how to use the data to inform instructional practices to increase student learning.

Consistency is crucial for developing data literate environments. While some building-level leaders may allocate time and resources for improving data literacy of their staff, district leaders should consider creating a comprehensive plan for data use and assessment literacy in all schools. A data-literate environment should allow not only time for teachers and principals to collaborate, but also provide them the technical support needed to bring clarity about how and why data are used.

The respondents in the Emerald County School District expressed concerns about the accuracy of the new evaluation program. Many teachers are initially skeptical of using student growth measures in summative teacher evaluation programs, but with collaboration and transparency, teachers become more accepting of using student growth

scores to measure teacher effectiveness (Behrstock-Sherratt et al., 2013). Teachers and principals need to know and understand how the district creates, monitors, reviews, compares, and analyzes assessment data derived from the SLOs. Providing teachers and principals with an understanding of how the SLO assessments are valid and reliable measures by regularly analyzing and comparing them with teacher observation ratings and other measures that predict future student success can improve perceptions of the merit and value of the assessments and decrease skepticism. Teachers need to see how the SLO data will improve all performance measures over time. The district should also increase teachers' positive perceptions of the evaluation program by clearly declaring the evaluation results to be a guide for promoting professional learning for all teachers.

It is worth noting the mean scores for both the principals and teachers in in the study fell between the 2.0 of *disagree* and the 3.0 of *agree*. As a group, neither teachers nor principals fully agreed nor fully disagreed, perhaps indicating ambivalence toward the evaluation program. These perceptions may be a result of the limited time allotted for rolling out the new evaluation program. Additionally, both the professional practices and SLO components were concurrently implemented giving rise to concerns about the quality and depth of the program. Districts should not underestimate the value of explaining the underlying assumptions of the framework and SLOs, in addition to the purpose and procedural aspects of the new observation process to both administrators and teachers. Without continuous professional development opportunities to improve the attributes described in the framework and SLOs, the teachers will underestimate the potential of both the Danielson framework and the SLOs for promoting their professional growth. The focus of a mid-course correction should include integrating both the

professional practices and SLO components of the new evaluation system with the JCSEE. Continued implementation efforts should concentrate on the quality and consistency of the new evaluation program's capacity to strength classroom practices and improve teaching rather than approaching implementation of the components as disjointed and piecemeal.

Tuytens and Devos (2009) found that teachers grew professionally because of the positive perceptions of their evaluation experiences. More importantly, because of its connection to the Danielson framework, the district should consider the results from research conducted by Jiang, Sporte, and Luppescu (2015) concerning the perceptions and experiences of teachers and administrators during the first year of Recognizing Educators Advancing Chicago implementation. The evaluation program included an observation tool adapted from the Danielson framework for teaching. Jiang et al. found that school administrators and teachers expressed positive views of the potential of the teacher practice component to support teacher growth and professional development.

Implications

Emerald County School District is in the third year of implementing the new evaluation program for teachers; yet the perceptions of teachers and administrators responding to the study reveal concerns regarding the accuracy, utility, and feasibility of the evaluation program. Zimmerman and Deckert-Pelton (2003) reported that teachers consistently expressed a desire for "reciprocal, communicative relationships with their evaluators" (p. 32) and a need for constructive feedback on their individual strengths and weaknesses. It appears that the responding teachers in this study slightly favored using the professional practices component more than the SLO component of the evaluation.

Using SLO data in teacher evaluation is a relatively new and unknown dynamic for the teachers in this study; therefore, school administrators need to understand how teachers perceive this change as it relates to teacher support of the changes to the teacher evaluation process (Bryk, Camburn, & Louis, 1999; Schneider & Bryk, 2000; Turnbull, 2002). Providing teachers opportunities to understand how student performance data can support and balance other areas, such as the professional practices component of the evaluation program, may garner and sustain support from all stakeholders.

Although principals in the Emerald County School District study agreed more than teachers did on many of the items in the questionnaire, the impact of changes in new policies, particularly a new evaluation program, cannot be ignored. The time principals extend in conducting new evaluation programs often prevents them from enhancing and engaging in supports that increase teaching performance and student achievement. Principals are expected to be adept at facilitating change. Effective principals are not only skillful in their practices to bring about change, but they also promote and nurture programs that encourage professional staff development for improved learning outcomes (National Governors Association, 2008). Simultaneously, the effects of the change process take a personal toll on principals. Implementation of new policies and programs often require not only new learning along with new school practices, but also a shift in paradigms related to novel policies and procedures that are externally mandated (Derrington & Campbell, 2015) . These externally mandated new accountability policies and procedures can cause principals to experience feelings of frustration, inadequacy, and disorder; therefore, it important to provide interventions that address principal concerns for successful implementation of an evaluation program (Hall & Hord, 2015).

Many leaders can positively influence student growth even if instructional practices remain unchanged (Supovitz, Sirinides, & May 2010; Witziers, Bosker, & Kruger, 2003). Leithwood and Jantzi (2005) contended that principals could enhance student achievement by providing powerful visions, a strong academic mission, robust organizational goals, and high expectations. Principals need opportunities to facilitate instructional quality by supporting student opportunities to learn (Harris & Herrington, 2006). They need to work with staff in developing and using data systems to inform and monitor decisions (Lachat & Smith, 2005). They must develop school cultures that encourage learning through aligning school actions with the vision and mission to ensure all students not only participate, but have ownership in the school (Bryk, Sebring, Allensworth, Easton, & Luppescu, 2010; Sebastian & Allensworth, 2012). Lastly, they must provide alignment and cohesiveness to all school actions.

Charalambous, Komitis, Papacharalambous, and Stefanou (2014) delineated the importance of teachers' perceptions toward validating specific criteria in new teacher evaluation programs. Charalambous et al. found that teachers' perceptions of empowerment increased when they had a voice in the process, and implementation was improved when districts asked teachers to identify their concerns in implementing the new evaluation criteria and instructional practices in their teaching. Charalambous et al. emphasized the influence of perceptions when implementing a new evaluation programs.

If districts are to safeguard the fidelity, implementation, and sustainability of new evaluation programs, districts must acknowledge the influence that teacher perceptions have on endorsing implementation efforts toward change. Teachers' perceptions toward adjusting instructional practices to align with the standards and criteria of new evaluation

programs can either hinder or ensure the program implementation. Perhaps teachers in this study require more time and training to gain a theoretical understanding of the Danielson framework and the SLOs, as well as the evaluation program's tools and rubric. Most importantly, though, the teachers and principals need the opportunity to discuss effective teaching through the lens of student learning.

Donaldson's (2012) findings regarding how a new evaluation system affected how teachers planned for their lessons indicated that the new evaluation system had no direct effect on teachers' pedagogy. Considering that both teachers and principals' scores in the current study did not meet the ratings for either *disagree* or *agree*, thereby indicating possible ambivalence coupled with the pressures of increased accountability for teachers and principals, districts will need to find ways to support teachers taking risks with different instructional strategies and pioneering ideas to foster student learning.

Teachers less experienced with an evaluation system may perceive the new evaluation program less positively due to the lack of understanding of the purpose, expectations, and worth of the program (Hopkins, 2013). If those who are most affected by the new evaluation program do not understand or embrace the new program, they may be less likely to identify and eventually apply the evaluation programs merit for improving professional growth and development. Sartain et al. (2011) acknowledged the need for depth and quality of training for both teachers and administrators in order for proper implementation.

Recommendations for Practice

Teachers tended to *disagree* and administrators tended to *agree* that the Emerald County School District's evaluation system met the attributes of the JCSEE. Both

teachers and administrators' levels of agreement were highest on the propriety and lowest on the accuracy of the evaluation system. Teachers were less likely than were administrators to agree to all items measuring the JCSEE attributes. Moreover, there was more agreement from both groups that the use of professional practices data was more in line with the JCSEE standards than using student learning objectives data.

This is not surprising, as evaluations having limited alignment with the JCSEE standards may produce negative perceptions from the evaluatees. Vekeman et al. (2015) found that most teachers initially fear that the new evaluation programs would be solely summative in rating teacher job performance. Tuytens and Devos (2009) found that teachers expressed concerns regarding how schools implemented the new teacher evaluation policy even though the teachers expressed perceptions that were positive toward the new teacher evaluation policy. Reflecting on the current implementation practices with greater consideration for providing teachers and principals opportunities for on-going authentic engagement is strongly recommended. It is recommended that district leaders focus on instructional quality in their efforts to implement that new teacher evaluation program. This can be accomplished by providing key instructional connections that strength the quality of the SLOs and the professional practices.

Other research found that teachers having summative expectations regarding the new evaluation program resulted in their increased fear that the new evaluation systems would result in greater control over teachers and more sanctions (Flores, 2012; Morgado & Sousa, 2010; Stronge & Tucker, 1999). However, evaluations that have a greater alignment with the JCSEE standards may improve teacher perceptions toward the evaluations. Vekeman et al. (2015) reported teachers' perceptions were more favorable in

schools where expectations of the evaluation were both formative and summative regarding the implementation of the new teacher evaluation policy in their school. Therefore, it is recommended that the school district consider conducting an audit or assessment for applying the JCSEE standards to the current evaluation program.

The teachers' perceptions of the JCSEE standards of propriety, utility, feasibility, and accuracy of the district's new teacher evaluation system were significantly lower than the administrators' perceptions. While teachers did not display outright resistance, the analysis did not find teachers leaning toward agreement in their perceptions of the evaluation system. There is ample research regarding teachers' perceptions toward teacher evaluation systems suggesting that teachers may not choose to support some teacher evaluation systems (Peterson, 2000) due the inherent emotionally and politically laden challenges of designing and implementing teacher evaluation systems (Stronge & Tucker, 1999).

Teachers or subgroups of teachers may display resistance toward new evaluation programs (Monyatsi, Steyn, & Kamper, 2006). This resistance can be attributed to lack of communication about the evaluation (Heneman et al., 2006), lack of collaboration in designing the system, or lack of organizational commitment on the part of local educational leadership (Stronge & Tucker, 2003). Teacher and administrator buy-in is critical to implementing and sustaining a new evaluation program for teachers. Research literature shows that when teachers accept and respond positively to evaluation systems they take optimal advantage of the systems to improve teaching practice (Donaldson, 2012; Mielke & Frontier, 2012).

Teachers' perceptions of any evaluation process are derived from their experience with evaluation. These perceptions have the capacity to influence the climate and quality of instructional practices in their classroom. Teachers will do what they perceive is best to serve their students and themselves (Donaldson, 2012; Mielke & Frontier, 2012). In today's educational climate teachers perceive evaluation as a shallow, sporadic event that is detached from their daily classroom teaching and learning. Teachers perceive their roles as passive recipients of external judgment (Mielke & Frontier, 2012). Teachers have also expressed concerns that evaluations do nothing to help them improve their teaching practices due to the lack of feedback from school leaders (Danielson & McGreal, 2000). In light of the extant research, it is recommended that the district consider providing both teachers and principals deeper clarity and a practical understanding of how the professional practice and SLO components of the evaluation program can influence and improve their instructional practices.

Helping teachers build capacity for improved assessment literacy specifically relating to assessing SLO growth would improve perceptions of the validity and reliability of SLOs. Recommendations for the district relating to creating valid SLO assessments should include recommendations from Herman, Heritage, and Goldschmidt (2011) and Gareis and Grant (2015). Herman et al. (2011) provided five essential elements for valid and reliable assessments of SLOs: (a) the standards clearly define what students are expected to learn, (b) the assessment instruments are designed to accurately and fairly address what students are expected to learn, (c) Student assessment scores accurately and fairly measure what students have learned, (d) student assessment scores accurately and fairly measure student growth, and (e) student growth scores (based on the

assessments) can be accurately and fairly attributed to the contributions of individual teachers. The following checklist for valid and reliable SLO assessments is based on research by Gareis and Grant (2015).

Alignment to standards. Is the learning object clearly reflected in the assessment measurement?

- All items in the assessment align to the standard(s) addressed in the SLO.
- The assessment tool addresses the full range of topics and skills included in the SLO.
- The focus of the assessment mirrors the focus of the curriculum and standards.
- The items or task match the full range of cognitive thinking required during the course.
- The assessment requires students to engage in higher-order thinking where appropriate.

Stretch. Will all students be able to demonstrate growth on this assessment?

- The test includes items that cover prerequisite knowledge and skills from prior years and appropriate, content-relevant items that will challenge the highest performing students.
- Test items cover knowledge and skills that will be of value beyond the school year.

Validity and reliability. Is the assessment measure a valid and reliable tool for the intended purpose?

- The assessment does not include overly complex vocabulary.

- Items or tasks are written clearly and concisely.
- Clear scoring rubrics or guidance exists for open-ended questions or performance-based assessments.
- The teacher has a plan for administering assessments consistently across classes.

Moreover, helping teachers and principals develop a more practical and meaningful understanding for aligning the evaluation components and the JCSEE standards with strategies for reaching success, may enhance teachers' perceptions of the new evaluation program. The district can include supports for improving the principals' capacity for viewing the evaluation program as pragmatic and meaningful for improving teaching practices.

Principals in turn should create a culture in their schools that enables teachers to not only view the new evaluation program as a fundamental part of the school system's mission to improve instructional practices, but also provide teachers with innovative strategies that use both SLOs and Danielson's framework in planning and practice. Additionally, the district should consider establishing engagement teams of stakeholders to review the degree to which the district's curriculum, assessments, and instructional strategies align with the philosophical underpinnings of the JCSEE standards, the Danielson framework, and SLOs. Subsequently, these engagement teams would collaborate with schools to develop strategies that align with the district's vision for sustaining a successful evaluation program (Behrstock-Sherrat et al., 2013).

Both principals and teachers responding to the study showed a tendency toward *disagreeing* with the SLO evaluation component while tending slightly more toward

agreeing with the professional practice components in all areas of the JCSEE standards. The JCSEE personnel evaluation standards stress the importance of evaluation programs providing defensible performance decisions that are valid and reliable. The research regarding the psychometric value in using SLOs to measure teacher performance in high-stakes evaluation is limited.

After examining seven research studies, Gill et al. (2013) concluded there was limited evidence of the statistical properties for using SLOs to measure student and teacher performance in high-stakes evaluation and compensation programs. Furthermore, Gill et al. questioned the ability of SLOs to discriminate accurately among the differences in teacher performance. Although Gill et al. found that SLOs show more promise than previous evaluation metrics to better distinguish teachers based on performance, research is limited on the reliability of the SLOs to yield ratings that correlate with other measures of teacher performance. Therefore, it is recommended that school districts consider using SLOs primarily for instructional planning prior to incorporating them into high-stakes teacher evaluations, until research provides more evidence relating to the statistical properties for using SLOs to measure student and teacher performance. Because SLOs are tailored to individual teachers and constructed on the professional judgments of teachers and principals, creating a valid and reliable SLO is difficult. Therefore, districts incorporating SLOs into their evaluation program would benefit in providing teachers and administrators with not only extensive training time and continuous support in the creation and instructional application of SLOs, but also with resources and training for reducing the time demands required to create SLOs.

Recommendations for Policy

Although the district is in the early stages of developing and modifying both policies and tools for the new evaluation system for teachers, the district can continue to build their policy framework for sustaining the evaluation program. The district can also use evaluation data in future policies to expand and improve instructional practices that will increase, in turn, student achievement. Now that the district is moving further along in implementation practices for the new evaluation program, the district should address various emerging challenges that are inevitable with change and school reform.

Not only should the district provide continuous communication and feedback that is essential to the success of any reform efforts, the district must also establish robust monitoring systems and feedback mechanisms for identifying the strengths and weaknesses of the new teacher evaluation policies. Doing so will provide the district with the capacity to make informed decisions about ways to improve the effectiveness and avoid low-fidelity implementation. Research, such as this present study, that examines and provides specific feedback for evaluation models using SLOs and the Danielson framework can be used and reviewed by policymakers.

In as much as districts have adopted standards for student learning and the professional practices of both teachers and administrators, developing policies that support significant efforts to implement and enforce standards for professional development will further the quality of teaching and learning in the district. The implications of the district's policies regarding the use of effective evaluation to identify ineffective teaching practices are significant and profound. If implemented well, and if school administrations act on the results, district policies can be further developed that

consider actions for struggling teachers and for the development of fair but rigorous policies for addressing persistently ineffective teachers.

Teacher evaluation needs to be in tandem with individualized, rigorous, and concentrated professional development that provides opportunities for growth. The idea of drive-by, single-session, and whole-district workshops should be avoided. Professional development needs should be aligned with evaluation outcomes and individualized using technology. Non-traditional methods of professional development, wikis and/or blogs, online district courses, action research, and study groups should be considered.

Schools must make an effort to sustain the professional learning community model until it becomes deeply embedded in the culture of the school. Professional learning communities shift the focus of school reform from restructuring to re-culturing and engagement (Louis, 2007). Schools must offer more opportunities for such collaboration and engagement between teachers and stakeholders. Teachers and principals need opportunities to learn, engage, and share their voice.

Recommendations for Further Research

A deeper and richer understanding of the current situation in the Emerald County School District could be achieved using a qualitative study that includes individual interviews and focus groups. Interviews could provide additional insight into how teachers perceive the alignment of the evaluation components with JCSEE standards. Providing teachers opportunities to express their perceptions verbally negates the likelihood of error in interpreting the results of responses to survey questions. As this study was conducted in the early phase of a new state mandate, adequate time may not have been allotted, with respect to teachers' exposure and experience with the new

evaluation system. Repeating the study in a few years would determine if trends associated with this study change over time as both educators and supervisors gain more experience with the new evaluation process. Future research could address the inherent bias of teacher ratings based on the ability levels of the students.

Summary

Research over the past 20 years substantiates that teaching is foundational for improved educational outcomes. Because teaching matters tremendously, the evaluation of teachers matters. For evaluations to be effective, district leaders should understand teacher and principal perceptions of new evaluation programs.

APPENDIX A: PERMISSION TO USE SURVEYS

Paul <paulthopkins@cox.net>
Jan 19, 2015

Mrs. Finnegan,

I grant you permission to use the requested materials for your dissertation. My dissertation was published in the ProQuest database. If you have any questions, please email me at my new email address above. Best of luck.

Paul Hopkins
Sent from my iPhone.

From: Ruth Finnegan [mailto:rsfinnegan@email.wm.edu]
Sent: Saturday, January 17, 2015 3:40 PM
To: Paul Hopkins
Subject: Permission to Use and Adapt Research Materials

Dr. Hopkins,

I am currently a doctoral candidate at the College of William and Mary. My research efforts are directed toward both administrators and teachers perceptions of students growth data and professional practices in the teacher evaluations. Your research regarding teachers' perceptions on the use of student growth data in teacher evaluations came to my attention through Dr. Stronge. Your research was insightful and furthered my understanding of the complexities districts face in devising new evaluations for teachers. As I prepare a proposal for my doctoral dissertation, I am requesting permission to use and adapt the items listed below. I request permission to

- Adapt and use your Teacher Perception Survey: Teacher Perceptions to the Use of Student Performance Data in Teacher Evaluation as a data collection instrument.
- Use Table 1: Cronbach's Alpha Analyses for Four Evaluation Standards from your dissertation. I plan to insert the table in my dissertation to support a discussion on the reliability of the survey.

If you grant permission and I decide to continue on my current path for my dissertation, I will credit you appropriately. Additionally, please describe any further parameters for use. One more question. Has your dissertation been published? If it is still in publication, could you please provide me an unpublished copy so that I might properly cite page numbers of any direct quotes from your research?

Thank you for your consideration,
R. Shannon Finnegan

Stronge, James H <jhstro@wm.edu>
11/9/14

Dear Shannon,

I owe the intellectual property rights for the SABJE teacher evaluation survey. Please accept this email as permission to adapt and use the survey in your dissertation at the College of William and Mary.

Best wishes,
James Stronge

APPENDIX B: QUESTIONNAIRE

	Strongly agree	Agree	Disagree	Strongly disagree
1. I believe I have a clear understanding of the expectations of the teacher's job performance in the new teacher evaluation system.	1	2	3	4
2. I believe the county provides clear and concise documentation of procedures and guidelines outlining the policies and procedures for the new teacher evaluation system.	1	2	3	4
3. I believe using Student Learning Objectives data in the new teacher evaluation system encourages professional discussion during follow-up conferences.	1	2	3	4
4. I believe using Professional Practice data in the new teacher evaluation encourages professional discussion during follow-up conferences.	1	2	3	4
5. I believe using Student Learning Objectives data in the new teacher evaluation system documents teachers' areas of strengths, as well as areas for improvement.	1	2	3	4
6. I believe using Professional Practice data in the new teacher evaluation documents teacher's areas of strengths, as well as areas for improvement.	1	2	3	4
7. I believe using Student Learning Objectives data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.	1	2	3	4
8. I believe using Professional Practice data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.	1	2	3	4
9. I believe the administrators/ evaluators implementing the use of Student Learning Objectives data in the new evaluation system for teachers are qualified to evaluate this component.	1	2	3	4
10. I believe the administrators/ evaluators implementing the use of Professional Practices data in the new evaluation system for teachers are qualified to evaluate this component.	1	2	3	4
11. I believe the criteria for using Student Learning Objectives in rating teacher performance is clear and accurate.	1	2	3	4
12. I believe the criteria for using Professional Practices in rating teacher performance is clear and accurate.	1	2	3	4
13. I believe using Student Learning Objectives data in the new evaluation system for teachers informs administrators/evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	1	2	3	4

	Strongly agree	Agree	Disagree	Strongly disagree
14. I believe using Professional Practice data in the new evaluation system for teachers informs administrators/evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	1	2	3	4
15. I believe using Student Learning Objectives data in the new evaluation system for teachers is a responsible use of student assessment data.	1	2	3	4
16. I believe using the Student Learning Objectives data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/evaluators and the teacher.	1	2	3	4
17. I believe using the Professional Practices data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/evaluators and the teacher.	1	2	3	4
18. I believe the time required of teachers for employing Student Learning Objectives data in the new teacher evaluation system is feasible.	1	2	3	4
19. I believe the time required of teachers for employing Professional Practices data in the new teacher evaluation system is feasible.	1	2	3	4
20. I believe using Student Learning Objectives data in the new evaluation system for teachers accurately contributes to evaluating my teaching.	1	2	3	4
21. I believe using Professional Practice data in the new evaluation system for teachers accurately contributes to evaluating my teaching.	1	2	3	4
22. I believe using Student Learning Objectives data in the new evaluation system for teachers will make my evaluation more objective.	1	2	3	4
23. I believe using Professional Practice data in the new evaluation system for teachers will make my evaluation more objective.	1	2	3	4
24. I believe using Student Learning Objectives data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	1	2	3	4
25. I believe using Professional Practices data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	1	2	3	4
26. I believe the use of Student Learning Objectives data in the new evaluation system for teachers helps administrators identify low-performing/ineffective teachers.	1	2	3	4
27. I believe the use of Professional Practices data in the new evaluation system for teachers helps administrators identify low-performing/ineffective teachers.	1	2	3	4

What is your age? _____

What is your gender?

_____ Female
_____ Male

Which race/ethnicity best describes you? (***Please choose only one***.)

_____ Black/African American
_____ Hispanic American
_____ White/Caucasian
_____ Other

What is your highest degree earned?

_____ Bachelor's
_____ Master's
_____ Specialist
_____ Doctorate

How many years of teaching experience do you have? ***Count this year as 1 year***. _____

At what level do you teach?

_____ PreK – Elementary
_____ Middle
_____ High

APPENDIX C. CORRESPONDENCE OF QUESTIONNAIRE ITEMS TO SURVEYS BY HOPKINS (2013) AND STRONGE (2013)

Items measuring propriety

	Previous studies	
Items on study questionnaire	Hopkins (2013)	Stronge (2013)
1. I believe I have a clear understanding of the expectations of the teacher's job performance in the new teacher evaluation system.		The handbook tells me what I need to know to prepare for the evaluation process.
2. I believe the county provides clear and concise documentation of procedures and guidelines outlining the policies and procedures for the new teacher evaluation system.		The handbook is clearly written and easy to understand.
3. I believe using Student Learning Objectives data in the new teacher evaluation system encourages professional discussion during follow-up conferences.		The new observation forms promote dialogue during follow-up conferences.
4. I believe using Professional Practice data in the new teacher evaluation encourages professional discussion during follow-up conferences.		The new observation forms promote dialogue during follow-up conferences.
5. I believe using Student Learning Objectives data in the new teacher evaluation system documents teachers' areas of strengths, as well as areas for improvement.	I believe the use of student performance data will more accurately document my strengths and weaknesses as a teacher.	
6. I believe using Professional Practice data in the new teacher evaluation documents teacher's areas of strengths, as well as areas for improvement.	I believe the use of student performance data will more accurately document my strengths and weaknesses as a teacher.	

Items measuring utility

	Previous studies	
Items on study questionnaire	Hopkins (2013)	Stronge (2013)
7. I believe using Student Learning Objectives data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.		The teacher performance standards, indicators, and rubrics will improve teaching and learning by providing a mechanism for specific evidence-based feedback.
8. I believe using Professional Practice data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.		The teacher performance standards, indicators, and rubrics will improve teaching and learning by providing a mechanism for specific evidence-based feedback.
9. I believe the administrators/ evaluators implementing the use of Student Learning Objectives data in the new evaluation system for teachers are qualified to evaluate this component.		
10. I believe the administrators/ evaluators implementing the use of Professional Practices data in the new evaluation system for teachers are qualified to evaluate this component.		
11. I believe the criteria for using Student Learning Objectives in rating teacher performance is clear and accurate.		The new teacher evaluation system clearly indicates what teachers are expected to do.
12. I believe the criteria for using Professional Practices in rating teacher performance is clear and accurate.		The new teacher evaluation system clearly indicates what teachers are expected to do.
13. I believe using Student Learning Objectives data in the new evaluation system for teachers informs administrators/ evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	I believe the use of student performance data in my evaluation will provide my evaluator/ administrator with sufficient information to suggest meaningful content-specific professional development activities for me.	

Items on study questionnaire	Previous studies	
	Hopkins (2013)	Stronge (2013)
14. I believe using Professional Practice data in the new evaluation system for teachers informs administrators/ evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	I believe the use of student performance data in my evaluation will provide my evaluator/ administrator with sufficient information to suggest meaningful content-specific professional development activities for me.	

Items measuring feasibility

Items on study questionnaire	Previous studies	
	Hopkins (2013)	Stronge (2013)
15. I believe using Student Learning Objectives data in the new evaluation system for teachers is a responsible use of student assessment data.	I believe the use of student performance data as one performance standard in my evaluation is a responsible use of student assessment data.	
16. I believe using the Student Learning Objectives data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/ evaluators and the teacher.		The new teacher evaluation system promotes two-way communication between the evaluator and the teacher.
17. I believe using the Professional Practices data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/evaluators and the teacher.		The new teacher evaluation system promotes two-way communication between the evaluator and the teacher.
18. I believe the time required of teachers for employing Student Learning Objectives data in the new teacher evaluation system is feasible.		The time required to implement the new teacher evaluation system is feasible.
19. I believe the time required of teachers for employing Professional Practices data in the new teacher evaluation system is feasible.		The time required to implement the new teacher evaluation system is feasible.

Items measuring accuracy

	Previous studies	
Items on study questionnaire	Hopkins (2013)	Stronge (2013)
20. I believe using Student Learning Objectives data in the new evaluation system for teachers accurately contributes to evaluating my teaching.		The teacher performance standards, indicators, and rubrics provide a meaningful and accurate measure of teacher performance.
21. I believe using Professional Practice data in the new evaluation system for teachers accurately contributes to evaluating my teaching.		The teacher performance standards, indicators, and rubrics provide a meaningful and accurate measure of teacher performance.
22. I believe using Student Learning Objectives data in the new evaluation system for teachers will make my evaluation more objective.	I believe the use of student performance data in my evaluation will make my evaluation more objective.	
23. I believe using Professional Practice data in the new evaluation system for teachers will make my evaluation more objective.	I believe the use of student performance data in my evaluation will make my evaluation more objective.	
24. I believe using Student Learning Objectives data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	I believe the use of student performance data in my evaluation will direct my attention to potential achievement gaps for students in my classroom.	
25. I believe using Professional Practices data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	I believe the use of student performance data in my evaluation will direct my attention to potential achievement gaps for students in my classroom.	
26. I believe the use of Student Learning Objectives data in the new evaluation system for teachers helps administrators identify low-performing/ineffective teachers.	I believe the use of student performance data in teacher evaluations will help administrators identify low-performing/ineffective teachers.	
27. I believe the use of Professional Practices data in the new evaluation system for teachers helps administrators identify low-performing/ineffective teachers.	I believe the use of student performance data in teacher evaluations will help administrators identify low-performing/ineffective teachers.	

APPENDIX D: ALIGNMENT OF RESEARCH QUESTIONS, SURVEY QUESTIONS, AND JCSEE PERSONNEL EVALUATION STANDARDS

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date? a. What are the perceptions of the propriety of the evaluation system as implemented to date? 2. Are there differences between teachers at different levels in their perceptions regarding the propriety of the evaluation system as implemented to date? 3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the propriety of the evaluation system as implemented to date?	1. I believe I have a clear understanding of the expectations of the teacher's job performance in the new teacher evaluation system.	Propriety	P1 Service Orientation- Personnel evaluations should promote sound education, fulfillment of institutional missions, and effective performance of job responsibilities, so that the educational needs of students, community, and society are met
	2. I believe the county provides clear and concise documentation of procedures and guidelines outlining the policies and procedures for the new teacher evaluation system.	Propriety	P2 Appropriate Policies and Procedures- Guidelines for personnel evaluations should be recorded and provided to the evaluatee in policy statements, negotiated agreements, and/or personnel evaluation manuals, so that evaluations are consistent, equitable, and fair.

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
	3. I believe using Student Learning Objectives data in the new teacher evaluation system encourages professional discussion during follow-up conferences.	Propriety	P4 Interactions with Evaluatees-The evaluator should respect human dignity and act in a professional, considerate, and courteous manner, so that the evaluatee's self-esteem, motivation, professional reputations, performance, and attitude toward personnel evaluation are enhanced or, at least, not needlessly damaged
	4. I believe using Professional Practice data in the new teacher evaluation encourages professional discussion during follow-up conferences.	Propriety	P4 Interactions with Evaluatees-The evaluator should respect human dignity and act in a professional, considerate, and courteous manner, so that the evaluatee's self-esteem, motivation, professional reputations, performance, and attitude toward personnel evaluation are enhanced or, at least, not needlessly damaged
	5. I believe using Student Learning Objectives data in the new teacher evaluation system documents teacher's areas of strengths, as well as areas for improvement.	Propriety	P5 Balanced Evaluation Personnel evaluations should provide information that identifies both strengths and weaknesses, so that strengths can be built upon and weaknesses addressed.
	6. I believe using Professional Practice data in the new teacher evaluation documents teacher's areas of strengths, as well as areas for improvement.	Propriety	P5 Balanced Evaluation Personnel evaluations should provide information that identifies both strengths and weaknesses, so that strengths can be built upon and weaknesses addressed.

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date? What are the perceptions of the utility of the evaluation system as implemented to date? 2. Are there differences between teachers at different levels in their perceptions regarding the utility of the evaluation system as implemented to date? 3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the utility of the evaluation system as implemented to date?	7. I believe using Student Learning Objectives data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.	Utility	U1 Constructive Orientation - Personnel evaluations should be constructive, so that they not only help institutions develop human resources but encourage and assist those evaluated to provide excellent services in accordance with the institution's mission statements and goals
	8. I believe using Professional Practice data in the new evaluation system for teachers improves teaching and learning in the classroom through explicit evidence-based feedback.	Utility	U1 Constructive Orientation - Personnel evaluations should be constructive, so that they not only help institutions develop human resources but encourage and assist those evaluated to provide excellent services in accordance with the institution's mission statements and goals
	9. I believe the administrators/evaluators implementing the use of Student Learning Objectives data in the new evaluation system for teachers are qualified to evaluate this component.	Utility	U3 Evaluator Qualifications- The evaluation system should be developed, implemented, and managed by persons with the necessary qualifications, skills, training, and authority, so that evaluation reports are properly conducted, respected and used.

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
	10. I believe the administrators/evaluators implementing the use of Professional Practices data in the new evaluation system for teachers are qualified to evaluate this component.	Utility	U3 Evaluator Qualifications- The evaluation system should be developed, implemented, and managed by persons with the necessary qualifications, skills, training, and authority, so that evaluation reports are properly conducted, respected and used.
	11. I believe the criteria for using Student Learning Objectives in rating teacher performance is clear and accurate.	Utility	U4 Explicit Criteria- Evaluators should identify and justify the criteria used to interpret and judge evaluatee performance, so that the basis for interpretation and judgment provide a clear and defensible rationale for results
	12. I believe the criteria for using Professional Practices in rating teacher performance is clear and accurate.	Utility	U4 Explicit Criteria- Evaluators should identify and justify the criteria used to interpret and judge evaluatee performance, so that the basis for interpretation and judgment provide a clear and defensible rationale for results
	13. I believe using Student Learning Objectives data in the new evaluation system for teachers informs administrators/evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	Utility	U6 Professional Development-Personnel evaluations should inform users and evaluatees of areas in need of professional development, so that all educational personnel can better address the institution's missions and goals, fulfill their roles and responsibilities, and meet the needs of students.
	14. I believe using Professional Practice data in the new evaluation system for teachers informs administrators/evaluators in recommending content-specific professional development activities for improving teacher's instructional practices.	Utility	U6 Professional Development- Personnel evaluations should inform users and evaluatees of areas in need of professional development, so that all educational personnel can better address the institution's missions and goals, fulfill their roles and responsibilities, and meet the needs of students.

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date? What are the perceptions of the feasibility of the evaluation system as implemented to date? 2. Are there differences between teachers at different levels in their perceptions regarding the feasibility of the evaluation system as implemented to date? 3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the feasibility of the evaluation system as implemented to date?	15. I believe using Student Learning Objectives data in the new evaluation system for teachers is a responsible use of student assessment data.	Feasibility	F1 Practical Procedures - Personnel evaluation procedures should be practical, so that they produce the needed information in efficient, non-disruptive ways
	16. I believe using the Professional Practice data in the new evaluation system for teachers is a responsible use of student assessment data.	Feasibility	F1 Practical Procedures - Personnel evaluation procedures should be practical, so that they produce the needed information in efficient, non-disruptive ways
	17. I believe using the Student Learning Objectives data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/evaluators and the teacher.	Feasibility	F2 Political Viability- Personnel evaluations should be planned and conducted with the anticipation of questions from evaluatees and others with a legitimate right to know, so that their questions can be addressed and their cooperation obtained
	18. I believe using the Professional Practices data in the new evaluation system for teachers provides opportunity for two-way communication between the administrators/evaluators and the teacher.	Feasibility	F2 Political Viability- Personnel evaluations should be planned and conducted with the anticipation of questions from evaluatees and others with a legitimate right to know, so that their questions can be addressed and their cooperation obtained

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
	19. I believe the time required of teachers for employing Student Learning Objectives data in the new teacher evaluation system is feasible.	Feasibility	F3 Fiscal Viability- Adequate time and resources should be provided for personnel evaluation activities, so that evaluation can be effectively implemented, the results fully communicated, and appropriate follow-up activities identified.
	20. I believe the time required of teachers for employing Professional Practices data in the new teacher evaluation system is feasible	Feasibility	F3 Fiscal Viability- Adequate time and resources should be provided for personnel evaluation activities, so that evaluation can be effectively implemented, the results fully communicated, and appropriate follow-up activities identified.
1. What are the perceptions of Emerald County School District teachers and school building administrators regarding the evaluation system as implemented to date? What are the perceptions of the accuracy of the evaluation system as implemented to date? 2. Are there differences between teachers at different levels in their perceptions regarding the accuracy of the evaluation system as implemented to date? 3. Are there differences in the perceptions of the Emerald County School District teachers and school administrators regarding the accuracy of the evaluation system as implemented to date?	21. I believe using Student Learning Objectives data in the new evaluation system for teachers accurately contributes to evaluating my teaching.	Accuracy	A1 Validity Orientation -The selection, development, and implementation of personnel evaluations should ensure that the interpretations made about the performance of the evaluatee are valid and not open to misinterpretation
	22. I believe using Professional Practice data in the new evaluation system for teachers accurately contributes to evaluating my teaching.	Accuracy	A1 Validity Orientation -The selection, development, and implementation of personnel evaluations should ensure that the interpretations made about the performance of the evaluatee are valid and not open to misinterpretation

Research question	Survey question	JCSEE standard	Justification/concepts for survey question
	23. I believe using Student Learning Objectives data in the new evaluation system for teachers will make my evaluation more objective	Accuracy	A6 Reliable Information- Personnel evaluation procedures should be chosen or developed and implemented to assure reliability, so that the information obtained will provide consistent indications of the evaluatee's performance
	24. I believe using Professional Practice data in the new evaluation system for teachers will make my evaluation more objective.	Accuracy	A6 Reliable Information- Personnel evaluation procedures should be chosen or developed and implemented to assure reliability, so that the information obtained will provide consistent indications of the evaluatee's performance
	25. I believe using Student Learning Objectives data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	Accuracy	A9 Analysis of Information- The information collected for personnel evaluations should be systematically and accurately analyzed, so that the purposes of the evaluation are effectively achieved
	26. I believe using Professional Practices data in the new evaluation system for teachers directs attention to potential achievement gaps for students in individual classrooms.	Accuracy	A9 Analysis of Information- The information collected for personnel evaluations should be systematically and accurately analyzed, so that the purposes of the evaluation are effectively achieved
	27. I believe the use of Student Learning Objectives data in the new evaluation system for teachers helps administrators identify low-performing/ineffective teachers.	Accuracy	A10 Justified Conclusions- The evaluative conclusions about the evaluatee's performance should be explicitly justified, so that evaluatees and others with a legitimate right to know can have confidence in them

APPENDIX E: CONSENT FORM

Thank you for agreeing to participate in the following survey. This informed consent outlines the facts, implications, and consequences of the research study. Upon reading, understanding, and signing this documentation, you are giving consent to participant in the research study.

Voluntary Nature of the Study
Your participation in this study is strictly voluntary. Your decision whether or not to participate will not affect your current or future relations with the researcher or the participating schools. If you initially decide to participate, you are still free to withdraw later without affecting those relationships.

Risks and Benefits of Being in the Study
No study is without risk. The risks are minimal, no more than the participant would encounter in everyday life. There are no risks associated with participating in this study and there are no short or long-term benefits. In the event you experience stress or anxiety during your participation in the study, you may terminate your participation at any time. You may refuse to answer any questions you consider invasive or stressful.

Confidentiality
The records of this study will be kept private and all subjects will remain unidentified and anonymous. I will take every precaution to protect participant identity by not linking survey information to participant identity. In any part of this study is published, the researcher will not include any information that will make it possible to identify schools and participants. The survey will be located on SurveyMonkey.com. Data stored by Survey Monkey is in a secure location protected by pass card and biometric recognition; it is conceivable that engineering staff at the web hosting company may need to access the database for maintenance reasons. The researcher will also store all research documentation on a protected computer database on her personal computer used for educational and university purposes that requires a secure password to access.

Contacts and Questions
I understand that should I have any questions about this research and its conduct, I should contact any of the following:

The researcher conducting this study is Shannon Finnegan [rsfinnegan@wm.email.edu]. You may ask any questions you have any via email. If you have additional questions later regarding the form and content of study, you are encouraged to contact the researcher's faculty advisor Dr. Tschannen-Moran [mxtsch@wm.edu]. If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher and advisor, you are encouraged to contact the Institutional Review Board, [Dr. Ray McCoy], Chair, [The College of William and Mary at 757-221-2783], or email at [rwmcco@wm.edu]

Electronic Signature

By clicking on the submit button to begin the survey, I am indicating that I have read the information provided and give my consent to be a participant in the research. I understand that when I complete the electronic survey, I am indicating that I have agreed to participate in this research project.

REFERENCES

Aaronson, D., Barrow, L., & Sanders, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95–135.

Acheson, K., & Gall, M. (2003). *Clinical supervision and teacher development: Preservice and inservice applications* (5th ed.). New York, NY: John Wiley & Sons.

Ahn, T. (2013). The missing link: Estimating the impact of incentives on teacher effort and instructional effectiveness using teacher accountability legislation data. *Journal of Human Capital, 7*(3), 230–273.

Alkin, M. (2004). *Evaluation roots*. Thousand Oaks, CA: SAGE.

Amrein-Beardsley, A. (2008). Methodological concerns about the Education Value-Added Assessment System (EVAAS). *Educational Researcher, 37*(2), 65–75.

Anast-May, L., Penick, D., Schroyer, R., & Howell, A. (2011). Teacher conferencing and feedback: Necessary but missing! *International Journal of Educational Leadership Preparation, 6*(2), 1–7.

Austin Independent School District. (2012). *AISD REACH Program Update, 2010–2011: Participant feedback* (Department of Research and Evaluation Report 10.86 RB). Austin, TX: Author. Retrieved from http://www.austinisd.org/sites/default/files/dre-reports/rb/10.86RB_AISD_Reach_Participant_Feedback_2010–2011_0.pdf

Auty, W., Bielawski, P., Deeter, T., Hirata, G., Hovanetz-Lassila, C., Rheim, J., & Williams, A. (2008). *Implementer's guide to growth models*. Washington, DC: Council of Chief State School Officers.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26.

Barry, R. A. (2010). *Teaching effectiveness and why it matters.* Retrieved from http://www.chalkboardproject.org/images/CB_TeachEffctPaper_A3.pdf

Bascia, N., & Hargreaves, A. (2000). Teaching and leading on the sharp edge of change. In N. Bascia & A. Hargreaves (Eds.), *The sharp edge of educational change* (pp. 131–155). New York, NY: Routledge Falmer.

Behrstock-Sherratt, E., Rizzolo, A., Laine, S., & Friedman, W. (2013). *Everyone at the table: Engaging teachers in evaluation reform.* San Franciso, CA: Jossey-Bass.

Blumer, H. (1969). *Symbolic interactions*. Englewood Cliffs, NJ: Prentice Hall.

Bonk, W., Copa, J., Gibson, N., Gillin, T., Nau, J., Peoples, A. L., Woolard, C. (2012). *Growth models issues and advice from the states*. Washington, DC: Institute of Education Sciences Statewide Longitudinal Data Systems Grant Program.

Borman, G. D., & Kimball, S. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elementary School Journal, 106*, 3–20.

Boykin, A. (2011). *Creating the opportunity to learn*. Alexandria, VA: ASCD.

Bransford, J. D., & Donovan, M. S. (2005). *How students learn: History, mathematics, and science in the classroom*. Washington, DC: National Academies Press.

Bridges, W., & Bridges, S. (2009). *Managing transitions: Making the most of change* (3rd ed.). Philadelphia, PA: Da Capo Press.

Bryk, A. S., Camburn, E., & Louis, K. S. (1999). Professional community in Chicago elementary schools: Facilitating factors and organizational consequences. *Educational Administration Quarterly, 35*(5), 751–781.

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Burns, S. F., Gardner, C. D., & Meeuwsen, J. (2009). *An evaluation of teacher and principal experiences during the pilot phase of AISD REACH: A strategic compensation initiative.* Nashville, TN: Peabody College at Vanderbilt University.

Castellano, K. E., & Ho, A. D. (2013). *A practitioner's guide to growth models*. Washington, DC: Council of Chief State School Officers.

Castillo, B. A. (2005). *The impact of teacher evaluation on teacher practice: A case study* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 180449)

Cavanagh, S. (2011). Statehouses in ferment over K–12. *Education Week, 30*(32), 1, 18–19.

Charalambous, C. Y., Komitis, A., Papacharalambous, M., & Stefanou, A. (2014). Using generic and content-specific teaching practices in teacher evaluation: An exploratory study of teachers' perceptions. *Teaching and Teacher Education, 41*, 22–33.

Chenoweth, K. (2010). *It's being done: Academic success in unexpected schools*. Cambridge, MA: Harvard Education Press.

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). *Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects.* Washington, DC: Urban Institute.

Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002, April). *Teacher evaluation: A review of literature*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Community Training and Assistance Center. (2013). *It's more than the money: Teacher incentive fund—Leadership for educators' advanced performance: Charlotte-Mecklenburg Schools*. Boston, MA: Author.

Conley, S., & Glasman, N. (2008). Fear, the school organization, and teacher evaluation. *Educational Policy, 22*, 63–85. doi:10.1177/0895904807311297

Conley, S., Muncy, D. E., & You, S. (2005). Standards-based evaluation and teacher career satisfaction: A structural equation modeling analysis. *Journal of Personnel Evaluation in Education, 18*, 39–65.

Connell, J., Kubisch, A., Schorr, L., & Weiss, C. (1995). *New approaches to evaluating community initiatives: Concepts, methods, and contexts*. Washington, DC: Aspen Institute.

Cousins, B., & Leithwood, K. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56*(3), 331–364.

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. New York, NY: Routledge.

Danielson, C. (1996). *A framework for teaching*. Alexandria, VA: ASCD.

Danielson, C. (2002). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: ASCD.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35–39.

Danielson, C. (2012). Observing classroom practice. *Educational Leadership, 70*(3), 32–37.

Danielson, C. (2013). *The framework for teaching: Evaluation instrument* (2013 ed.). Princeton, NJ: The Danielson Group.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice.* Alexandria, VA: ASCD.

Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, and practice for democratic education. *Educational Researcher, 25*(6), 5–17.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1), 6–14.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York, NY: Teachers College Press.

Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as part of a comprehensive system for teaching. *American Educator, 38*(1), 4–13, 44.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher, 31*(9), 13–25.

Datnow, A., & Castellano, M. (2000). Teacher's responses to success for all: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal*, 37, 775–799.

Davis, J. D., Chopin, J., Drake, C., & McDuffie, A. R. (2014). Factors underlying middle school mathematics teachers' perceptions about the CCSSM and the instructional environment. *Middle Grades Research Journal, 9*(3), 11–26.

Derrington, M. L., & Campbell, J. (2013). *The changing conditions of instructional leadership: Principals' perceptions of teacher evaluation accountability measures.* In B. G. Barnett, A. R. Shoho, & A. J. Bowers (Eds.), *School and district leadership in an era of accountability* (pp. 231–254). Charlotte, NC: Information Age.

Derrington, M. L., & Campbell, J. W. (2015). Implementing new teacher evaluation systems: Principals' concerns and supervisor support. *Journal of Educational Change, 16*(3), 305–326.

DiPaola, M., & Hoy, W. (2012). *Principals improving instruction: Supervision, evaluation, and professional development*. Charlotte, NC: Information Age.

DiPaola, M., & Hoy, W. (2014). *Improving instruction through supervision, evaluation, and professional development*. Charlotte, NC: Information Age.

Donaldson, M. (2012). *Teachers' perspectives on evaluation reform*. Washington, DC: Center for American Progress.

Drago-Severson, E., & Blum-DeStefano, J. (2014). Tell me so I can hear: A developmental approach to feedback and collaboration. *JSD, 35*(6), 16–22.

Duke, D., & Stiggins, R. (1986). *Five keys to growth through teacher evaluation.* Portland, OR: Northwest Regional Educational Laboratory.

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, *17*(1), 101–128.

Emery, K., & Ohanian, S. (2004). *Why is corporate America bashing our public schools?* Portsmouth, NH: Heinemann.

Engram, D. A. (2007). *Elementary teachers' perceptions of the Georgia Teacher Evaluation Program* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3253626)

Feeney, E. J. (2007). Quality feedback: The essential ingredient for teacher success. *The Clearing House*, *80*(4), 191–197.

Fink, D., & Stoll, L. (1996). *Changing our schools: Linking school effectiveness and school improvement.* Bristol, PA: Open University Press.

Flores, M. A. (2012). The implementation of a new policy on teacher appraisal in Portugal: How do teachers experience it at school? *Education Assessment, Evaluation and Accountability*, *24*(4), 351–368.

Fowler, F. C. (2009). *Policy studies for educational leaders* (3rd ed.). Boston, MA: Allyn & Bacon.

Frase, L. (2001). Constructive feedback on teaching is missing. *Education, 113*, 176–181.

Fullan, M. (2005*). Leadership and sustainability: Systems thinkers in action.* Thousand Oaks, CA: Corwin.

Fullan, M. (2008). *Six secrets of change*. San Francisco, CA: Jossey-Bass.

Fullan, M. (2011). *Leading in a culture of change.* San Francisco, CA: Jossey-Bass.

Gareis, C. R., & Grant, L. (2008). *Teacher-made assessments*. Larchmont, NY: Eye on Education.

Gates Foundation. (2013). *Measures of effective teaching (MET).* Retrieved from http://www.gatesfoundation.org/united-states/Pages/measures-of-effective-teaching-fact-sheet.aspx

Giliya, Z. (2006). *A study of teachers' perceptions of the evaluation process*. Retrieved from ProQuest Dissertations and Theses database. (AAT No. 3314402)

Gill, B., Bruch, J., & Booker, K. (2013). *Using alternative student growth measures for evaluating teacher performance: What the literature says* (REL 2013–002). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2010). *Supervision and instructional leadership: A developmental approach* (8th ed.). Boston, MA: Allyn & Bacon.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *Review of Economics and Statistics, 89*(1), 134–150.

Goldhaber, D., & Brewer, D. (2001). Evaluating the evidence on teacher certification: A rejoinder. *Educational Evaluation and Policy Analysis, 23*(1), 79–86.

Goldhaber, D., & Walch, J. (2011). Strategic pay reform: A student outcomes-based evaluation of Denver's ProComp teacher pay initiative. *Economics of Education Review, 31*(6), 1067–1083.

Hackman, J., & Oldham, G. (1980). *Work redesign.* Reading, MA: Addison-Wesley.

Halawah, I. (2005). The relationship between effective communication of high school principal and school climate. *Education, 126*(2), 334–345.

Hall, G. E. (2013). Evaluating change processes: Assessing the extent of implementation (constructs, methods, and implications). *Journal of Educational Administration, 51*(3), 264–289.

Hall, G. E., & Hord, S. M. (2015*). Implementing change: Patterns, principles, and potholes*. Boston, MA: Pearson Education.

Hallinger, P., & Heck, R. H. (2011). Conceptual and methodological issues in studying school leadership effects as a reciprocal process. *School Effectiveness and School Improvement, 22*(2), 149–173.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100,* 84–117.

Hanushek, E. A. (2010). *The economic value of higher teacher quality* (Working Paper 16606). Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466–479.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). *Teachers, schools, and academic achievement* (Working Paper No. 6691). Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A., & Lindseth, A. A. (Eds.) (2009). *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America's public schools*. Princeton, NJ: Princeton University Press.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267–271. doi:http://dx.doi.org/10.1257/aer.100.2.267

Hargreaves, A. (2000). Emotional geographies of teaching. *The Teachers College Record, 103*(6), 1056–1080.

Hargreaves, A. (2001). Classrooms, colleagues, communities and change: The sociology of the century. *Asia-Pacific Journal of Teacher Education and Development, 4*(1), 101–129.

Hargreaves, A. (2004). Inclusive and exclusive educational change: Emotional responses of teachers and implications for leadership. *School Leadership & Management, 24*(3), 287–309.

Hargreaves, A., & Dawe, R. (1990). Paths of professional development: Contrived collegiality, collaborative culture, and the case of peer coaching. *Teaching and Teacher Education, 6*(3), 227–241.

Harris, D. N., & Herrington, C. D. (2006). Accountability, standards, and the growing achievement gap: Lessons from the past half-century. *American Journal of Education, 112*(2), 209–238.

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.

Heneman, H. G., III, Kimball, S., & Milanowski, A. (2006). *The teacher sense of efficacy scale: Validation evidence and behavioral prediction* (WCER Working Paper No. 2006–7). Madison, WI: Wisconsin Center for Education Research.

Herman, J. L., Heritage, M., & Goldschmidt, P. (2011). *Developing and selecting assessments of student growth for use in teacher evaluation systems*. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hill, H. C., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371–384.

Holland, P. (2005). The case for expanding standards for teacher evaluation to include an instructional supervision perspective. *Journal of Personnel Evaluation in Education, 18*(1), 67–77.

Holland, P. E., & Adams, P. (2002). Through the horns of a dilemma between instructional supervision and the summative evaluation of teaching. *International Journal of Leadership in Education, 5*(3), 227–247.

Honig, M., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher, 33*(8), 16–30.

Hopkins, P. T. (2013). *Teacher perspectives of the use of student performance data in teacher evaluations* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3578118)

Interstate New Teacher Assessment and Support Consortium. (1992). *Model standards for beginning teacher licensing, assessment and development: A resource for state dialogue*. Retrieved from http://thesciencenetwork.org/docs/BrainsRUs/ Model 2Standards forBegTeaching_Paliokas.pdf

JCSEE. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators.* Newbury Park, CA: SAGE.

JCSEE. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd ed.). Thousand Oaks, CA: Corwin.

JCSEE. (2014a). *About JCSEE*. Retrieved from http://www.jcsee.org/about

JCSEE. (2014b*). Personnel evaluation standards*. Retrieved from http://www.jcsee.org/personnel-evaluation-standards

Jiang, J. Y., Sporte, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH Students. *Educational Researcher, 44*(2), 105–116.

Johnson, J. (2012). *You can't do it alone: A communications and engagement manual for school leaders committed to reform*. Lanham, MD: Rowman & Littlefield.

Kelsey, A. (2009). *Principal collaboration using student growth data to improve student success* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3338892)

Kilgore, S., & Reynolds, K. (2011). *From silos to systems: Reforming schools for success*. Thousand Oaks, CA: Corwin.

Kimball, S., White, B., Milanowski, A., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education, 79*(4), 54–78.

Kleinhenz, E., & Ingvarson, L. (2004). Teacher accountability in Australia: Current policies and practices and their relation to the improvement of teaching and learning. *Research Papers in Education, 19*, 31–49.

Knight, C. (2008). *The effects of high-stakes testing on teacher accountability* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3316096)

Kovach, J., & Manning, J. (2003). The continuing challenges of excellence and equity. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices* (2nd ed., pp. 25–47). Alexandria, VA: ASCD.

Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk, 10*(3), 333–349.

Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education, 21*(8), 899–916.

Leithwood, K., & Jantzi, D. (2005). Transformational leadership. In B. Davies (Ed.), *The essentials of school leadership* (pp. 31–43). Thousand Oaks, CA: SAGE.

Leithwood, K., Strauss, T., & Anderson, S. (2007). District contributions to school leaders' sense of efficacy: A quantitative analysis. *Journal of School Leadership, 17*(6), 735–770.

Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality.

Louis, K. (2007). Changing the culture of schools: Professional community, organizational learning, and trust. *Journal of School Leadership, 16*, 477–487.

Marion, S., DePascale, C., Domaleski, C., Gong, B., & Diaz-Biello, E. (2012). *Considerations for analyzing educators' contributions to student learning in non-tested subjects and grades with a focus on student learning objectives*. Dover, NH: Center for Assessment.

Marks, M. (2005). *Comparison of teacher perceptions of teacher evaluation systems in two school districts* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3196215)

Marshall, K. (2013*). Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap* (2nd ed.). San Francisco, CA: Jossey-Bass.

Maryland Department of Education. (2012). *Maryland teacher and principal evaluation guidebook* (v. 3). Baltimore, MD: Author.

Marzano, R. S., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Aurora, CO: Mid-continent Research for Education and Learning.

Maslow, V., & Kelley, C. (2012). Does evaluation advance teaching practice? The effects of performance evaluation on teaching quality and system change in large diverse high schools. *Journal of School Leadership, 22*(3), 1–33.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2004). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Education.

McGuinn, P. (2006). *No Child Left Behind and the transformation of federal education policy, 1965–2005*. Lawrence, KS: University Press of Kansas.

McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy, 26*(1), 136–159.

McNeil, M. E., Hood, A. W., Kurtz, P. Y., Thousand, J. S., & Nevin, A. (2006, April). *A self-actualization model for teacher induction into the teaching profession: Accelerating the professionalization of beginning teachers.* Paper presented at the annual meeting of the Teacher Education Division, Council for Exception Children, San Diego, CA.

Mehan, H. (2000). Beneath the skin and between the ears: A case study in the politics of representation. In B. Levinson et al. (Eds.), *Schooling the symbolic animal: Social and cultural dimensions of education* (pp. 259–279). Lanham, MD: Rowman & Littlefield.

Mendro, R., Jordan, H., Gomez, E., Anderson, M., & Bembry, K. (1997, April). *An application of multiple linear regression determining longitudinal teacher effectiveness*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: Guilford Press.

Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership, 70*(3), 10–13.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33–53. London, UK/New York, NY: Routledge/Falmer.

Milanowski, A., & Heneman, H. G., III. (2001). Assessment of teacher reactions to a standards-based evaluation system: A pilot study. *Journal of Personnel Evaluation in Education*, *15*(3), 193–212.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, NJ: Merrill/Pearson.

Mintrop, H., & Sunderman, G. L. (2009). *Why high stakes accountability sounds good but doesn't work: And why we keep on doing it anyway*. Los Angeles, CA: The Civil Rights Project/Proyecto Derechos Civiles at UCLA.

Monyatsi, P., Steyn, T., & Kamper, G. (2006). Teacher perceptions of the effectiveness of teacher appraisal in Botswana. *South African Journal of Education*, *26*(3), 427–441.

Morgado, J. C., & Sousa, F. (2010). Teacher evaluation, curricular autonomy and professional development: Trends and tensions in the Portuguese educational policy. *Journal of Education Policy*, *25*(3), 369–384.

Muñoz, M. A., Prather, J. R., & Stronge, J. H. (2011). Exploring teacher effectiveness using hierarchical linear models: Student- and classroom-level predictors and cross-year stability in elementary school reading. *Planning and Changing*, *42*(3/4), 241–273.

Muñoz, M. A., Scoskie, J. R., & French, D. L. (2013). Investigating the "black box" of effective teaching: The relationship between teachers' perception and student

achievement in a large urban district. *Educational Assessment, Evaluation and Accountability, 25*(3), 205–230.

Namaghi, S. (2010). A data-driven conceptualization of teacher evaluation. *The Qualitative Report*, *1*5(6), 1504–1522.

National Governors Association. (2008). *Benchmarking for success: Ensuring US students receive a world-class education*. Washington, DC: Author

National Governors Association. (2011). *Common core state standards initiative.* Retrieved from http://www.nga.org/cms/home/special/col2-content/common-core-state-standards-init.html

Nias, J. (1999). Teachers' moral purposes: Stress, vulnerability, and strength. In R. Vandenberghe, & A. M. Huberman (Eds.), *Understanding and preventing teacher burnout. A sourcebook of international research and practice* (pp. 223–237). Cambridge, UK: Cambridge University Press.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Nye, B., Konstanopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237–257.

Ovando, M. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *Journal of Personnel Evaluation in Education, 15*(3), 213–231.

Ovando, M., & Ramirez, A. (2007). Principals' instructional leadership within a teacher performance appraisal system: Enhancing students' academic success. *Journal of Personnel Evaluation in Education*, *20*(1–2), 85–110.

Pearlman, M. A., & Tannenbaum, R. (2003). Teacher evaluation practices in the accountability era. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.),

International handbook of educational evaluation (pp. 609–641). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International University Press.

Pizzi, J. D. (2009). *Urban secondary teachers' perceptions of a standards-based teacher evaluation system* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No: 3344939)

Popham, W. (2013). Tough teacher evaluation and formative assessment: Oil and water? *Voices from the Middle, 21*(2), 10–14.

Price, H. (2012). Principal-teacher interactions: How affective relationships shape principal and teacher attitudes: *Educational Administrators Quarterly, 48*(1), 39–85.

Proctor, D., Walters, B., Reichardt, R., Goldhaber, D., & Walch, J. (2011). *Making a difference in education reform: ProComp external evaluation report 2006–2010. Prepared for the Denver Public Schools.* Denver, CO: University of Colorado, The Evaluation Center.

Raymond, M. E., & Hanushek, E. A. (2003). High-stakes research: The campaign against accountability has brought forth a tide of negative anecdotes and deeply flawed research. *Education Next, 3*(3), 48.

Reeves, D. B. (2004). *Accountability for learning: How teachers and school leaders can take charge.* Alexandria, VA: ASCD.

Reineke, R. A., Willeke, M. J., Walsh, L. H., & Sawin, C. (1988). Review of personnel evaluation system: A local application of standards. *Journal of Personnel Evaluation in Education, 1(*4), 373–378.

Richardson, V., & Placier, P. (2001). Teacher change. In V. Richardson (Ed.), *Handbook on research on teaching* (4th ed., pp. 905–947). Washington, DC: American Research Association.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: Free Press.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press

Rothstein, R., Jacobsen, R., & Wilder, T. (2008). *Grading education: Getting accountability right.* Washington, DC: Economic Policy Institute.

Roussin, J., & Zimmerman, D. (2014). Inspire learning, not dread: Create a feedback culture that leads to improved practices. *JSD, 35*(6), 36–39.

Ryser, G., & Rambo-Hernandez, K. (2014). Using growth models to measure school performance: Implications for gifted learners. *Gifted Child Today, 37*(1), 17–23.

Sanders, J. B. (1999, April). *General background on the Joint Committee on Standards for Educational Evaluation*. Paper presented at the 1999 annual meeting of the National Council on Measurement in Education, Montreal, Canada.

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

Sarason, S. (1996). *Revisiting"the culture of the school and the problem of change."* New York, NY: Teachers College Press.

Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago, IL: Consortium on Chicago School Research at the University of Chicago.

Schein, E. H. (2004). *Organizational culture and leadership* (3rd ed.). San Francisco, CA: Jossey-Bass.

Schmidt, M., & Datnow, A. (2005). Teachers' sense-making about comprehensive school reform: The influence of emotions. *Teaching and Teacher Education, 21*, 949–965. doi:10.1016/j.tate.2005.06.006

Schmitt, L., & Ibanez, N. (n.d.). *AISD REACH program update: 2009–2010 Texas Assessment of Knowledge and Skills (TAKS) results and student learning objectives (SLOs)* (Department of Program Evaluation Publication 09.83 RB). Retrieved from http://www .austinisd.org/sites/default/files/dre-reports/rb/09.83_RB_AISD_REACH_ Program_Update_2009-2010_Texas_Assessment_of_Knowledge_and_Skills_ TAKS_Results_and_Student_Learning_Objectives_SLOs.pdf

Schneider, B., & Bryk, A. S. (2000, April). *Studying trust in school reform in Chicago*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly, 48*(4), 626–663.

Sergiovanni, T. J. (1992). Why we should seek substitutes for leadership. *Educational Leadership, 5,* 41–45.

Seyfarth, J. T. (2001). *Personnel management for effective schools* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Shadish, W. R., Campbell, D. T., & Cook, T. D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23.

Spillane, J., Reiser, B., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72,* 387–431.

Stronge, J. H. (2007). *Qualities of effective teachers* (2nd ed.). Alexandria, VA: ASCD.

Stronge, J. H. (2010a). *Effective teachers = student achievement: What the research says*. Larchmont, NY: Eye on Education.

Stronge, J. H. (2010b). *Evaluating what good teachers do: Eight research-based standards for assessing teacher excellence*. Larchmont, NY: Eye on Education.

Stronge, J. H., Gareis, C. R., & Little, C. A. (2006). *Teacher pay and teacher quality: Attracting, developing, and retaining the best teachers*. Thousand Oaks, CA: SAGE.

Stronge, J. H., & Tucker, P. D. (1999). The politics of teacher evaluation: A case study of new system design and implementation. *Journal of Personnel Evaluation in Education, 13,* 339–359.

Stronge, J. H., & Tucker, P. D. (2003). *Teacher evaluation. Assessing and improving performance.* Larchmont, NY: Eye on Education.

Stronge, J. H., Tucker, P. D., & Hindman, J. L. (2004). *Handbook for qualities of effective teachers*. Alexandria, VA: ASCD.

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education, 62*(4), 339-355.

Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (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/s11092-008-9053-2

Stufflebeam, D. (1968, April). *Evaluation as enlightenment for decision making*. Paper presented at the Association for Supervision and Curriculum Development Conference on Assessment Theory, Sarasota, FL.

Stufflebeam, D. (2004). The 21st century CIPP model. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influence* (pp. 245–266). Thousand Oaks, CA: SAGE.

Stufflebeam, D. (2007). *The CIPP evaluation model checklist*. Retrieved from http://www.wmich.edu/evalctr/archive_checklists/cippchecklist_mar15.pdf

Stufflebeam, D., & Sanders, J. R. (1990). Using the personnel evaluation standards to improve teacher evaluation. In J. Millman, & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pp. 416–428). Newbury Park, CA: SAGE.

Stufflebeam, D., & Shinkfield, A. J. (2007). *Evaluation theory, models, & applications*. San Francisco, CA: Wiley.

Sullivan, S., & Glanz, J. (2005). *Supervision that improves teaching: Strategies and techniques*. Thousand Oaks, CA: Corwin.

Sunderman, G. (2010). Evidence of the impact of school reform on systems governance and educational bureaucracies in the United States. *Review of Research in Education*, *34*(1), 226–253.

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly, 46*(1), 31–56.

Taut, S. M., & Alkin, M. C. (2003). Program staff perceptions of barriers to evaluation implementation. *American Journal of Evaluation, 24*(2), 213–226.

Tennessee Department of Education. (2012). *Teacher evaluation in Tennessee: A report on year 1 implementation.* Nashville, TN: Author.

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership, 66*(2), 32–37.

Tollefson, N. (2000). Classroom applications of cognitive theories of motivation. *Educational Psychology Review*, *12*(1), 63–83.

Tomlinson, C. (2007). Learning to love assessment. *Educational Leadership*, *65*(4), 8.

Troman, G., & Woods, P. (2001). *Primary teachers' stress*. New York, NY: Routledge/Falmer.

Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student achievement.* Larchmont, NY: Eye on Education.

Tucker, P. D., Stronge, J. H., & Gareis, C. R. (2002). *Handbook on teacher portfolios*. Larchmont, NY: Eye on Education.

Turnbull, B. (2002). Teacher participation and buy-in: Implications for school reform initiatives. *Learning Environments Research, 5*, 235–252.

Tuytens, M., & Devos, G. (2009). Teachers' perceptions of the new teacher evaluation policy: A validity study of the Policy Characteristics Scale. *Teaching and Teacher Education, 25*(6), 924–930.

Tuytens, M., & Devos, G. (2010). The influence of school leadership on teachers' perceptions of teacher evaluation policy. *Educational Studies, 36*(5), 521–536.

U.S. Department of Education. (2004). *New No Child Left Behind flexibility: Highly qualified teachers* (Fact Sheet). Washington, DC: Author.

U.S. Department of Education. (2010a). *A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act*. Washington, DC: Author.

U.S. Department of Education. (2010b*). Teacher incentive fund-awards*. Retrieved from http://www2.ed.gov/programs/teacherincentive/awards.html

van den Berg, R. (2002). Teacher's meaning regarding educational practices. *Review of Educational Research, 72*(4), 577–625.

van Veen, K., & Lasky, S. (2005). Emotions as a lens to explore teacher identity and change: Different theoretical approaches (Introduction to special issue on emotion, teacher identity and change). *Teaching and Teacher Education, 21*(8), 895–898.

van Veen, K., Sleegers, P., & Van de Ven, P. (2005). One teacher's identity, emotions and commitment to change: A case study into the cognitive-affective processes of a secondary school teacher in the context of reforms. *Teaching and Teacher Education, 21*(8), 917–934.

Vekeman, E., Devos, G., & Tuytens, M. (2015). The influence of teachers' expectations on principals' implementation of a new teacher evaluation policy in Flemish secondary education. *Educational Assessment, Evaluation and Accountability, 27*(2), 129–151. doi:10.1007/s11092-014-9203-4

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wacha, M. (2013). *Teacher perceptions of the evaluation process* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1536043)

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and students achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wechsler, M. E., & Shields, P. M. (2008). *Teaching quality in California: A new perspective to guide policy*. Santa Cruz, CA: Center for the Future of Teaching and Learning.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: TNTP.

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*, 12.

Winslow, R. A. (2015). *Administrative feedback following classroom observations as part of a Danielson-based teacher evaluation system: Teacher and administrator perceptions* (Unpublished doctoral dissertation). Southern Illinois University, Edwardsville, IL).

Witziers, B., Bosker, R. J., & Krüger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*(3), 398–425.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: SAGE.

Youngcourt, S., Leiva, P., & Jones, R. (2007). Perceived purposes of performance appraisal: Correlates of individual and position focused purposes on attitudinal outcomes. *Human Resource Development Quarterly, 18*(3), 315–343.

Zembylas, M., & Barker, H. (2007). Teachers' spaces for coping with change in the context of a reform effort. *Journal of Educational Change, 8*(3), 235–256.

Zepeda, S. (2011). Professional development: What works? *Journal of Educational Administration, 51*(5), 728–730. doi:10.1108/JEA-11-2012-0128

Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin, 87*(636), 28–37.

VITA

R. Shannon Finnegan

Birthplace: West Palm Beach, Florida

Education

2013-2016	The College of William and Mary Williamsburg, Virginia Doctor of Education
1999-2001	Old Dominion University Norfolk, Virginia Master of Education
1987-1989	University of Florida Gainesville, Florida Bachelor of Science

Experience

2009–Present	Vice-Principal, Berry Elementary, Charles County, Maryland
2007–2009	Instructional Specialist, Special Education, Charles County, Maryland
2006–2007	Instructional Specialist, Charles County, Maryland
2005–2006	Gifted Education Resource Teacher, Charles County, Maryland
2001–2005	Special Education Elementary Teacher, Charles County, Maryland
1999–2001	Special Educator Teacher, King George County, Virginia
1996–1999	Middle School ELA Teacher, Fredericksburg Christian School, Virginia
1994–1996	Middle/High ELA Teacher, Covenant Christian School, Chula Vista, California
1990–1991	ELA Teacher Saint Paul's Episcopal Day School, Atlantic Beach, Florida