TEACHER PERCEPTIONS OF PENNSYLVANIA'S TEACHER EFFECTIVENESS MODEL

A Dissertation Submitted to the Faculty of

Immaculata University

By

Marshall E. Hoffritz

In partial fulfillment of the

requirements for the degree of

Doctor of Education

Immaculata, Pennsylvania March 2014

TITLE OF DISSERTATION

Teacher Perceptions of Pennsylvania's Teacher Effectiveness Model

AUTHOR: Marshall E. Hoffritz

Mary D. Calderone, Ed. D.
Chairperson

Stephen A. Catanzaro, Ed. D.
Committee

Johnna L. Miller, Ed D.
Committee

Paul W. Brown Ed.D
Reader

ON BEHALF OF IMMACULATA UNIVERSITY

Janet F. Kane, Ed.D.
Dean, College of Graduate Studies

Thomas Compitello, Ed D.
Chairperson, Education Division

DATE: March 24, 2014

Abstract

This qualitative study investigated teacher perceptions on the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model and its influence on teacher instruction, collaboration, and professional growth. The study was conducted in four public school districts that serve students from Kindergarten to 12th grade. The 4 school districts participating included an Intermediate Unit, 2 suburban districts in southeastern Pennsylvania, and a rural district located in eastern-central Pennsylvania. Data instruments used in the study incorporated a Likert scale survey, open-ended questions, and individual interviews. Thirty five teachers participated in the questionnaire portion of the study. Additional insights were gathered from 13 teacher interviews. The results of the study revealed that the teacher evaluation process has impacted teacher instructional practices and professional growth. Teachers acknowledged that the evaluation process was a good indicator for determining the quality of teaching. Themes from their responses emerged related to self-reflection, feedback, and utilizing student data to guide teacher planning and instruction. The data also illustrated that teachers participate in collaborative inquiry to discuss ways to improve teaching through professional conversations. However, it is not clear whether the Pennsylvania teacher evaluation has had a direct impact on teacher collaboration. Respondents did identify that their evaluation has been used to direct their professional development activities that focus on student achievement that directly affected classroom instruction to enhance classroom learning.

Acknowledgements

This dissertation would not have been possible without the help of so many people. It is the product of a larger intellectual conversation with people who have navigated my academic career.

I would like to express many thanks, appreciation, and gratitude to my committee chair, Dr. Mary Calderone. Her belief, guidance, genuine caring, and calm nurturing manner meant more than I can express. Without her insight and support this dissertation would not have been possible.

I would also like to express my sincere appreciation to the members of my dissertation committee, Dr. Johnna Weller and Dr. Stephen Catrambone. Thank you for serving on my dissertation committee and for your continuous encouragement and guidance throughout the process. Your extensive knowledge, feedback, and suggestions provided me with the motivation and determination to persevere.

A special thank you goes to Kelly DiPrinzio for taking time away from her family to help a colleague. Your support and scholarly expertise have been greatly appreciated.

Lastly, I would also like to express my genuine appreciation to my friends and family for their support and words of encouragement which cannot be overstated. I would especially like to thank my mom and my dad for instilling in me the belief that I could accomplish anything that I set out to do in life.

Dedication

It is with all my heart that I dedicate this dissertation to my greatest supporter, my wife, Tamara. Throughout this journey, the doctoral process, graduate education, and our years together, she has shown unwavering support and commitment to me with persistent encouragement and love. Her belief in me gave me the strength to pursue and the courage to continue my educational work. I will always be grateful. She has sacrificed much, putting aside her education so I could continue with mine. She maintained our home while I worked and still she unselfishly reviewed pages upon pages and offered constructive feedback. Tamara, for these reasons and so many more, I dedicate my work to you. This achievement is as much yours as it is mine. Thank you, I love you.

To my son, Seth, and my daughter, Julia, none of this would have been possible without your love, patience, and support. I realize that this process has taken time away from our family, yet both of you have been so understanding. I hope I have instilled in you the importance of education, working hard, and pursuing your goals. Yes, I am finally done!

Table of Contents

Abstract iv

Acknowledgements v

Dedication vi

Table of Contents vii

Lists of Tables x

Chapter One – Introduction 1

Overview 1

Need for the Study 5

Statement of the Problem 9

Definition of Terms 10

Limitations 12

Research Questions 12

Summary 13

Chapter Two – Literature Review 14

Introduction 14

Reform Policies 15

Teacher Accountability 21

Standards Based Model 28

Collaborative Inquiry 32

Multiple Measures of Appraisals 37

Teacher Value Added Model 39

Performance Rubrics 43

Peer Observation 46
Summary 47
Chapter Three – Methodology 49
Introduction 49
Setting 49
Subjects 52
Instruments 53
Reliability and Validity 54
Design of the Study 55
Procedure 56
Data Analysis 57
Summary 58
Chapter Four – Results 59
Introduction 59
Analysis of Data Results 61
Research Question One 61
Research Question Two 69
Research Question Three 76
Summary 83
Chapter Five – Discussion 86
Summary of the Study 86
Summary of the Results 86

Limitations Found in the Study 93
Relationship to Other Research 94
Recommendation for Further Research 99
Conclusion 100
References 104
Appendices 116
A. Teacher Evaluation Survey 116
B. Interview Questions 123
C. RERB Approval 124

List of Tables

Table **Page**

1.1 Pennsylvania Teacher Evaluation Phase in Plan 8

4.1 Subject Demographic Data 60

4.2 Teachers' Perceptions of the Impact of the Evaluation on Instructional Practice 63

4.3 Teachers' Perceptions of the Impact of the Evaluation Process on Collaboration 71

4.4 Teachers' Perceptions of the Impact of the Evaluation Process on Professional Growth 78

Chapter One – Introduction

Overview

Teacher quality is the most important school factor related to student academic achievement (Chait, 2009; Hightower, Lloyd, Wittenstein, Sellers, & Swanson, 2011). The federal No Child Left Behind Act (NCLB) of 2001, in pursuit of improving school performance, requires schools to employ only "highly qualified" teachers as a means to improve teacher quality. It mandates that states require teachers to demonstrate that they are "highly qualified" through a certification process (United States Department of Education, 2004b). For teachers to meet the law's standard in Pennsylvania, a teacher must minimally hold a valid Pennsylvania teaching certification and a bachelor's degree, and pass a content area test in their teaching assignment (Pennsylvania Department of Education [PDE], 2013b). However, teacher performance is not included as part of the NCLB legislation.

More recently, the Obama administration released a competitive educational grant program entitled Race to the Top (U.S. Department of Education, 2009). The Race to the Top fund provides competitive grants to encourage and reward states that are creating the conditions for education innovations and reforms. This grant has instigated an unprecedented wave of teacher evaluation reform (McGuinn, 2012). For states to meet the Race to the Top requirements, states are tasked with reworking the teacher evaluation process to be based on student achievement (U.S. Department of Education, 2009). Currently, the states' education agencies are engaged in methods to realign and adapt the teacher evaluation system. This process is complicated, and involves balancing multiple issues related to instruction: traditional methods of teacher evaluation, emerging

approaches that are tied to student achievement, and strategies that focus on multiple measures of teacher and student performance (Hightower et al., 2011; McGuinn, 2012).

According to Ron Tomalis, in Pennsylvania's Secretary of Education 2012 budget report, 99.4 % of all teachers received a satisfactory rating on their evaluations in the 2009-10 school year. Despite this high percentage, the teacher evaluation process lacks the ability to assist teachers to improve professionally and distinguish those who are rated "successful" from those who are rated "needs improvement" or "unsatisfactory" (Danielson, 2007). Typically, teacher evaluations are summative, year-end reviews including an overall rating of satisfactory or unsatisfactory (Johnson, 2011). Baratz-Snowden (2009) noted that teacher evaluations are used both for the improvement of teaching and learning and for accountability purposes. Further, Baratz-Snowden suggested that it is important to identify exceptional teachers, and develop policies to cultivate and nurture quality teachers in every classroom. Koops and Winsor (2005) stated that the goal of evaluation is to "promote professional excellence and improve the skills of teachers by providing feedback about performance" (p. 7). Moreover, they acknowledged that it is crucial to focus on identifying and removing weak teachers. This would lead to ensuring that all students are taught by effective teachers (Baratz-Snowden, 2009).

Marshall (2012) suggested that teaching effectiveness cannot be accurately measured without looking at multiple results. Teacher evaluation must consider classroom observation, student achievement, and feedback from students. The different methods and measures are designed to validate one another and support student achievement through effective teaching and analyzing student data (Little, Goe, & Bell,

2009). Research completed by Rockoff and Speroni (2010) examined how both subjective and objective evaluations of new teachers impact student achievement. They concluded that teachers had a greater confidence in evaluation systems that reduced subjectivity by incorporating both subjective observations made by trained professionals and objective performance data. Furthermore, objective data and subjective evaluations presented "meaningful information about a teacher's future success in raising student achievement" (Rockoff & Speroni, 2010, p. 266).

Kimball, White, Milanowki, and Borman (2004) suggested that a "valid and consistent" (p. 55) standards-based teacher evaluation that utilizes multiple measures has to be developed to adequately determine the teacher's impact on student achievement. The evaluation process would then be informative in an assortment of educational purposes including instructional improvement and accountability. Kimball et al. stated, "Teacher performance assessment results could then be considered to be valid measures of teaching practice and to have the expected positive relationship to student achievement" (p. 55). According to Marshall (2012), evaluating individual teachers based on students' test scores alone is a weak measure, because it "fails to take into account the work done by pullout teachers, specialists, tutors, and teachers in previous grades, all of whom contribute to student outcomes" (p. 52). Student achievement cannot be the primary and indispensable outcome (Measures of Effective Teaching Project [MET], 2013). Policymakers consider student achievement to be only one among many elements of good teaching. The Center for Education Policy and Practice [CEPP] (2010) noted in its review of teacher evaluations, that it is problematic to use standardized test scores in isolation to evaluate teachers, and multiple measures provide

richer and more accurate results. The Center emphasized that judgments based on one method are vulnerable to flawed conclusions, while multiple measures of evidence using different types of data are more likely to result in accurate evaluation. Multiple data sources should be used to gather information about both the practice and the practitioner, including observations, artifacts, student work, teacher work, administrator work, assessment results, and survey data (CEPP, 2010).

Marshall (2009) argued that the current design behind evaluations is flawed and that the process rarely changes teachers' performance and subsequently student achievement growth. Marshall believed that for teacher growth and student achievement to occur, a change with the way teachers are supervised and evaluated must be adapted. Beyond the use of multiple measures for individual teacher evaluations, Marshall concurred with other researchers (DuFour, DuFour, Eaker, & Karhanek, 2004; Glaser, 2005; Reeves, 2007) in his belief in the benefits of teacher teams working collaboratively. Teacher collaboration contributes to instructional improvement and student achievement. It can be most effective when the purpose of teacher teams is to enhance teacher effectiveness for the ultimate benefit of students (DuFour, 2011; Reeves & Flach, 2011; Servage, 2009; Wiggins & McTighe, 2007).

The benefits of teacher collaboration as part of teacher evaluations have gained much attention from school leaders and researchers in recent years. Numerous researchers have linked the benefits of teacher collaboration to teacher growth and student achievement (DuFour & Marzano, 2009; Fullan, 2006; Koops & Winsor, 2005). DuFour et al. (2004) supported empowering teacher teams to determine common valid, local assessments that will provide baseline data assessment of every student. They

believed that teacher teams need to set specific goals, teach, assess, and share student data throughout the year. As a circulative process, DuFour et al. recommended that teacher teams assess all students at the end of the year and present the results for evaluation. Teacher evaluations can be designed to assist evaluators and teacher teams in examining team and individual data, giving the team a collective evaluation on student value added data, and including a team score as factors in the evaluation (Koops & Winsor, 2005; Marshall, 2012).

Need for the Study

Teacher evaluation continues to be a subject of public debate in Pennsylvania and across the United States. Pennsylvania is currently in the process of developing a new standards-based teacher evaluation system, with a goal of more accurately measuring teacher effectiveness (Pennsylvania Department of Education, Standards Aligned System, 2012). The Pennsylvania Department of Education's new evaluation system for educators will include multiple measures of student growth as part of the final evaluation. Until 2013, the Pennsylvania system in place for evaluating the quality of educators has been unchanged for over 40 years (Aument, 2011). In 2011, Pennsylvania State Representative Aument (2011) stated, "In the 2009-10 academic year, 99.4% of public school teachers in the Commonwealth received an evaluation rating of satisfactory" (para. 4). Aument believed that those high percentages represent an outdated evaluation system.

The development of a standards-based evaluation has emerged in response to deficiencies in the traditional evaluation practices. Traditional evaluation systems are outdated, utilizing limited evaluative measures that do not focus on student achievement

(Danielson & McGreal, 2000). Teacher observations and evaluations have been traditionally driven by administrative observation (Toch & Rothman, 2008). The lack of accurate observable data has resulted in an inflation of satisfactory evaluation ratings (Danielson & McGreal, 2000).

While traditional evaluations tend to be one dimensional, relying on a single observation of a teacher in a classroom, a comprehensive multiple-form observation and evaluation model more accurately portrays a better picture of a teacher's performance (Donaldson, 2009). Koops and Winsor (2006) believed that observations, supervision, and evaluation together provide a system that encourages and motivates teachers. Koops and Winsor agreed that the formal written evaluation that judges the overall performance of a teacher should only be one part of a larger process. Recently, new demands for greater accountability by both local communities and governments have emerged, creating an interest in making teacher evaluation fairer and more accurate in its structure (Kohut, Burnap, & Yon, 2007; Toch & Rothman, 2008). A standards-based teacher evaluation system that uses multiple measures of appraisal is being considered by schools as a method to improve instruction and accountability (MET, 2013; Papay, 2012).

In the 2013-14 school year, Pennsylvania implemented the use of *Enhancing Professional Practice: A Framework for Teaching* (Pennsylvania Department of Education [PDE], 2013b) as the standards-based model for teacher evaluations. The state's Indicators of Student Growth and Educator Practices has been developed as a criteria and rubric guide to inform discussions on what constitutes multiple measures of student growth (Pennsylvania State Education Association [PSEA], 2011). These

categories allow educators to bring evidence from multiple sources into the evaluation process.

Pennsylvania has spent three years, previous to the start of implementation, piloting the new evaluation system. The first round of Pennsylvania's pilot program took place during the 2010-11 school year with 10 districts. Phase two (2011-12) of the program was expanded to 100 districts; and in phase three, (2012-13) 300 districts (McGuinn, 2012; PDE, 2013c).

The Pennsylvania Teacher Effectiveness model is comprised of two parts. Fifty percent of the new evaluation system requires all classroom teachers to be evaluated using traditional observations based on the Danielson Framework (PDE, 2013c). The state approved observation forms measure performance in four domains: (1) Planning and Preparation, (2) Classroom Environment, (3) Instruction, and 4) Professional Responsibilities. The finalized evaluation system will calculate the other 50% of a teacher's overall rating from multiple measures of student achievement and will be implemented in subsequent years. The 50% will be split into the following three categories: (1) 15% will be based on the school building data such as PSSA achievement, graduation and attendance rates; (2) 20% will be based on elective data, which are student learning objectives developed by the district, such as district designed test, National test, IEP growth, projects, portfolios, and surveys; and (3) 15% based on teacher specific Pennsylvania Value Added Assessment System (PVAAS) student growth scores (PDE, 2013c).

The new evaluation system is being phased in over a three year period starting in the 2013-14 school year. For the 2013-14 school year, 85% of the teacher rating will be

comprised of the Danielson Framework and 15% will be building data. For the 2014-15 school year, 50% of the teacher rating will be comprised of the Danielson Framework, 15% will be building data, and 35% will be based on elective data. For the 2015-16 school year, the full evaluation system will be in place. Teacher evaluations will be comprised of 50% from the Danielson Framework, 15% will be from building data, 20% elective data, and 15% teacher specific data (Pennsylvania Department of Education, 2013c). Table 1.1 depicts the Pennsylvania's teacher evaluation three year phase in plan.

Table 1.1

Pennsylvania Teacher Evaluation Phase in Plan

Year	Observation	Building Score	Teacher Specific Data	Elective Data
2013-14	85%	15%	N/A	N/A
2014-15	50%	15%	N/A	35%
2015-16	50%	15%	15%	20%

Note. Adapted from Pennsylvania Department of Education. (2013c). *Measuring Educator Effectiveness*. Retrieved from http://www.education.state.pa.us

There exists a need to explore and comprehend the effects of Pennsylvania's revised teacher evaluation system. The implementation of this system could drastically alter the evaluation criteria of professionals working as highly qualified teachers in Pennsylvania ("Educator Effectiveness," 2013; Lane & Horner, 2011). Teachers are currently evaluated by an annual evaluation based largely on classroom observations (PSEA, 2011). Effective for the 2013-14 school year, teachers will be evaluated by the new model (PDE, 2013b). The purpose of the reform is to bring accountability and instructional growth, making the evaluation process less subjective and providing students the opportunity to receive a more effective education (PDE, 2013c). While this

process is designed to improve teacher quality, it is important to monitor and collect data on the new evaluation process and measure its effects on teacher performance, teaching methods, and collaboration (Lane & Horner, 2011). Therefore, this study will examine teachers' perceptions of the impact of Pennsylvania's standards-based teacher evaluation model's influence on teacher instruction, collaboration, and professional growth.

Statement of the Problem

Consensus has been growing that the current teacher evaluation system fails to measure teacher performance accurately (DuFour & Marzano, 2009; Marshall, 2009). States across the country are reforming the evaluation systems to better address this need. New, standards-based teacher evaluation practices have recently emerged in response to deficiencies in evaluation practices and to improve instruction and accountability (Baratz-Snowden, 2009). The goal is for states to develop a comprehensive system that provides support for supervision and professional learning, while recognizing both competent teachers as well as those who need support (Darling-Hammond, Cook, Jaquith, & Hamilton, 2012; Papay, 2012). Not only do these evaluation reforms provide new directions in school personnel evaluation practices, but also their results may represent useful information regarding teacher effects on student achievement (Milanowski, 2004).

The purpose of this study was to investigate teachers' perceptions of the effectiveness of the new standards-based teacher evaluation process by qualitatively studying its impact on instruction, collaboration, and professional growth. The intent of the study was to contribute to the body of research in teacher evaluation, instructional collaboration, and professional growth and development.

Definition of Terms

For the purpose of this study, the following terms are defined:

Building Data – school-wide student information such as PSSA achievement, Keystone scores, graduation and attendance rates (PDEb, 2013).

Classroom Observation – a method to gather and document evidence of effective teaching by evaluators (peers, mentors, supervisors, and administrators) (Danielson, 2007).

Collaboration – an interactive process that enables teachers to work together as effective team based learning communities for the benefit of teacher growth and student achievement (Glaser, 2005).

Framework for Teaching – a research-based set of comprehensive criteria comprised of four domains that identifies teacher responsibilities for promoting student learning (Danielson, 2011b).

Highly Qualified Teacher – a teacher who has at least a bachelor's degree, holds a valid Pennsylvania teaching certificate, and demonstrates subject matter competency in core content exams (Pennsylvania Department of Education, 2013a).

Instructional Collaborative Inquiry – an approach to improve student achievement where teachers work together to identify common challenges, analyze student data, and examine classroom instruction (David, 2009).

Multiple Measures – indicators of evidence of student learning from various sources including state and local assessments gathered at multiple points in time within and across subject areas (Darling-Hammond et al., 2012).

Professional Growth and Development – a continuous process of teachers' engaging in professional learning to increase pedagogical knowledge through the process of collaboration, reflection, teaching, and learning (Danielson & McGreal, 2000).

Professional Learning Communities – an ongoing, formative process in which teacher teams work collaboratively and interdependently to achieve common goals (DuFour et al., 2004).

Pennsylvania Teacher Effectiveness – a state program to assess educators' teaching practices on student achievement using multiple measures ("Educator Effectiveness," 2013).

Pennsylvania Value Added Assessment System (PVAAS) – a statistical analysis tool developed to measure a school/teacher's impact on the academic progress rates of students from year to year (PDE, 2013d).

Standards-Based Evaluation – teaching standards that are measured using rubrics, which broadly reflect evidence-based practices, comprehensive standards, and detailed criteria for teachers. The standards that are established are consensus based, transparent, and provide detailed expectations about performance (Kimball, 2002).

Student Achievement – a measurement of proficiency of what students have learned over a defined period of time used for teacher feedback, lesson planning, and evaluation (Tucker & Stronge, 2005).

Teacher Evaluation – a systematic method of review for making professional judgments about teacher performance for the purposes of improving teacher instruction and personnel decision-making (Tucker & Stronge, 2005).

Limitations

Inherent in this study are several limitations. The region that was selected for this study does not represent all regions of Pennsylvania. A small number of schools were included when compared to the total number of schools across the Commonwealth of Pennsylvania. Participation in this study was voluntary. Therefore, the data included self-reported perceptions of teachers and were subject to individual bias. The educators who participated in the study had different levels of experience in their respective positions and a wide and varied knowledge base. The results may not be generalizable to contexts that are located in a different geographical area or that have a different setting. It cannot be assumed that the results of the study will reflect the awareness, use, or effectiveness of Pennsylvania's standards-based teacher evaluation process.

Research Questions

The purpose of this study was to explore teacher perceptions of the impact of a standards-based teacher evaluation model on teacher instruction, collaboration, and professional growth. The following research questions guided the study:

1. What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on their instructional practice?
2. What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on teacher collaboration?
3. What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on their professional growth?

Summary

Across the United States, states and school districts are developing, redesigning, or implementing new teacher evaluation systems based on student achievement data (Little et al., 2009). These reforms have emerged due to the shortcomings of previous evaluation models. Researchers have agreed that teacher evaluation systems have not accurately measured teacher quality because they have failed to accurately discriminate between effective and ineffective teachers (Danielson & McGreal, 2000). Additionally, evaluation systems have not supported teachers in professional growth and development in best practices (Center for Education Policy and Practice [CEPP], 2010).

Beginning in the 2013-2014 school year, teachers in Pennsylvania will be annually rated based on a standards rubric based on the Danielson Framework and multiple measures of data. The evaluation system is planned to be phased in over a three year period starting in the 2013-14 school year. For the 2013-14 school year, 85% will be comprised of the Danielson Framework and 15% will be building data. By the 2015-16 school year, classroom observations will comprise 50% of a teacher's total score. The other 50% will include multiple measures of student achievement: 15% from building-level data, 15% on teacher-specific data, and the remaining 20% will come from elective data (PDE, 2013b). Accordingly, the purpose of this study was to gain teachers' perceptions towards the effectiveness of the new standards-based teacher evaluation process by qualitatively studying its impact on instruction, collaboration, and professional growth.

Chapter Two – Literature Review

Introduction

Teacher quality has been cited as the single most important factor in determining student success (Borman & Kimball, 2004; Milanowski, 2004). There is consensus among researchers and reformers that the traditional, observation-based teacher evaluation system has done little to improve teacher quality and student achievement (Marzano, 2012; Toch & Rothman, 2008). Across the country, there is a growing movement to develop a more "holistic approach to the teacher evaluation system" (Hanover Research, 2012, p. 2).

State and federal education reforms have resulted in creating new systems of teacher accountability that are designed to improve teacher quality and student achievement through the teacher evaluation process (Baratz-Snowden, 2009). These systems require clear and objective standards for evaluation, with multiple measures of accountability. These measures must be grounded in evidence-based practice, and should reflect all parameters of instruction, including planning, preparation, direct instruction, and student performance (Danielson, 2011b). When a comprehensive evaluation process is implemented, teacher quality and student achievement levels should improve (Tucker & Stronge, 2005).

This review of the related literature will focus on the purpose of teacher evaluations and examine the approaches that have been employed to advance the teacher evaluation model. Topics of research such as teacher effectiveness, evaluation methods, and policies will be discussed. This Literature Review includes a body of research specifically focused on the following: reform policies, teacher accountability, standards-based model, multiple measures of appraisals, teacher value-added model, performance rubric, peer observations, and professional learning communities.

Reform Policies

Teacher evaluation systems are dramatically changing as a result of state and federal governments' demands to reform education in the public schools. Throughout the nation, legislative reformers have increasingly attempted to hold educators accountable for student achievement. Veir and Dagley (2002) noted, "A recurring theme in the reform movement has been the manner in which school personnel are evaluated" (p. 2).

The quality of the educational system in the United States has been called into question by legislative leaders for decades (Chait, 2009; Furtwengler, 1995). Politicians began linking the state of America's schools to the nation's economic productivity and sought to correct the problems of low student achievement and the health of the economy by enacting educational reforms (Towe, 2012). Beginning with A Nation at Risk, to No Child Left Behind (NCLB), to the adoption of the American Recovery and Reinvestment Act (ARRA), legislation has continued to create an educational climate to improve and reform the public school teacher evaluation system (Chait, 2009; Hightower et al., 2011). These educational reforms were designed to create conditions for change in the evaluation of teachers with attempts to "identify highly effective teachers based on student growth and other factors" (U.S. Department of Education, 2010, p. 4). NCLB in 2002 required all students to demonstrate proficiency in reading, mathematics, and science by the year 2014 (United States Department of Education, 2004b). Furthering NCLB's requirement, President Obama signed ARRA into law in 2009. This act funded the Race to the Top grant, designed to encourage states to create conditions for improving student achievement (U.S. Department of Education, 2009).

Decades prior to NCLB, the National Commission on Excellence in Education was created on August 26, 1981. The commission was created due to the "public perception that something (was) seriously remiss in our educational system" (National Commission on Excellence in Education [NCEE], 1983, p. 7). The commission cited several recommendations specific to teacher evaluations and student achievement. The commission found that teacher preparation programs needed substantial improvement and that the professional working conditions of teachers were on the whole unacceptable. The commission made recommendations in the area of teacher evaluations and wages stating, "Salary, promotion, tenure, and retention decisions should be tied to an effective evaluation system that includes peer review so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated" (NCEE, 1983, p. 30).

In response to *A Nation at Risk* report, states have targeted teacher evaluations to move toward a standardized platform. Previous to the report, teacher evaluations' statutes and policies had been left to local districts' discretion (Hazi & Rucinski, 2009). *A Nation at Risk* provided the basis for other reform-oriented reports from interest groups such as the Task Force on Education for Economic Growth, The Carnegie Corporation, the National Governors' Association, and Research and Policy Committee of the Committee for Economic Development (NCEE, 1983). Using teacher accountability as their framework, each of these reports expressed concerns about evaluation of teaching personnel (Furtwengler, 1995; Veir & Dagley, 2002).

A review of state statutes and regulations for teacher evaluations by Furtwengler (1995) concluded that during the years 1983 to 1985, in response to governmental

reformers, 20 states enacted their first requirements for teacher evaluations. Nearly half of the states passed policies requiring school districts to evaluate personnel based on state-developed evaluation systems and mandated that local school districts use them. The research further identified 29 states that developed specific criteria for evaluating teachers. States tinkered with specific criteria, procedures, tenure, and state instruments. This level of state involvement and mandated policy and protocol was a radical change from the previous policies that only guided local school districts (Furtwengler, 1995; National Commission on Teaching & America's Future [NCTAF], 1996). "These actions support the premise that the first wave of reform was an accountability movement, and personnel evaluation was one vehicle used by policy makers in an attempt to insure assessment of personnel" (Furtwengler, 1995, p. 3). Despite these early state efforts, teacher evaluation was still viewed as inadequate (Veir & Dagley, 2002).

The Goals 2000 Educate America Act was signed into law in March 1994 and was amended in 1996. The Act established a framework with which to identify academic standards, to measure student progress, and to provide the support that students may need to meet the standards (U.S. Department of Education, 1998). The Act required six original education goals concerning (a) school readiness, (b) school completion, (c) student academic achievement, (d) leadership in mathematics and science, (e) adult literacy, and (f) safe and drug-free schools. It added two new goals which encouraged teacher professional development and parental participation (U.S. Department of Education, 1998). The report, based on the premise of outcome-based education, was that students will reach higher levels of achievement when more is expected of them. It established a student-centered learning philosophy that focused on the empirical

measurement of student performance. In accordance with measuring student performance, the report specifically states, “Substantial education reform is an iterative process that requires ongoing action, assessment, evaluation, and corrective action. As such, both data and research take on valuable roles in informing continuous improvements” (U.S. Department of Education, 1998, p. 29).

In a 1996 landmark report, *What Matters Most: Teaching for America's Future,* the National Commission on Teaching and America's Future established that teachers were critical to improving student achievement. The report challenged the nation to install high quality teachers in every classroom in America by 2006. The Commission set a goal of educating every student in the nation with "access to competent, caring, qualified teaching"(NCTAF, 1996, p. 21). The recommendations emphasized the influence of teacher quality on learner outcomes, and were based on research that explored school districts’ hiring practices, teacher qualifications, and student achievement. The Commission (1996) highlighted major flaws in the preparation, recruitment, and retention of teachers and concluded that public education reform depended on the restructuring of the teaching profession (Davis, Williams, & Griffin, 2003). The report offered five interrelated recommendations for systemic change: (a) get serious about standards, for both students and teachers; (b) reinvent teacher preparation and professional development; (c) fix teacher recruitment and put qualified teachers in every classroom; (d) encourage and reward teacher knowledge and skill; and (e) create schools that are organized for student and teacher success (NCTAF, 1996).

As a part of the Great Society programs, the Elementary and Secondary Education Act (ESEA) was passed in 1965 (United States Department of Education, 2004a). ESEA

emphasized equal access to education, and established high standards, accountability, and professional development for teachers. The law authorized federally funded education programs that are administered by the states (Hanna, 2005). In 2002, Congress amended ESEA and reauthorized it as the No Child Left Behind Act (United States Department of Education, 2004a). The reauthorization of NCLB of 2002 represented a major expansion of the federal government's role in public education. Rather than providing financial assistance to states in their efforts to set standards and improve student achievement, the act explicitly mandated, through consequential accountability, compliance to high standards, sanctioning states and schools that fail to meet a set criteria (Kress, Zechmann, & Schmitten, 2011; United States Department of Education, 2004b).

Those criteria dramatically changed the federal government's role in education. NCLB required schools in every state to raise student proficiency to 100% in mathematics and reading by 2014, but it allowed each state to devise its own plan as to how to get students there (United States Department of Education, 2004). According to Toch and Rothman (2008), the law has mandatory criteria to improve teacher quality by requiring that schools employ only highly qualified teachers. It mandated that states use the qualifications, degrees, and certifications that teachers earned as the measure of whether teachers meet the standard. The act permits teaching with a temporary, provisional, or emergency teaching certificate. However there is a debate among education researchers as to whether or not the NCLB "highly qualified" definition is sufficient assurance that a teacher will be successful in the classroom (Learning Point Associates [LPA], 2007). LPA (2007) believed that school districts should move beyond

the NCLB mandates and emphasize teacher effectiveness through evidence and accountability to improve teacher quality.

The new reform policies of accountability on school districts placed an enormous amount of pressure on teachers to provide proof of student growth. Teachers are held to an increased accountability standard under NCLB and are continually scrutinized regarding student achievement (Donaldson & Peske, 2010; Hightower et al., 2011). Learning Point Associates (2007) recommended that school districts need to adopt standards that define good teaching, establish procedures for the teacher evaluation process, and determine the role that student achievement should play in the evaluation.

Most recently, the federal government established the largest competitive education grant in U.S. history. Under the American Recovery and Reinvestment Act of 2009, The Race to the Top program frameworks a comprehensive reform designed to encourage and reward states that create the conditions for education innovation and reform (U.S. Department of Education, 2009). The program provides incentives for states in achieving significant improvement in student outcomes, by increasing teacher effectiveness through the tracking of student achievement data (MET, 2013).

In response to Race to the Top grants, states have made changes to their teacher evaluation systems to incorporate multiple measures of teacher effectiveness, including student performance information. In partnership with the United States Department of Education, the Reform Support Network (2012) completed a two year study surveying 11 states to understand concerns over changes made to the teacher evaluation systems based on the new requirements mandated by Race to the Top grants. Their findings indicated that states are making changes to policies for how evaluations are conducted including

observations, identifying methods for collecting data on teachers' classroom practice, and laying out how to incorporate different types of evidence for rating teacher performance (Reform Support Network, 2012).

As a matter of federal law, the enactment of the NCLB and Race to the Top reinforced the states' responsibility to hold school districts accountable for student achievement (CEPP, 2010; Kress et al., 2011). Society's demands and expectations of the public education system have increased dramatically. The intention of these reform requirements is to ensure the quality of the nation's teachers and thereby raise student achievement (LPA, 2007). Measures of Effective Teaching Project (2013) reported that the shift in policies towards increasing teacher quality and accountability represents an important acknowledgement of how effective teaching can improve student achievement.

Teacher Accountability

Teacher evaluations are used to measure teacher performance with two potential purposes, to identify quality educators and to identify areas of professional growth (Danielson & McGreal, 2000; MET, 2013). Marzano (2012) described two failures of past evaluation systems. First, teacher evaluation systems have not accurately measured teacher quality. The traditional and most common method of teacher evaluation relied on one or two classroom observations (Danielson & McGreal, 2000; Marshall, 2005) but failed to meaningfully discriminate between effective and ineffective teachers. The second failure Marzano (2012) recognized was that these evaluation systems did not aid in developing highly skilled teachers, which is one of the most important factors for student achievement.

School reformers have suggested the idea of using student achievement data, including standardized test scores, as part of the documentation that teachers present

when they are evaluated. School reformers DuFour & Marzano, (2009) and Toch and Rothman (2008) debated that evaluating teachers on the basis of their students' achievement is a reasonable strategy and a direct method to measure teacher performance.

Kimball et al. (2004) examined the relationship between teacher behavior, as measured through the evaluation system, and the amount of student achievement attributable to teachers. The results of the study suggested that there is "sufficient reliable variation in student achievement at the teacher level to be related to teacher evaluation scores" (p. 64). Researchers have argued that teacher evaluation and student achievement should be linked (Reeves, 2010; Tucker & Stronge, 2005). However, Braun (2005) stated that testing alone does not adequately capture the complexities and extent of student learning because of the many variables involved.

Marshall (2009), along with Danielson and McGreal (2007), advised that it is necessary that teachers present evidence of student learning through test results and other formative and summative assessments as a part of the tenure system. Toch and Rothman (2008) suggested that calculating a teacher's performance based on how much students' test scores increase over the course of a school year provides important information on teacher effectiveness in the classroom.

Colvin, Flannery, Sugai, and Monegan (2009) believed that focusing on using improved student scores would result in an increased accountability for teachers. Using scores for this reason will enhance effective instructional and classroom practices. Danielson (2007) commented that to be effective, administrators must establish

assessment and feedback systems that give educators specific information on what they are doing and what effect their actions have on student performance.

Often teacher evaluations are lenient and fail to adequately differentiate teachers' effectiveness in the classroom (Chait, 2009; Toch & Rothman, 2008). Marshall (2009) argued, "The theory of action behind supervision and evaluation is flawed and the conventional process rarely changes what teachers do in their classrooms" (p. 724). Donaldson (2009) postulated that teacher evaluations have come under fire because most teachers are given at least satisfactory ratings. Stating that evaluations reflect what is observable and measurable in a formal observation, Donaldson noted that traditional evaluations lack other factors such as student performance and professional growth. In addition, Donaldson acknowledged that the evaluators themselves may impede good evaluations, because they may lack the time or skill sets necessary for best evaluation practices.

In his review of research on raising the quality of teacher evaluations, Donaldson (2009) identified cultural issues such as the non-interference that causes administrators to give broad discretion to teachers in their classrooms. Cultural norms of non-interference with teachers by administrators are characterized by a lack of evaluator willingness to break norms and long standing school culture. These issues create barriers to improving the teacher evaluation process (Donaldson, 2009).

The rationale of developing quality teacher evaluations can be further observed by the high percentage of teachers meeting successful criteria for tenure as opposed to the low number of teachers identified as being in need of improvement (MET, 2013). Teacher evaluation results have shown that only a small statistical percentage of teachers

have received unsatisfactory evaluations. Donaldson's (2009) research showed that only "14 of 11,000 teachers in Philadelphia were fired in 2003-04; 12 of 2,600 in Pittsburgh were dismissed or non-renewed" (p. 12). This equates to at least satisfactory ratings for 99.9% of all teachers evaluated in those districts (Donaldson, 2009).

A study of teacher evaluations conducted in Chicago between 2003 and 2005 found that the majority of teachers had highly inflated performance ratings (New Teacher Project, 2007). Over the four-year period, 93% of Chicago teachers earned the two highest ratings "superior" or "excellent", and only three in 1,000 received "unsatisfactory" ratings. Even in 87 schools that had been identified as failing, 79% did not award a single unsatisfactory rating to teachers between 2003 and 2005. Donaldson (2009) reported that between 1995 and 2005, only one in every 930 teachers (.1) in Illinois received an unsatisfactory rating. On average, 99% of Oregon teachers were rated satisfactory each year (Donaldson, 2009).

Research by Marshall (2009) on effective systems of teacher evaluations reinforced Donaldson's (2009) remarks, stating that traditional supervision and evaluation processes are not effective systems for improving teaching and learning. Marshall's research pointed out that many school districts required formal observations in an attempt to compensate for the limited time that principals spend in individual classrooms. These systems typically involved brief observations of teaching by the principal or another school administrator, with results recorded on a form that is often organized as a checklist. The evaluator assesses overall teacher performance using a rating scale ranging from "outstanding" to "unsatisfactory" (Marshall, 2009)

Baratz-Snowden (2009) completed a study of states' tenure laws, comparing state evaluation systems to determine whether they sustain high quality teaching. The study supported Donaldson's (2009) and Marshall's (2009) criticisms, and identified problems with the current teacher evaluation and tenure models as systems in need of fixing. Baratz-Snowden's findings concluded that there are multiple concerns that school systems need to overhaul when evaluating teacher effectiveness for tenure purposes. The study reported that states' systems lack evidence of students' learning. States were found to lack the development of standards for teaching and implementation of these standards. Baratz-Snowden found that states need to research teacher effectiveness, specifically examining how to measure, develop, and sustain qualities of effective teaching, and integrate them into teacher evaluations.

Based on those findings, Baratz-Snowden (2009) recommended a set of criteria for granting employment based on defensible evidence. The report recommended six specific areas for states to improve: (a) professional standards, (b) a system that is multidimensional and utilizes multiple data sources, (c) collaboration between teachers and administration, (d) evidence of student learning, (e) evidence of the teachers' teaching and learning environment, and (f) professional judgment to grant and revoke tenure. Baratz-Snowden surmised that a comprehensive evaluation system must "include transparent standards and rubrics for meeting those standards, access to professional development to assist in meeting the standards, and a labor and management group that examines the evidence" (p. 4).

More recently, Danielson (2010) discussed the deficiencies of traditional evaluation systems. She determined that traditional systems used outmoded evaluative

criteria, and used "simplistic evaluative comments, such as *needs improvement*, *satisfactory* and *outstanding*" (p. 35). Additionally, similar methods were used for both novice and experienced teachers, which could feel like a punitive process that is "done to" teachers (p. 36). Lastly, Danielson stated that there is a "lack of consistency among evaluators" (p. 36). Administrators' ratings of a teacher were found to vary greatly, resulting in inequity and unreliable credibility in the evaluation process.

A decade earlier, Danielson and McGreal (2000) described the teacher evaluation process as "a meaningless exercise" (p. 7) that yields little value for teachers. The authors noted that although well intended, the systems were burdensome, outdated, and lacked a shared understanding of what constitutes good teaching. The combination of these factors imparted a culture of "passivity and protection" (p. 6). These shortcomings, Danielson and McGreal concluded, enabled teachers to continue to stagnate rather than progress and grow in their teaching methods.

On January 28, 2010, a report released by the not-for-profit, non-partisan National Council on Teacher Quality (NCTQ) found that Pennsylvania's teacher policies largely worked against the nation's goal of improving teacher quality. The national focus on teacher quality, the broad range of state laws, and the regulations that govern the teaching profession were found to impede serious reform, rather than promote reform. The report summary indicated that Pennsylvania had to address these issues in order to make meaningful decisions when granting tenure. Additionally, it was noted that there are limited requirements for permanent certification and limited criteria for advancement once a teacher is hired. The report recommended a strengthening of policies regarding teacher compensation issues, evaluation, and tenure polices. The report concluded that:

> Pennsylvania's evaluation and tenure policies do not consider what should count the most about teacher performance: classroom effectiveness. Pennsylvania does not require any objective measures of student learning in teacher evaluations. It also does not require that districts collect or consider any evidence of teacher effectiveness as part of tenure decisions. Pennsylvania makes it too difficult for districts to attempt to dismiss poor performers by failing to articulate a policy for dismissing teachers for poor performance separate from dismissal policies for criminal and morality violations. (pp. 1-2)

In 2009, the Bill and Melinda Gates Foundation concluded a three-year study, the Measuring Effective Teaching Project (MET), which focused on determining the most effective methods for measuring teacher performance. The project captured over 23,000 lessons on video tape. While the study did not specify means to conduct effective teacher evaluations, the findings surmised that teaching is too complex to be captured accurately with any single measure of performance. The MET's findings suggested that states should consider whether their evaluation systems are measuring teacher effectiveness for student growth and assessment objectives. The report provided suggestions for ensuring high quality evaluations, arguing that a more balanced set of standards needs to be considered. The report detailed that evaluations that combined several performance measures will produce accurate results. MET's five recommendations for policymakers included: (a) basing teacher evaluations on multiple measures of performance data including student academic progress, (b) more frequent and robust classroom observations, (c) use of observation rubrics for consistent results, (d) training for evaluators, and (e) use of student surveys as a component of teacher evaluations

Standards Based Model

New ideas have emerged in response to the known deficiencies in the traditional evaluation practices. Many experts agreed that in order to improve teacher instruction and increase student achievement, the current teacher evaluation systems need to be reformed (Baratz-Snowden, 2009; Danielson, 2011a; Kimball et al., 2004; Marshall, 2012; Marzano, 2012). New statistical methodologies, databases with student achievement information, and changing assessment strategies allow for new ways to identify teacher effectiveness (Stronge, Ward, & Grant, 2011). Researchers have recommended developing a comprehensive system for identifying teacher effectiveness based on student achievement (Danielson, 2011a; Darling-Hammond et al., 2012; Marzano, 2012; Papay, 2012).

Marshall (2005) summarized that traditional evaluations tend to be one dimensional. Teacher observations and evaluations have been traditionally driven by administrative observation. Koops and Winsor (2006) stated that observations, supervision, and evaluation together provide a system that encourages and motivates teachers. Koops and Winsor further suggested that the formal written evaluation that judges the overall performance of a teacher should only be one part of a larger process.

Recent demands for greater accountability by both local communities and governments have created more measurable structures for observation and evaluation (Kimball et al., 2004; Kohut et al., 2007; Toch & Rothman, 2008). Research has suggested that quality teacher evaluation systems should be based on clear, objective standards of practice; be conducted by multiple, trained evaluators; and consider multiple observations and sources of data collected over time (Donaldson & Peske, 2010;

Marshall, 2012; Toch & Rothman 2008). Papay (2012) stated that a rigorous evaluation model based on multiple measurements provides quality targeted feedback. Moreover, Papay suggested that having a valid model that assessed teacher performance is important to improve student learning, but it must also promote teacher development.

Current research has stated that more rigorous professional standards need to be developed that reflect the complexity of teaching and learning (Hightower et al., 2011; Marzano, Toth, & Schooling, 2012; MET, 2013). Both teachers and administrators need a common understanding of all the facets of teacher effectiveness, as well as the indicators and evidence that teachers must present to demonstrate that they are meeting the standards (Baratz-Snowden, 2009). Such an evaluation process can be a very potent tool for improving teaching and learning, as teachers and administrators together examine teaching practices and come to agreement on what constitutes best practice (Baratz-Snowden 2009; Danielson, 2007; Marshall, 2009).

Darling-Hammond et al. (2012) recommended a strong need for a more comprehensive teacher evaluation system. The system must have credibility with teachers, use multiple evaluators and measures, offer professional development, provide accountability for evaluators, and should be integrated with a system-wide emphasis on quality teaching. Previous to these recommendations, Borman and Kimball (2004) suggested the need for a standards-based teacher evaluation system, based on common effective teaching behaviors, with assessment using multiple, authentic sources of teaching evidence.

Many states and districts have started to examine a performance-based assessment as a means for teacher evaluation (McGuinn, 2012; Reform Support Network, 2012).

This method involves observation, but assesses the teacher's instruction against an articulated set of performance standards translated into detailed descriptors within a rubric. Danielson's (2007) *Framework for Teaching* outlined a comprehensive system so evaluators can contribute to teachers' professional learning through the use of in-depth reflective questions. Danielson's *Framework for Teaching* was first published in 1996 as an observation-based evaluation tool for beginning teachers (Danielson, 2011b). The Framework has since been extended and adapted for a teacher evaluation system that includes four broad domains of practice covering instructional planning, classroom management, instructional interactions, and professional responsibilities (Danielson, 2011b).

Borman and Kimball (2004) conducted a study in Washoe County School District located in western Nevada. The study analyzed the first generation of Danielson's *Framework for Teaching* standards-based evaluation model. The model was used to evaluate 400 teachers and assessed their performance in comparison to the district's 7,000 students' academic achievement. The authors researched whether teachers with higher or lower standards-based evaluation ratings correlated with higher or lower student achievement scores. The findings indicated that lower achieving students tended to be in classrooms with lower teacher ratings on the standards-based system. The study further highlighted that better teachers may have been assigned to classes with higher achieving students. Borman and Kimball (2004) concluded that an evaluation model that included the domains, rubrics, and sources of evidence, was a good overall measure of teaching quality.

Milanowski (2004) studied the relationship between teacher evaluation scores and student achievement in the Cincinnati Public Schools system. The study was based on a teaching evaluation score on over 3,000 teachers in more than 70 schools and programs. According to Milanowski, Cincinnati's teacher performance system models Danielson's *Framework for Teaching*. Utilizing the four domains' ratings, Milanowski combined teacher scores and linked them with student achievement results. The results indicated that teachers with higher evaluations scores correlated with higher student achievement. Milanowki's findings suggested that a teacher evaluation rating based on a standards-based model demonstrated a positive relationship with student achievement.

The two studies performed in different states explored the relationship between teacher performance and student achievement based on Danielson's *Framework for Teaching.* Both the Cincinnati Public schools' (Milanowski, 2004) and the Washoe County School District's (Borman & Kimball, 2004) studies indicated that standards-based evaluation systems can be designed to measure teacher quality and link it to student learning. Danielson (2007) suggested that by shifting the focus of evaluation from an inspection format of checking off tasks to a collaborative reflection, educators will benefit from a qualitative evaluation that focuses on growth and improvement.

Without a shared set of standards of professional practice, teacher quality is at the mercy of the observer. In such a system, it is easy for form, "didactic instruction, small group collaborative reflections, and student enthusiasm to override substance, student engagement and learning" (Baratz-Snowden, 2009, p. 11). Additionally, it is easy for inappropriate and insufficient outcome measures such as standardized student test scores to overtake the process and narrow the concept of effective teaching (Goe et al., 2008).

Collaborative Inquiry

Teaching is not an isolated activity (Leana, 2011). A growing body of research has suggested that when teachers collaborate for instructional purposes, while informed by data from their students, teacher knowledge grows and their practices change (David, 2009; Nelson, Slavit, Perkins, & Hathorn, 2008). In their research of collaboration and inquiry, Nelson et al. (2008) noted that teachers using student achievement data for the purpose of teaching and learning can transform classroom pedagogy to meet the specific needs of the students. David (2009) believed that "collaborative inquiry is among the most promising strategies for strengthening teaching and learning" (p. 88).

Forms of formalized collaborative inquiry have been contextualized in conversations among school leaders and reformers to best practice literature on professional learning, (Nelson, Deuel, Slavit, & Kennedy, 2010) and data teams (Reeves & Flach, 2011). The benefits of teacher collaboration have gained much attention from school leaders and researchers (DuFour et al., 2004, Glaser, 2005; Schmoker, 2006). Effective instructional collaborative inquiry focuses on improving classroom practices in order to improve student learning (Nelson et al., 2010). Focused work in collaborative teams, structured through professional development opportunities, have helped teachers expand their pedagogy, modify lesson planning, and utilize differentiated learning strategies that positively impact student performance (DuFour et al., 2004).

DuFour et al. (2004) believed that all students should be expected to demonstrate the same knowledge and skills, regardless of the teacher to which they are assigned. For this to occur, DuFour et al. stated that teachers must work in collaborative teams to assess student learning. Servage (2009) and Reeves (2007) believed that teachers must create

common assessment measures to determine when students have mastered essential objectives. Developing common assessments and using student achievement data to inform instruction are crucial aspects to successful, collaborative teams. Development of formative assessments provides ongoing data throughout the learning process. The data provide the professional learning community teams with information on student achievement that is used to monitor, adjust, and reinforce sound instructional practices (Reeves, 2007; Servage, 2009).

Support focusing on staff interactions, expectations, and teaming directly impacts the overall effectiveness of the school. Koops and Windsor (2005) noted that teachers who are actively engaged in conversation and inquiry with their colleagues are more energized and encourage risk taking and innovation, which promotes an environment where "continued school improvement is the goal" (p. 63). Effective team based learning communities allow schools to create a culture of teaching that empowers teams to lead their own inquiries. Critical to the impact of teacher collaboration on improving student learning, is the teacher's willingness to engage in potentially different types of collegial discussions (Nelson et al., 2010). These conversations are less about sharing "activities, information and student anecdotes, and more about raising and pursuing questions about learning goals and in their instructional practices" (Nelson et al., 2010, p. 175).

Even though many teachers show willingness to engage in collaborative inquiry, Nelson and Slavit (2008) have argued that often "complex layers of support" (p. 99) are required. While common practice in some professions, the culture of many schools does not invite feedback for fellow professionals. Schools are more likely to reflect academic autonomy. There is a tendency for educators to avoid criticizing each other's

professional practices (Given et al., 2010). Yet McTighe (2007) suggested that honest, specific peer feedback is necessary for improvement.

The act of meeting as a team does not necessarily indicate that the meeting will be productive. Collective teacher inquiry and professional learning communities require reflection and constructive criticism in order to be transformational. Working collaboratively can be a setting of tension, mistrust, and conflicting goals (Given et al., 2010). Danielson (2007) cautioned, "Leading change within one's own department or team may require considerable interpersonal skill and tact. The success of such an effort also depends on a teacher leader's having to established credibility and trust with his or her colleagues" (p. 17).

DuFour and Marzano (2009) found that when team members build trust and provide one another with ongoing evidence of progress toward their shared goals, they become effective models for empowering teacher learning and improving student learning. Fullan (2007) contended that deprivatizing teaching changes culture and practice. The transparency that develops among team members creates opportunities to share evidence of student learning. This includes sharing common assessments and lesson studies, which can serve as an influential tool to facilitate positive teacher interactions and school improvement (Fullan, 2007).

Research by Given et al. (2010) focused on engaging teachers in collaborative inquiry using the Harvard Project Zero's model. The Harvard Project Zero requires teachers to use documentation as a practice to share information for teacher learning. This model provides a variety of methods for observing, recording, interpreting, and sharing information through regular examination. Each area is specifically designed to

deepen teacher learning, through continuous conversations and reflection, in a systematic way. The researchers worked with three groups of educators in three different school districts. Using this formalized process, the researchers evaluated the inherent tensions of working collaboratively, and over time analyzed group learning as it evolved and developed into a community culture of knowledge. The researchers concluded that when teachers use a formalized process with a specific focus on the aspects of learning and practice, they adopt an inquiry model that fosters "shared local knowledge" (Given et al., 2010, p. 43) among all teachers involved.

Marzano (2003) and DuFour et al. (2004) cited numerous studies linking the benefits of teacher collaboration to teacher growth and student achievement. When teachers are afforded the opportunity to collaborate, study, and learn from one another, teachers become empowered to control their own learning and modify their instructional practices and beliefs (DuFour et al., 2004).

DuFour et al. (2004) characterized professional learning communities as groups of educators who "work together to analyze and improve their classroom practice" (p. 9). Ongoing professional learning must be integral to teacher evaluation and tenure (Toch & Rothman, 2008). Professional learning communities (PLCs) involve systematic observations and analyses of classrooms and student work and ongoing collegial dialogue (Servage, 2009). Through the process, teachers will build effective pedagogical knowledge (Wood, 2007).

Roberts' (2010) study examined general education teachers' perceptions of the PLC process and the benefits in helping increase student achievement. The study included 247 elementary, middle, and high school English and mathematics teachers

working as parts of a professional learning community in a Midwestern school district. The results of the study had a high correlation between teachers' perceptions in creating a culture of collaboration and determining growth in student achievement. Teachers reported to know which students were making progress toward meeting objectives and which students needed more support. The study found that teacher collaboration was most beneficial to improving student achievement when teachers focus their PLCs on student data and instructional strategies.

In 2011, Leana completed a five year study of 199 New York City elementary schools of more than 1,200 kindergarten through fifth grade teachers. She found that students showed higher gains in mathematics achievement both when their teachers reported frequent conversations with their peers that centered on mathematics, and when there was a feeling of trust and closeness among teachers. The study revealed that teachers who interact frequently and closely with peers were more willing to share their difficulties and struggles. Through frequent collaboration, teachers were able to identify effective classroom strategies. Teacher collaboration was a significant predictor of student achievement gains above and beyond teacher experience or ability in the classroom (Leana, 2011).

Koops and Winsor (2005) provided a basic framework for the necessary attributes in developing a professional learning community including (a) shared beliefs, vision, and values; (b) supportive and shared leadership; (c) collective learning; (d) shared practices; and (e) supportive conditions. These characteristics support an evaluation process that has instructional collaboration and professional growth as key components and promote a

school culture where continuous school improvement is the goal (DuFour, 2004; Koops & Winsor, 2005; Wood, 2007).

Multiple Measures of Appraisals

Assessments that use multiple measurements and multiple evaluations are part of Donaldson's (2009) recommendations. She also said that valid and reliable instruments are needed for making those evaluations and that evaluations should be used to provide meaningful feedback to teachers on their performance instead of being used as they have been historically, for acclamation and inspiration. Marzano et al. (2012) supported Donaldson's recommendations and noted that a highly effective teacher evaluation model must reflect the learning process. An effective evaluation should hold teachers accountable while encouraging teachers to develop and use best practice pedagogy. The results of the Marzano et al. study found that the best method for identifying effective teaching is by using multiple measures.

Towe (2012) suggested that teacher evaluation should not be contingent solely on students' standardized test scores and advocated for attaching student growth measures from a variety of sources to validate teacher effectiveness. Towe noted, "Teachers should be evaluated over a period of time following instructional practice in order to determine student growth as a result of that instruction" (p. 157). Towe believed that multiple measures must be considered; for example, teacher collaboration, teacher portfolios, self-reflection, and self-evaluation, as well as student growth.

The objective of evaluating teachers should be to collect information that will be effective in developing appropriate strategies to improve instruction and therefore improve student achievement (Goe et al., 2008). The authors further suggested that this

approach to improve instruction should be meaningful and may include such measures as: (a) professional development, (b) individual work with curriculum specialist, (c) college coursework, and (d) study teams. Danielson (2010) supported this concept:

> If we want to design teacher evaluation systems that teachers find meaningful and from which they can learn, we must use processes that not only are rigorous, valid, and reliable, but also engage teachers in those activities that promote learning namely self-assessments, reflection on practice, and professional conversations. (p. 38)

The MET (2013) project suggested the use of classroom observations, student surveys, and student achievement as part of the teacher evaluation process. The study gave evidence that evaluation systems provided teachers with meaningful feedback. That feedback enabled teachers to grow instructionally and improve student achievement. The study affirmed that multiple measures of student achievement, in conjunction with classroom observations, provide the most valid method to measure teacher performance (MET, 2013; U.S. Department of Education, 2009).

Additionally, MET (2013) recommended a formative evaluation process that encourages teachers to reflect on student achievement with evidence from multiple data sources. MET suggested that such a system would ensure a high level of accountability in an environment that is supportive and encouraging, and would provide a catalyst for professional development and improved student achievement.

A standards based teacher evaluation practice that uses multiple measures of appraisal is being considered by schools as a method to improve instruction and accountability (Kane & Cantrell, 2012). Baratz-Snowden (2009) recommended the

development of rigorous standards of practice that reflect the pedagogy of teaching. Moreover, standards need to be research-based and specific to assist teachers in improving classroom practices. Baratz-Snowden concluded that developing defined standards of practice for teacher evaluation would enable teachers to have a tangible knowledge base for known expectations.

Baratz-Snowden (2009) recommended teachers, as stakeholders, need to have input in the decision-making process when developing evaluation systems. Baratz-Snowden stated that teachers have critical knowledge about best practice that can assist in designing the system. Further, she suggested that teachers should be involved in all aspects of the process, "from creating the standards, to articulating the rubrics of best practice and evidence of student learning, to identifying professional development" (p. 15).

Teacher Value-Added Model

A method many states have enacted to measure teacher effectiveness and improve teacher performance is the Value Added Model (VAM). Developed by statistician William Sanders (Sanders, 1994), the Education Value-Added Assessment System (EVAAS) has been adopted in several states (Amrein-Beardsley, 2008; Anderman et al., 2010; Darling-Hammond et al., 2012; Green, Baker, & Oluwole, 2012). Currently, the most influential VAM is the Tennessee Value Added Assessment System (TVAAS). Goe et al. (2008) stated:

> The increased availability of data in which student achievement is linked to teachers, along with statistical innovations in analyzing these data, may be partly

> responsible for what appears to be a growing emphasis on measuring teachers' contribution to student achievement. (p. 6)

Twenty-one states, including Colorado, Ohio, and Pennsylvania, are experimenting with or using a form of the TVAAS model for teacher evaluations (McGuinn, 2012; MET, 2013; Reform Support Network, 2012). This method uses a statistical probability growth score based on all prior achievement data for each student ("PVAAS," 2013). Beginning in the 1992-1993 school year, the statistical advances developed by Sanders made it possible to look at student achievement gains after adjusting for student and school characteristics (Sanders, 1994). Baker et al. (2010) argued that these approaches "measure growth using a fairer comparison of teachers than judgments based on their students' test scores at a single point in time or comparisons of student cohorts that involve different students at two points in time" (p. 2). For teachers, the projections for the current year offer a basis for differentiated instruction, identifying students needing an academic intervention and determining a student's readiness for advanced coursework (Anderman, Anderman, Yough, & Gimbert, 2010).

Sass (2008) described the VAM as an analysis growth score that measures teacher quality by "estimating the impact of teachers on student achievement, holding constant other factors that affect current student performance, including contemporaneous student ability and effort" (p. 1). The Pennsylvania Value Added Assessment System (2013) specified that the growth scores are based on certain assumptions. First, students will have a typical school experience; and secondly, the method will not take into account the effectiveness of a specific curriculum, assessment, and instruction in the current year. Rather, student growth scores reflect what the student is likely to score on summative

state assessments at the end of the current school year. Green et al. (2012) summarized that the VAM methodology for evaluating teachers by using numerous years of data of student achievement scores generates a stable estimation of a teacher's contribution to the students' achievement.

Using the VAM to measure teacher effectiveness is controversial (Amrein-Beardsley, 2008; Chetty, Friedman, & Rockoff, 2011). Anderman et al. (2010) stated that questions have emerged regarding how accurately it can measure consecutive years of student growth based on state achievement tests. There are assumptions in estimating a teacher's effectiveness. Amrein-Beardsley (2008) studied concerns about validity and the potential bias due to missing or inaccurate student test data. Amrein-Beardsley found that the VAM cannot accurately account for linking teachers to the actual subject taught and the percentage of the instructional time to individual students. In response, Sanders and Wright (2008) acknowledged concerns about "problems identifying which teachers taught which students" (p. 2), but contended that the those issues have been addressed by requiring teachers to identify and certify student rosters.

Anderman et al. (2010) supported Sanders and Wright (2008) and argued that using the value-added model can have a positive effect on student achievement if the data provided are used to influence instruction. Furthermore, Anderman et al. postulated that it provides a clear analysis of individual students' growth over time. This information offers teachers the opportunity to monitor student progress and could help teachers better align their instruction. Chetty et al. (2011) agreed that the VAM provides useful data for evaluation teachers. However, both Chetty et al. and Anderman et al. stated that the

value added growth score should not be used in isolation, and should be considered only a factor in the evaluation process.

Chetty et al.'s (2011) study of VAM tracked one million children from urban school districts from fourth grade to adulthood. The purpose of the study was to answer two questions. First, does the value-added approach accurately measure teachers' impact on student test scores or does it unfairly penalize teachers who may be assigned lower achieving students? Secondly, do high value-added teachers improve students' long term outcomes or are they better at teaching to a test? The results of the study showed that teachers rated highly on the value-added system created greater value and that test scores are helpful in identifying such teachers. It gave evidence that the value-added measures are informative about a teacher's long-term impact on student achievement. The study bolstered the belief in the use of the value added model for teachers effectiveness. Chetty et al. concluded "that the Value Added Model is a good proxy for a teacher's ability to raise students' test scores" (p. 36).

Chetty et al.'s (2011) earlier research contradicted Kupermintz's (2003) study. Kupermintz examined the Tennessee Value-Added Assessment System (TVAAS) model for its validity in predicting teacher effectiveness for evaluation purposes. The study concluded that TVAAS could not determine the effect that an individual teacher has on a student's learning growth. Suggesting that there are too many factors of bias that could be expressed in a statistical formula, Kupermintz noted that multiple teachers in different subjects have an impact on a student's achievement. A more recent report by Pearson Education (2008) supported Kupermintz's findings. Pearson Education specified that using averages introduces bias for teachers with small numbers of students or for students

with missing data, and found that "as a result, effective teachers may be undervalued while an ineffective teacher might escape notice" (p. 10).

A report prepared by Darling-Hammond et al. (2012) for the American Federation of Teachers reached the conclusion that the VAM for teacher evaluation purposes was unstable. The report stated that teacher ratings are significantly affected and influenced by student variables outside the teacher's control. Darling-Hammond et al. cautioned against the practice of "basing individual teacher evaluations on annual student test scores" (p. 20). The report recommended that such measures should be used only when scores are part of an integrated system with "multi-faceted evidence" (p. 38) supporting teachers' practices.

Performance Rubric

Many researchers suggested the use of rubrics as a measuring instrument as part of teacher evaluation reviews (McClellan et al., 2011; MET, 2013; Papay, 2012). The Center for Education Policy and Practice (2010) study on the Massachusetts Teachers Association Evaluation Model (MTA) recognized that the rubrics align with both Danielson's and Marshall's teacher evaluation rubrics. These rubrics described the critical behaviors that should inform decisions about the standards. The study suggested that district evaluation systems adopt rubrics as long as they describe essential activities related to the states' evaluation standards.

The focus of the Colorado Model Evaluation System for Teachers has been the development of the rubric. This standards-based instrument provides descriptions of professional practices for each performance rating level (Colorado Department of Education [CDE], 2013). Evaluators rate teachers on each element of each standard and

then the ratings to determine the overall rating on professional practices. At the end of the rubric for each standard is a section for summarizing individual element ratings for the standard. Summary ratings are also included in the summary rating worksheet and the one-page summary evaluation sheet, where the evaluator and the teacher being evaluated will review all standard and element ratings and determine the overall professional practices rating. The state of Colorado received feedback from over 2000 teachers to help shape its professional practice rubrics (CDE, 2013).

Both Marshall (2006) and Danielson (2007) considered rubrics to have distinct advantages over traditional evaluation methods. The traditional approach to evaluating teachers does not improve student achievement (Marshall, 2006). Rubrics are tools that enable the principal to give teachers clear feedback with respect to a standard expectation. They provided a common reference for communicating about substandard teaching practice. The rubric design is organized into a framework of six domains of planning and preparation, classroom management, instruction, monitoring and assessment, community outreach, and professional responsibilities (Marshall, 2006). This framework is intended to promote teachers' development by identifying techniques for assessing certain aspects of pedagogy (Danielson, 2007).

"A standards-based evaluation system requires rigorous instructional standards with clear rubrics that define success" (Papay, 2012, p. 135). Papay (2012) stated that evaluations are not one size fits all. Districts should have clear expectations about the level of evidence required to make an assessment and the extent of feedback to provide to teachers. Papay (2012) further explained that a rigorous evaluation based on a rubric

should incorporate evidence from multiple sources throughout the year to ensure reliability.

McClellan, Atkinson, and Danielson (2011) made recommendations based on the MET (2013) project results that rubrics must have observable order, clear definitions, and descriptions to reduce bias. Additionally, evaluators must have thorough training to build understanding and scoring consistency. During the MET (2013) study, researchers found that it was difficult to evaluate teachers accurately using classroom observations alone. The researchers used rubrics and conducted multiple observations of each teacher. They found that observation ratings did correlate somewhat with student achievement data.

Rubrics are important for generating a shared image of what good teaching and learning looks like and sounds like. Rubrics are not checklists for clinical supervision or walkthroughs (Marshall, 2006). They are best used to pull together all the impressions gathered in the course of a school year. Saphier (2012) referenced evaluation rubrics as useful because they are an observational standard of practice and expectation. Additionally, they invite conversation between the evaluator and teacher about what good practice means, looks, and sounds like in real instruction. That in turn can lead to teacher engagement and specificity about what teaching behaviors are expected (Saphier, 2012).

Peer Observation

Rubrics developed by Danielson (2011b) provide a framework for focusing observations and feedback through peer and mentor observation cycles. Danielson believed that teachers benefit from supportive, classroom-experienced evaluators. Peer or mentor evaluators have insight that provides a reliable picture of a teacher's effectiveness in the classroom (Danielson, 2011b).

Research by Kohut et al. (2007) on peer observations and their effectiveness as an important component of teacher tenure concluded that teachers and administrators find them valid and useful. The goals of the study were to compare the perceptions of observers and those observed regarding the process of peer observation, the reporting of peer observations, the usefulness of peer observation as an evaluation tool, and whether either group felt that the process improves teaching effectiveness. Kohut et al. investigated 80 untenured teachers and 143 tenured teachers and found that both observers and teachers reported that they valued the process. The results further showed that the faculty expressed trust in the process. The study revealed that teachers were willing to offer both constructive and critical comments in peer observation reports. The process further received positive perceptions from teachers regarding the validity and reliability of the peer observation reports. Recommendations concluded that in order for the peer observations to be effective, they must employ data that are perceived to be both valid and reliable (Kohut et al., 2007).

Kohut et al. (2007) explained that "only collectively does faculty have the experience and standards that are both credible and useful to individual faculty" (p. 19). They recommended that peer evaluation needs to be an essential element in any teacher evaluation system and argued that professionals need shared practice and an honest dialogue among people in the profession. MET (2013) researchers determined the process for peer evaluations is more reliable if several different people evaluate the same teacher, even for shorter periods over a school year, rather than one evaluator making several trips to the same classroom.

Danielson (2011b) and MET (2013) cited that peer observations have the means to capture dimensions of teaching such as student rapport and student engagement that are not typically accounted for through other evaluation processes. Teachers who have the opportunity to receive feedback, through classroom observations, by their peers in a supportive environment can be a beneficial component of a comprehensive, multi-faceted summative assessment of teaching.

Summary

The reform movement of the last four decades has placed an emphasis on greater teacher accountability and student achievement (Furtwengler, 1995). States have been reforming their teacher evaluation models to identify highly qualified effective teachers based on their impact on student achievement (Little et al., 2009). Numerous studies have discussed deficiencies of the traditional evaluation methods. Concerns over the value of teacher evaluations include the failure to adequately differentiate teachers' effectiveness, outmoded evaluative criteria, simplistic rating systems, and a lack of consistency (Danielson, 2010). Traditional evaluation systems have been recognized as subjective judgments relying on infrequent or brief formal classroom observations (Danielson & McGreal, 2000; Marshall, 2009; Marzano, 2012).

An effective system of teacher evaluation accomplishes two things: it ensures quality teaching and it promotes professional learning. The quality of teaching is the single most important determinant of student learning. A school district's system of teacher evaluation is the method by which it ensures that teaching is of high quality (CEPP, 2010). Standards-based models of teacher evaluations, such as the *Framework for Teaching* by Danielson (2007), support the literature that quality teacher evaluation

systems should be based on clear, objective standards of practice (Borman & Kimball, 2004; Marshall, 2009; Toch & Rothman, 2008). Teacher evaluations must be grounded in research and be reflective of generally accepted definitions of good teaching. When used effectively, observation, supervision, and evaluation provide a vehicle for encouraging and motivating teachers (Danielson & McGreal, 2000). This can help teachers gain realistic appreciation of their abilities and make administrative decisions about teaching assignments and evaluations (Marshall, 2009). The evaluation system must include transparent standards and rubrics for clear teacher expectations and provide feedback to guide professional development that will assist teachers in meeting the standards (Danielson, 2010).

Chapter Three – Methods and Procedures

Introduction

This qualitative study explored teacher perceptions on the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model. The study focused specifically on the Pennsylvania's Teacher Effectiveness evaluation model's influence on teacher instruction, collaboration, and professional growth. This research study identified and evaluated data utilizing three components; Likert survey questions, open-ended questions, and interviews. This chapter describes the design of the qualitative study by examining the setting and participants, as well as the research instruments utilized to obtain data.

Setting

This qualitative study was conducted in four public school districts that serve students from kindergarten to 12th grade. Two of the school districts selected are suburban districts in southeastern Pennsylvania. One school district is a rural district located in eastern-central Pennsylvania. One school district is an Intermediate Unit in Pennsylvania. The school districts are identified by the letters: A, B, C, and D.

School District A is one of Pennsylvania's 29 Intermediate Units. The Intermediate Unit has participated in a phase of Pennsylvania's Standards-Based Teacher Effectiveness model. The Intermediate Unit has multiple education sites and a range of programs including educational levels from pre-school to the secondary level. The Intermediate Unit administers approximately 25% of the special education services offered in its county and serves nearly 3500 students from ages 3 through 21 years of age. Programs include a Child and Career Development Center, alternative-education

programs, and a number of secondary public high schools specializing in Career and Technical Centers. Serving grades 9 through 12, the four secondary schools in the Intermediate Unit offer educational opportunities in career and technical education programs and academic courses. Each school has a variety of academic course offerings, which include direct instruction, career instruction, and college preparation and honors level classes. Students have the option to attend on a full time or part time basis. Approximately 70% of students continue their education either at a four-year college, a two-year college, or a trade/technical school. Information regarding employability and interview skills are taught to the students going directly into the workforce.

School District B is a rural district located in eastern-central Pennsylvania. The district encompasses approximately 125 square miles and serves a community of approximately 18,000. The district approved a budget for the 2013-14 school year of $37.5 million with more than $14 million in state funding. The district provides educational services to approximately 3,000 students and employs 224 teachers. The district is comprised of one high school, one middle school, and four elementary schools. The district's curriculum offers college preparatory, general level, and vocational curricula. The district also offers honors and Advance Placement courses. Approximately 72% of the students continue their education either at a four-year or two-year college. The district made Adequate Yearly Progress for the 2011-12 school year. Adequate yearly progress for the district indicated that 81% of the student population were proficient in reading and 81% were proficient in mathematics. The average attendance rate was 95% and the graduation rate was 95%.

School District C is a suburban school district located in southeastern Pennsylvania, with a community of approximately 35,000 people. The district encompasses approximately 21 square miles. There are approximately 3,500 students enrolled in grades kindergarten through 12. The district approved a budget for the 2013-14 school year of $76 million with more than $10 million in state funding. The district is comprised of four elementary schools, one middle school, and one high school. Courses are offered at the College Preparatory and Advanced levels. Special Education courses are offered in Life Skills, direct instruction, and co-teaching classroom environments. Advanced Placement courses are offered in all the academic areas. Additionally, a variety of opportunities are available to students through elective courses, independent study programs, and a wide range of extra-curricular activities. In 2012, the district did not make adequate yearly progress and was placed on warning status due to its high school's special education subgroup and the district's graduation rate. Adequate yearly progress for the district indicated that 84% of the student population were proficient in reading and 84% in mathematics. The average attendance rate was 96% and the graduation rate was 87%.

School District D is a suburban school district located in southeastern Pennsylvania. Geographically, the district spans approximately 73 square miles and serves a diverse socioeconomic community. It comprises three townships and three boroughs. The district approved a Budget for the 2013-14 school year of $95 million with more than $20 million in state funding. There are approximately 5,500 students enrolled in kindergarten to 12th grade. The district is comprised of six elementary schools, two middle schools, a freshman center, and one high school. The district offers

a cyber academy from kindergarten to 12th grade. The district's 2011-2012 operating budget of $88.8 million supports the district's educational programs. It offers a comprehensive educational program for all levels of learners and has adopted a standards-based academic and grading program. At the secondary level the district offers honors, Advance Placement, dual enrollment college courses, cyber, and engineering classes. In 2012, the district did not make adequate yearly progress and was placed on warning status due to subgroups including Latino/Hispanic, Special Education, Economically Disadvantaged, and the district's graduation rate. Adequate yearly progress for the district indicated that 82% of the student population were proficient in reading and 86% were proficient in mathematics. The average attendance rate was 96% and the graduation rate was 91%.

Subjects

Participants in the study were teachers who were teaching in the selected schools and had participated in the pilot of the Pennsylvania Teacher Effectiveness Model. All participants had completed at least one year of teaching. Teachers from all subject areas were included in the study. Participating teachers were highly qualified and maintained active Pennsylvania State Teaching Certification. During the time of this study, the Pennsylvania Teacher Effectiveness model completed the last year of the pilot phase. Each district identified a pilot group of teachers from their staff to participate in the Pennsylvania Teacher Effectiveness pilot. The number of participants varied based on the district's pilot group. Each district's pilot group was comprised of teachers from the elementary to secondary grade level. District A had 15 teachers participate in the pilot. District B had 20 teachers participate in the pilot. District C had 42 teachers participate

in the pilot. District D had 30 teachers participate in the pilot. A total of 117 subjects qualified for the research study. A total of 35 teachers participated in the Likert survey and questionnaire and 13 teachers participated in the interview. Participation in all phases of this research study was voluntary.

Instruments

This study employed a qualitative method of inquiry that provided data to analyze secondary teachers' perceptions of how the Pennsylvania standards-based Teacher Effectiveness model would impact teacher instruction, collaboration, and professional growth. Information was gathered from teachers' currently instructing in different grade levels and having varied years of experiences. This study utilized a survey consisting of Likert-scale and open-ended responses. Questions developed for the Likert-scale and open-ended responses were generated based on a thorough review of the literature. Participants who completed the survey were invited to consider taking part in interviews.

Survey. Marshall and Rossman (2011) stated that surveys are an appropriate tool for data collection when making inferences about a large group based on a small number of respondents. The survey (Appendix A) was comprised of a Likert scale of 25 questions and three open-ended questions. The instrument presented items with a Likert response set at a four point scale from strongly agree to strongly disagree. The Likert scale allowed the respondents to indicate the extent to which they agreed or disagreed with each statement. The survey was provided to the participants through an email link to Surveymonkey.com.

For the purpose of this study, three open-ended questions were added to obtain a more detailed and descriptive view of teachers' perceptions of the Pennsylvania's

Teacher Effectiveness evaluation model. Teachers were asked to complete the questions based on their own thoughts and perceptions of the Pennsylvania's Teacher Effectiveness evaluation model and its impact on instruction, collaboration, and professional growth.

Interviews. According to Merriam (2009), interviewing is necessary when the desired behavior is not observable. Teachers who completed the online survey were invited to participate, through voluntary consent, in an interview with the researcher. A list of questions (Appendix B) was developed to elicit expandable responses regarding key concepts associated with the survey questions. The interviews lasted approximately 15 to 20 minutes. The participants were asked identical questions and their responses were audiotaped with permission for accuracy purposes. The list of five interview questions was structured using a standardized open-ended interview. The open-ended nature of the interviews allowed participants to contribute as much detailed information as desired. It also permitted the researcher to ask probing questions as a means of follow-up. Gall, Gall, and Borg (2007) believed that a standardized open-ended interview reduces the researcher's bias within the study when the interviewing process involves multiple participants.

Reliability and Validity

An instrument must be reliable to be trusted. Reliability is the ability of an instrument to be consistent in producing similar results when characteristics being used are unchanged (Creswell, 2009; Merriam, 2009). Merriam (2009) contended that that repeating identical measures improves validity of the study establishes consistent results. To ensure validity the survey, open-ended questions, and the interview were conducted using identical instructions, guidelines, and questions. The instruments used for this

study were first piloted by a group of educators to evaluate the questions and validate the effectiveness of the instrument

The method for collecting the data for the study was designed to address validity through triangulation. Triangulating different data sources of information builds credibility for the methods of data analysis. Triangulation emphasizes the use of multiple methods of data collection for analysis. Creswell (2009) recommended triangulation to examine evidence to build justification for emerging themes. This study was designed with multiple methods. The use of interviews, survey, and open-ended questions ensured triangulation. Marshall and Rossman (2011) emphasized the importance of triangulation because using multiple methods strengthens the study's usefulness for other settings. A descriptive analysis was used to present and articulate the data collection. The data were analyzed and compared to broaden the depth and scope of understanding of the research.

Design of the Study

A qualitative study was chosen for the research design because the intent of this study was to describe teachers' perceptions of the effectiveness of the new standards-based teacher evaluation process by qualitatively studying its impact on instruction, collaboration, and professional growth. A qualitative mode of inquiry was an appropriate design to measure the teachers' perceptions of the Pennsylvania's Teacher Effectiveness evaluation model on instruction, collaboration, and professional growth. Marshall and Rossman (2011) suggested that a qualitative method is best used to help study the variations of complex human behavior. Using a qualitative method will enrich the results of the study with the participants' words and actions, to build a conceptual framework that will lead to patterns of understanding.

Procedure

Prior to conducting the study, the researcher gained written permission from the Superintendents of the districts where the study was conducted. After receiving permission from the Superintendents of Schools, the researcher developed the survey and interview questions. The survey and interview questions were piloted by a teacher panel to validate the effectiveness of the instruments. The panel of teachers were asked to review the items on the questionnaire and make suggestions and/or recommendations. The researcher elicited information from the responses to test for clarity and understanding of the survey and interview questions. Feedback related to the design of the instrument served as documentation that addresses the content aspect of validity (Merriam, 2009).

After receiving permission to conduct the study at the proposed school districts, the researcher requested approval from Immaculata's Research Ethics Review Board. Once approval was granted from Immaculata's Research Ethics Review Board (Appendix C), the researcher again contacted the Superintendents from each study site. The researcher explained the study, assuring that participation was voluntary and that confidentiality would be maintained. The researcher requested permission for the survey instrument to be sent to the appropriate faculty within each study site.

The overview of the study was introduced by the researcher through the Superintendent's office to all eligible participants. All eligible participants were invited to participate in this study through an invitation in an email. Participants were able to link to the survey through the email invitation. Potential participants were asked to read

an explanation of the study and sign an embedded consent form. Teachers had three weeks to complete the survey, for inclusion in the study.

Potential participants were invited to take part in the interview phase of the study through the original solicitation using the school's list serve. Those who completed the questionnaire were invited to participate in a follow up interview at the end of the questionnaire, with an invitation to contact the researcher through email. Teachers willing to participate in the interview process were contacted through email and/or phone to arrange an interview time. Interviews were scheduled at the subjects' convenience, and took place at a location that was mutually agreed upon.

The teachers were interviewed using a standard open-ended set of questions designed by the researcher to encourage discussion. An audio recorder was used with the teachers' permission to accurately capture the dialogue and collect data. Participants in the interview were presented with a letter overviewing the study and a consent form assuring participants that their responses would be recorded and confidentiality would be held in strict regards.

Results from the surveys were tallied and analyzed. Interviews were transcribed and analyzed to identify patterns in teachers' perceptions. A summary of the findings will be shared with the district.

Data Analysis

The results of the data analysis were used to describe the perceptions that teachers have regarding the impact of the Pennsylvania standards-based Teacher Effectiveness model. The data collected were used for comparison and analysis. The data analysis

identified common trends and themes that impacted instructional practice, teacher collaboration, and professional growth.

Open-ended questions on the survey provided detailed information to support information gained from the Likert survey. Responses from this portion elicited an opportunity to gather more in-depth information and authenticate findings for the survey. Questions asked of the interviewees were open-ended to promote discussion. The information provided was used to gather insight into attitudes toward the teacher evaluation system. The data gathered were analyzed in conjunction with the survey questions.

Summary

The purpose of this study was to gain teachers' perceptions towards the effectiveness of the new standards-based teacher evaluation process by qualitatively studying its impact on instruction, collaboration, and professional growth. The schools selected for the study are from Pennsylvania school districts and an Intermediate Unit that had participated in the pilot phase of the Pennsylvania's Standards-Based Teacher Effectiveness program. Teachers were invited to participate on a volunteer basis. Teachers had to have successfully completed one full year of teaching and have participated in a pilot study. This study utilized a survey consisting of a Likert-scale and open-ended questionnaire. Additionally, interviews were conducted with those subjects who gave consent. The reliability and validity were established through the triangulation of information gather by multiple methods. This produced data rich results that were collected for comparison and analysis. The results of the study are presented and described in Chapter Four.

Chapter Four – Results

Introduction

This qualitative research study investigated teacher perceptions on the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model. The purpose of this study was to gain teachers' perceptions towards the effectiveness of the new standards-based teacher evaluation process by qualitatively studying its impact on instruction, collaboration and professional growth.

This chapter reports the findings of the teachers' responses to the Teacher Evaluation Survey (Appendix A) which consisted of 25 Likert scale responses and three open-ended responses, and an interview (Appendix B) which consisted of five questions. Four school districts were chosen to participate in the study based upon their participation as part of the Pennsylvania Teacher Effectiveness pilot. Each district had a sampling of their staff participate in the pilot group. A total of 35 teachers participated in the Teacher Evaluation questionnaire. Thirteen teachers from two of the four districts participated in interviews with the researcher. Participants' responses were analyzed to identify trends in order to answer the study's three research questions.

The survey contained three questions which addressed the demographic data of the subjects. Information gathered was related to the (a) number of years teaching, (b) levels of education, and (c) years in current position. Table 4.1 depicts the demographic data of the subjects for the study.

Table 4.1

Subject Demographic Data

Number of Years Teaching Answer Choices	Responses
1 to 3 years	0% (0)
3 to 5 years	6% (2)
6 to 10 years	20% (7)
11 to 15 years	23% (8)
16 or more years	51% (18)
Level of Education Answer Choices	**Responses**
Bachelor's	6% (2)
Bachelor's plus 15	11% (4)
Master's	34% (12)
Master's plus 15	14% (5)
Master's plus 30	26% (9)
Doctorate	9% (3)
Number of Years in Current Position Answer Choices	**Responses**
1 to 4 years	20% (7)
5 to 9 years	29% (10)
10 to 14 years	26% (9)
15 to 19 years	17% (6)
20 to 15 years	6% (2)
26 or more years	3% (1)

Note. N=35.

Analysis of Data Results

Research Question One. *What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on their instructional practice?*

Research Question One was addressed utilizing data from the survey responses and interview data. Likert-scale questions 4 through 12 and open-ended response question #26 addressed this research question. The survey was designed to ascertain the participants' perceptions about the impact of the teacher evaluation process on their instructional practices. Additionally, interview questions 1, 3, and 4 asked participants to describe how their instructional practices were impacted, influenced, and/or changed due to the Pennsylvania's teacher evaluation process.

Analysis of survey responses. To determine teachers' perspectives regarding the impact of the teacher evaluation process on their instructional practices, a Likert scale was provided to allow the respondents to indicate the extent to which they agreed or disagreed with each statement. The four-point response scale ranged from strongly agree to strongly disagree. Reflecting on the entire evaluation process, teachers gave their perceptions of the evaluations system's impact on their instructional practices.

In response to the statement that the evaluation process is an effective model to determine the quality of teaching, 66% (23 out of 35) of teachers agreed with 8% (3 out of 35) strongly agreeing. In response to the belief that the evaluation process provided useful feedback for planning and instruction purposes, 60% (21 out of 35) of the teachers agreed with 17% (6 out of 35) of teachers strongly agreeing. The responses were more evenly balanced to the statement that the teacher effectiveness model provided a more

accurate observation of instruction, 43% (15 out of 35) of the teachers agreed, with 11% (4 out of 35) strongly agreeing for a total of 54%. Conversely, 37 % (13 out of 35) disagreed and 8% (3 out of 35) strongly disagreed with the statement.

Teachers were divided with their responses to the statements that they changed instructional strategies or increased their planning and preparing time for lessons since the implementation of the new effectiveness model. Both questions received the same responses with 5% (2 out of 35) strongly agreeing and 42% (15 out of 35) agreeing. A slight majority of 48% (17 out of 35) disagreed, and only 3% (1 out of 35) strongly disagreed with the statement. In response to teachers' use of student data for the purpose of planning instruction, 14% (5 out of 35) strongly agreed. The majority, 54% (19 out of 35), agreed with the statement. Thirty one percent (11 out of 35) disagreed with the statement and no one strongly disagreed.

Responding to the statement that the evaluation rubrics provide guidelines for developing lessons, only 9% (3 out of 35) strongly agreed. The majority, or 51% (18 out of 35), agreed. Forty percent (14 out of 35) disagreed with the statement. No teacher strongly disagreed with the statement. In the statement that asked teachers if the teacher evaluation system requires them to utilize a variety of instructional strategies, 9% 93 out of 35) strongly agreed. Overwhelmingly, 83% (29 out of 35) agreed with the statement. Only 9% (3 out of 35) disagreed, with no teachers' strongly disagreeing with the statement. Lastly, teachers were asked to respond to the statement that the teacher evaluation had no impact on their instructional practices. Only 6% (2 out of 35) strongly agreed and 23% (9 out of 35) agreed with the statement. The majority of the teachers, 54% (19 out of 35), disagreed with the statement. Fourteen percent (5 out of 35) of the

responding teachers strongly disagreed. Table 4.2 provides a summary of the responses teachers provided regarding their perceptions of the impact of the evaluation process on their instruction.

Table 4.2

Teachers' Perceptions of the Impact of the Evaluation Process on Instructional Practice

	Participant Response			
Statement	SA	A	D	SD
4. Is effective in determining the quality of my teaching.	9% (3)	66% (23)	20% (7)	6% (2)
5. Has provided useful feedback on my planning and instruction.	17% (6)	60% (21)	17% (6)	6% (2)
6. Provides a more accurate picture of my teaching ability.	11% (4)	43% (15)	37% (13)	9% (3)
7. Has changed the instructional strategies in classroom.	6% (2)	43% (15)	49% (17)	3% (1)
8. Has increased the time required to plan lessons and prepare for classes'.	6% (2)	43% (15)	49% (17)	3% (1)
9. Encourages use of student achievement data to plan my instruction.	14% (5)	54% (19)	31% (11)	0%
10. Provides guidelines for developing future lessons through rubrics.	9% (3)	51% (18)	40% (14)	0%
11. Requires utilization of a variety of instructional strategies.	9% (3)	83% (29)	9% (3)	0%
12. Has had no impact on instructional practices.	6% (2)	26% (9)	54% (19)	14% (5)

Note. N=35. SA= Strongly Agree; A=Agree, D=Disagree, SD=Strongly Disagree

The overall analysis of the questions showed that the majority of participants (92%) agreed that the teacher evaluation process encouraged teachers to use a variety of instructional techniques. The majority of teachers also agreed that the teacher evaluation

model was able to determine the quality of teaching (77%), provide useful feedback (77%), and use student data for planning instruction (68%). Additionally, the majority of teachers also agreed that the evaluation rubric was a beneficial guide for developing lessons (60%). The majority of teachers also reported that the new evaluation gave a more accurate picture of instructional practices (54%). However, the results were evenly balanced regarding the teacher evaluation changing teachers' instructional strategies. The results were varied on whether the teacher evaluation process required teachers to increase their time to plan lessons and prepare for classes. Overall, more than half of the participants agreed that the teacher evaluation process had impacted their instructional practices.

Analysis of open-ended responses. In addition to the Likert-scale survey items, the researcher gathered additional data through the use of open-ended question #26. The question asked participants how the PA standards-based teacher evaluation impacted their classroom instruction. Thirty-four participants responded to the question. One participant responded with "not applicable."

Eighteen respondents, or 51%, responded positively specifying that they thought the PA standards-based teacher evaluation process impacted their classroom instruction. Respondents commented that the Danielson Framework enabled teachers to self-reflect on instructional practices and utilize student data to guide teacher planning and instruction. One participant stated, "It is an impetus for providing self-reflection as an educator. Prior to this process, I mostly relied on my own past teaching practices." Similarly, another participant commented, "I am more aware of how each activity I plan affects the overall outcome and whether or not it achieves my goals." Another

participant commented, “The evaluation has made me stop and think about what is relevant and what is not when planning instruction.”

Participants acknowledged utilizing student data for planning purposes. Several participants recognized the use of formative assessment for instructional purposes. A participant stated the PA standards-based teacher evaluation “…has encouraged me to use formative assessment to guide my instruction.” Likewise, one participant stated, “The PA standards-based teacher evaluation has assisted me in better incorporating formative assessment into my overall lesson planning and implementation of the lessons.” Another acknowledged, “It is a lot of work, but it has definitely made me analyze data more deeply than before to pull small groups and remediate students.”

Sixteen respondents, or 46%, specified that the PA standards-based teacher evaluation had little to no impact on their classroom instruction. Many of the participants commented that their instructional practices were already student-focused, reflective, and pedagogically sound. One participant remarked, “I do not believe the evaluation has had a direct connection to my classroom instruction. I have always been a dynamic individual who encourages students and develops instructional practices with colleagues based on student strengths and needs.” Another participant shared,

> My direct classroom instruction and interaction with students has not changed with this … I do have a lot of my own strategies and techniques that I have perfected for use in my classroom. As for reflecting on lessons, I hope all teachers do so always. I have ever since I started teaching, but I learned this technique from a leadership training course which complements my educational training.

Participants also identified professional development as a more effective means to improve classroom instruction. One participant stated, "I don't feel it has impacted my instruction nearly as much as my professional development committees I am involved in." Likewise, another participant commented, "It [teacher evaluation] seems to be such a snapshot of my teaching. I use professional development opportunities to improve my instruction."

Analysis of interview data. Interview response questions provided further insight into the first research question. Thirteen teachers from two of the four districts participated in individual interviews with the researcher to further explore their perceptions of the impact of the PA standards-based teacher evaluations on classroom instruction. Three interview questions asked teachers to reflect on (a) the Danielson Framework and its influence on teaching, (b) building-level student data and their impact on instruction, and (c) feedback from the evaluation on instructional practices.

Each participant identified that the Danielson Framework influenced their teaching and instruction. All 13 participants indicated that the Danielson Framework helped focus their instructional goals and provided the time for lesson planning and reflection. Participant P#10 commented, "I think [The Danielson Framework] has probably given me more time to reflect on my teaching…It's given me more awareness to reflect on my teaching." Participant P#1 elaborated, "[The Danielson Framework] made me focus more on my goals for getting what I wanted to accomplish." Participant P#3 found the Danielson Framework was beneficial for lesson planning and reflection. Participant P#3 described using the Danielson Framework "to reflect back before and after the lesson helped me to drive instruction." Furthermore, Participant P#6 articulated

the use of the Danielson Framework for lesson reflection stating, "It made me think about what I really want in the lesson: What I was going for, what objectives I was trying to achieve, and what I was really trying to accomplish." Participant P#5 commented, "[The Danielson Framework] has allowed me to focus more on improving what I do, doing better with the delivery, doing better with the student focus, focusing more on all the important information."

Nine the of the participants' responses from the interviews showed that they use building-level data for instructional purposes. During the interviews it was noted that the participants referred to formative and summative data as building data. Participants shared that the data provide valuable information on student performance. The use of the data for instruction was stressed by Participant P#4 who commented:

> It gave me immediate feedback for me to say "whoa, I really need to step back. I thought these guys were getting it, but they really aren't and I need do something different." Before I covered as much content as I could, I just plugged through the material. If students weren't doing well they needed to come see me and I worked with them one on one, but I wouldn't slow the class down. Using student data, I learned that more than just that one or two students were still struggling. I didn't see that prior to using data.

Participant P#6 supported Participant P#4's comments regarding the use of data for instruction and stated, "[Data are] fantastic in terms of pointing out areas on exams, content areas, types of things that my students weren't doing well in… and that's extraordinarily helpful."

Participant P#7 referred to using building-level student achievement data from state assessments as an instructional guide stating:

> The Keystones impacted [instruction] more… I want my students to be successful on the Keystone and I've spent countless hours looking at Keystone standards and anchors and the eligible content making sure that I'm covering all that material so that my students can be successful on that exam for their own good and graduate.

Participant P#10 referenced how data have impacted the instructional focus of the building stating, "There's definitely a building-wide effort to have everyone teach literacy, everyone teach proper writing style and drill down to the common core skills." Similarly, Participant P#13 described how the use of building data has impacted her teaching to students in a non-core subject area and stated:

> It's influenced my delivery of information to students who haven't tested well. So to isolate and know who those students are… it's helpful to know so that I am watching or engaged with that student who needs that extra help in whatever remedial help they need.

Eleven of the participants indicated that they made specific changes to their instructional practices from the feedback they received during the evaluation process. Participants commonly shared that as professionals they are continually looking to grow and improve their classroom instruction. Participant P#7 noted that he is always looking for feedback to improve stating, "I feel like it's just always changing and always evolving. Any feedback that I received in an end of the year evaluation I would take into consideration for the future." The importance of instructional feedback was also stressed by Participant P8. "Even with the experienced teachers, there is no perfect teacher out

there… I just think looking to grow is important and if there's anything that can help me on the evaluation, then I absolutely use it." Participant P#4 described how feedback has changed his instructional practices and commented that he "uses more formative assessments so that I can tweak what I'm doing in the classroom as far as my lesson plans and classroom management as well."

The results of the interviews supported the analysis the Likert-scale survey responses. When analyzing the interview questions, several themes emerged. All the respondents stated that the Danielson Framework had an impact on instructional practices. They noted that the Danielson Framework helped teachers focus their instructional goals and necessitated time for lesson planning and reflection. A majority of teachers generally thought that building-level data were beneficial in providing performance data for instructional purposes. Lastly, the feedback received from the evaluation was valuable and was utilized to improve classroom instruction.

Research Question Two. *What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on teacher collaboration?*

Research Question Two was addressed utilizing data from the survey responses and interview data. The survey included nine Likert-scale questions and one open-ended response. The survey was designed to ascertain the participants' perceptions about the impact of the teacher evaluation process on professional collaboration. Two interview questions asked participants to describe how their collaboration and collegial interactions were altered by the PA teacher evaluation process.

Analysis of survey responses. To determine teachers' perspectives regarding the impact of the teacher evaluation process on collaboration, Likert scale questions 13

through 19 were provided to allow the respondents to indicate the extent to which they agreed or disagreed with each statement. The four-point response scale ranged from strongly agree to strongly disagree.

The most significant agreement regarding collaboration and the teacher evaluation was the perception that teacher learning is supported through a combination of collaborative strategies. The responses indicated that 26% (9 out of 35) of the participants strongly agreed and 66% agreed (23 out of 35) with the statement. Only 6% (2 out of 35) disagreed and 3% (1 out of 35) strongly disagreed. Participants also agreed that teachers in their building meet to discuss ways to improve their teaching and learning, with 57% (20 out of 35) agreeing and 23% (8 out of 35) strongly agreeing. Fourteen percent (5 out of 35) disagreed and 6% (2 out of 35) strongly disagreed with the statement. For the statement concerning teachers collaboratively reviewing student work, the majority of participants 57% (20 out 35) agreed and 17% (6 out of 35) strongly agreed. Only 14% (5 out of 35) disagreed and 11% (4 out of 35) strongly disagreed. In response to the statement that the teachers have professional conversations regarding building level student achievement data, 15% (5 out of 35) of the participants strongly agreed, 48% (16 out of 35) agreed, 25% (9 out of 35) disagreed, 9% (3 of 35) strongly disagreed, and 6% (2 out of 35) did not respond. Participants also agreed that they had more conversations about ways to improve teaching methods with administrators. Fifty-one percent (18 out of 35) agreed and 9% (3 out of 35) strongly agreed with the statement. However, 31% (11 out of 35) disagreed and 9% (3 out of 35) strongly disagreed with the statement. When asked to respond to the statement about collaborating more with other teachers, 29% (10 out of 35) agreed and 9% (3 out of 35)

strongly agreed. However, the majority of participants, 51% (18 out of 35), disagreed and 11% (4 out of 35) strongly disagreed with the statement. Lastly, participants respond to the statement that the teacher evaluation had no impact on collaboration. Thirty-seven percent (13 out of 35) of the participants agreed and 17% (6 out of 35) strongly agreed. Thirty-four percent (12 out of 35) disagreed and 11% (4 out of 35) strongly disagreed with the statement. Table 4.3 provides a summary of the responses regarding their perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on teacher collaboration.

Table 4.3

Teachers' Perceptions on the Impact of the Evaluation Process on Collaboration.

	Participant Response				
Statement	SA	A	D	SD	NR
13. Teachers have conversations regarding student achievement data.	15% (5)	48% (16)	25% (9)	9% (3)	6% (2)
14. Teachers collaborate to review student work and improve instruction.	17% (6)	57% (20)	14% (5)	11% (4)	0% (0)
15. Teachers collaborate more because of the evaluation process.	9% (3)	29% (10)	51% (18)	11% (4)	0% (0)
16. Teachers and administrators have more conversations to improve instruction.	9% (3)	51% (18)	31% (11)	9% (3)	0% (0)
17. Teacher learning is supported through a combination of strategies.	26% (9)	66% (23)	6% (2)	3% (1)	0% (0)
18. Teachers discuss ways to improve their teaching and learning.	23% (8)	57% (20)	14% (5)	6% (2)	0% (0)
19. The evaluation process has had no impact on collaboration.	17% (6)	37% (13)	34% (12)	11% (4)	0% (0)

Note. N=35. SA= Strongly Agree; A=Agree, D=Disagree, SD=Strongly Disagree, NR=No Response

The overall analysis of the questions showed that the majority of participants agreed that teachers support each other through a combination of collaborative strategies (92%), meet and discuss ways to improve the teaching (80%), and collaboratively review and share student work to improve instructional practices (74%). Additionally, the majority of participants (63%) agreed that they have professional conversations regarding building level student data. Furthermore, 60% of the participants agreed that they have had more conversations with their administrators to improve their teaching methods. However, the participants disagreed that they are meeting more with their peers (62%) and they agreed that the teacher evaluation has had no impact on collaboration with their peers (54%).

Analysis of open-ended responses. In addition to the Likert-scale survey items, the researcher gathered additional data through the use of an open-ended question. Open-ended question #27 asked participants to describe how the teacher evaluation process promoted collaboration with their colleagues. A total of 34 participants responded to the question. One participant responded with "not applicable."

Seventeen respondents, or 49%, specified that they thought the PA standards-based teacher evaluation promoted collaboration with their colleagues. Six of the participants indicated that they collaborate specifically about student data and instructional strategies due to the implementation of the teacher evaluation process. One participant stated, "The teacher evaluation process promoted an urge to have more collaboration with my colleagues." Another commented, "I seek out teachers who will help me learn to use data to inform my instruction." Another participant remarked, "Our building collaboration has always been very open and productive; however, this process

has challenged us to be more specific targeting areas of overall need based on student data." Another commented, "We share instructional strategies and collaborate on lesson plan development." Likewise, another participant noted, "We frequently stop in to discuss strategies that we've found effective and even those that failed miserably." Participant P#13 discussed collaborating with specialists in the building for improving instruction and assisting underperforming students. Participant P#13 commented:

> I speak to the reading teachers, the reading specialists for our reading strategies classes because they're there to help the students perform better. So we'll compare notes because I often wonder is it in my class they're not achieving, what am I doing wrong, are they struggling in their other classes? So that data helps me to know just which students I need to be reinforcing.

Seventeen respondents, or 49%, stated that the teacher evaluation was not a factor in promoting collaboration with their colleagues. However, many of the participants responded that they have been regularly collaborating with their colleagues regardless of the teacher evaluation process. One participant commented, "I have always collaborated heavily with colleagues." Similar comments were made by several other participants. For instance, one participant specified, "I have always collaborated with colleagues and will continue to do so." Another participant responded, "Good colleagues collaborate regardless of the evaluation process." Likewise, another participant acknowledged, "I have sought to increase collaboration with my colleagues, but that has nothing to do with the standards-based teacher evaluation system." One participant shared:

> I do not believe it has made a difference for me. I have always worked collaboratively with any team I am on, and my participation in the pilot did not

change that fact. I am currently working with a team of first grade teachers who believe in consistent and meaningful collaboration.

Analysis of interview data. Interview response questions provided further insight into the second research question. Thirteen teachers from two of the four districts participated in individual interviews with the researcher to further explore their perceptions related to the PA standards-based teacher evaluation and their collaboration. The two interview questions asked teachers to (a) reflect on building-level data and their impact on collaboration and (b) describe any differences in collegial interactions because of the standards based evaluation system.

All participants indicated that collaboration was part of their professional practice. However, it was not clear that collaboration was directly impacted based on the PA standards-based teacher evaluation. Seven out of the 13 respondents indicated that discussions concerning building-level data were taking place because of the teacher evaluation process. Regardless of the participants' responses in the interview, 11 of the respondents indicated that they regularly collaborate with their colleagues as part of a committee or with their departments for the purposes of discussing student achievement data, aligning curriculum, and assessment development. Participant P#11 stated, "We have begun to look at the CDT [Classroom Diagnostic Tool]…it's a diagnostic tool that gives us some ideas on the students' progress." Participant P#11 continued about the significance of looking at and discussing data with colleagues stating, "It needs to be discussed. I want to share the good, the bad and the ugly. And so I think we're moving in the right direction." Participant P#10 elaborated on his collaboration within his department stating:

> We've been working together more and more with the skill sets and common assessments… what skills do we need [students] to know… getting away from the content-based instruction and focusing more on the skills and then using the content to teach the skill rather than just trying to stick the skill into the content. That has really been a team effort across the board. Even bringing some of the English people in with the writing skills and what we require and what we should look for.

Participant P#4 expounded on the collegial interactions through referencing his building's committees:

> We've always had the team approach…I am on the Formative Assessment Committee and the Grading Committee, and so those two things, along with the evaluations, have all kind of come together. We have a lot of dialogue in those meetings about how can we use [data]…That happens at our committees.

In addition, Participant P3 commented on how collaboration through committees impacts the classroom by stating, "I'm doing more because I'm in a couple of committees… I'm getting more immediate feedback." Participant P3 noted, "The assessment committee and the grading committee are the ones that are really making an immediate impact in my classroom… and truly make me improve my professionalism." Participant P3 further stated, "If I wait for some other evaluation, it's not as fast and it's not going to be as effective for the kids in the classroom."

Participant P6 described collaboration through conversation within the department stating, "I work in a great department and we're always collegial. You get a good bunch of people who work well together and that kind of collaboration is always going to

happen." Similar comments were made by Participant P#7 who stated, "I really work closely with the other [subject] teachers…we have a common planning time." Participant P#7 described how the time is used for developing content remarking, "We've built Wiki-pages to put content up, we've built pages of podcasts so students can go home to find information and learn from a different sources."

Research Question Three. *What are teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on their professional growth?*

Research Question Three was addressed utilizing data from the survey responses and interview data. The survey included six Likert-scale questions and one open-ended response. The survey was designed to ascertain the participants' perceptions about the impact of the teacher evaluation process on professional growth. Two interview questions asked participants to describe how their professional growth was influenced by the Pennsylvania teacher evaluation process.

Analysis of survey responses. To determine teachers' perspectives regarding the impact of the teacher evaluation process on their professional growth, six Likert-scale questions (#20 through #25) were provided to allow the respondents to indicate the extent to which they agreed or disagreed with each statement. The four-point response scale ranged from strongly agree to strongly disagree. Reflecting on the entire evaluation process, teachers gave their perceptions of how the evaluations system influenced their professional growth. Thirty-five participants responded to the survey. One participant did not respond to statement #20: Professional development that teachers participate in improves instructional strategies in the classroom.

The response data indicated that the participants significantly supported the statement that professional growth work directly impacts student achievement and success. Overwhelmingly, 97# of participantx (34 out of 35) were in agreement, with 71% (25 out of 35) agreeing and 26% (9 out of 35) strongly agreeing, with the statement. No participant disagreed and only 3% (1 out of 35) strongly disagreed with the statement.

Participants responded favorably to the statement that the teacher evaluation usually contain a good balance between encouraging, positive comments and suggestions for improvement and growth. The data specified that 74% of participants agreed (26 out of 35) and 14% (5 out of 35) strongly agreed. Only 9% (3 out of 35) of the participants disagreed and 3% (1 out of 35) strongly disagreed. Among the participants surveyed, a total of 65% agreed (23 out of 35) and 26% (9 out of 35) strongly agreed, that teachers chose professional growth options that directly affected classroom instruction. Only 9% (3 out of 35) of the participants disagreed and no participants strongly disagreed. Similarly, the majority of participants, 68% agreed (23 out of 34) and 12% (4 out of 34), strongly agreed that the professional development improves the instructional strategies used in the classroom. A small minority, 15% (5 out 34) of participants, disagreed and 6% (3 out of 34) strongly disagreed. Participants were slightly more in agreement, 54% agreed overall, that the evaluation has helped teachers establish goals for professional growth. Forty percent (14 out of 35) agreed, 14% (5 out of 35) strongly agreed, while 43% percent (15 out of 35) disagreed and the 3% (1 out of 35) strongly disagreed. Lastly, 54% (19 out of 35) of the participants reported that their evaluation has been used to direct their professional development activities. No participants strongly agreed with the statement. A total of 46% disagreed with the statement, with 37% of the participants

(13out of 35) disagreeing and 9% (3out of 35) strongly disagreeing. Table 4.4 provides a summary of the responses teachers provided regarding their perceptions of the of the Pennsylvania standards-based teacher evaluation process on their professional growth.

Table 4.4

Teachers' Perceptions of the Impact of the Evaluation Process on Professional Growth.

	Participant Response				
Statement	SA	A	D	SD	NR
20. Professional development improves instructional strategies.	12% (4)	65% (23)	15% (5)	6% (2)	3% (1)
21. Professional growth directly impacts student achievement.	26% (9)	71% (25)	0% (0)	3% (1)	0% (0)
22. Evaluation established goals for professional growth.	14% (5)	40% (14)	43% (15)	3% (1)	0% (0)
23. Evaluations direct professional development activities.	0% (0)	54% (19)	37% (13)	9% (3)	0% (0)
24. Professional growth directly affects classroom instruction.	26% (9)	66% (23)	9% (3)	0% (0)	0% (0)
25. Evaluations contain comments and suggestions for improvement and growth.	14% (5)	74% (26)	9% (3)	3% (1)	0% (0)

Note. N=35. SA= Strongly Agree; A=Agree, D=Disagree, SD=Strongly Disagree, NR= No Response

The overall analysis of the questions showed that the majority of participants agreed that the Pennsylvania standards-based teacher evaluation contains comments and suggestions for improvement and growth. Participants also agreed that professional growth impacts student achievement, directly affects classroom instruction, and improves instructional strategies used in the classroom. Lastly, a slight majority of participants agreed that the evaluations have directed their professional growth goals and development.

Analysis of open-ended response. In addition to the Likert scale survey items, the researcher gathered additional data through the use of an open-ended question. The

open-ended question asked participants to describe how the teacher evaluation process has affected their professional development. A total of 34 participants responded to the question. One participant responded with "not applicable."

Seventeen respondents, or 49%, specified that they thought the PA standards-based teacher evaluation had affected their professional development. After analyzing the participants' statements, a theme emerged around enhancing classroom learning. Participants identified that their teacher evaluation focused their professional development on topics such as formative assessments, teaching strategies, and self-reflection. One respondent commented, "I am looking into taking more classes each year to enhance my professional development and, in turn, enhance my classroom teaching skills." Another participant specified, "I am focusing more on the skill sets of the students rather than content." Another respondent stated, "My view on professional development has changed. I am trying to apply what we are learning on formative assessment and data analysis." One respondent remarked:

> I find that I am working harder than I ever have in my 23 years of teaching. I thought I was a good teacher, but this allows me to improve my teaching strategies, making sure that I am targeting all learners. I am constantly re-evaluating my lessons and making note of what I need to change for next time.

Seventeen respondents, or 49%, specified that they thought the PA standards-based teacher evaluation had not affected their professional development. A similarity emerged when comparing the statements from respondents who believed that the evaluation had an effect on professional development and statements from respondents who believed the evaluation had no effect. The common attributes included themes of

enhancing professional knowledge, self-reflection, and professional improvement. One participant whose response suggested that the teacher evaluation had no effect on professional growth wrote, "I'm always trying to improve my skills and enhance my teaching regardless of the evaluation system." Another participant concluded, "I do nothing differently. I prefer to attend college courses or weeklong workshops in my field [of mathematics] to improve my knowledge base and teaching background." A different person commented, "I am always learning, self-evaluating and retooling." One more participant specified, "I am not really doing anything differently. My choices of professional development programs have always been geared towards becoming a more effective teacher."

Analysis of interview data. In addition to the survey responses and open-ended items, interviews provided further insight into the research question. Thirteen teachers from two of the four districts participated in individual interviews with the researcher to further explore their perceptions related to the PA standards-based teacher evaluation and their professional growth. The two interview questions asked teachers to reflect on (a) building-level data and their impact on professional growth and (b) the influence of the evaluation on professional goals.

Participants' interview responses were similar to those identified in the survey and open-ended questions. Seven of the participants referenced using building data to develop their professional goals. Through their participation in building committees such as the formative assessment and grading committees, participants discussed how they use data to develop their professional goals. Participant P#3 stated, "I'm getting more immediate feedback in what I'm doing in committees. Committees are truly making the

impact for me to improve my professionalism the following year." Participant P#3 continued to describe the committees' goals stating, "We're looking at how we're grading and why we grade the way we do. It's going to be quite the paradigm shift from where we're going to and away from." Participant P#4 is part of a formative assessment committee, and joined the committee as a professional goal. Working as part of a committee has changed Participant P#4's classroom instruction affirming, "I used to grade everything. Now I pretest the units. That's a big change for me."

Nine of the participants stated that the evaluation process has an impact on developing professional goals. The participants recognized the development of professional goals as an opportunity to enhance their effectiveness in the classroom. Participant P#5 declared, "I personally consider myself a highly effective teacher…there's always room for personal growth and personal development." Participant P#1 similarly stated, "Everybody knows where they could be a little bit stronger."

Participant P#6 commented about the influence of the evaluation on professional goals explaining, "I think the reflection is really important. You could teach for 25 years or you could teach the same year for 25 years." Participant P#6 identified that this year's goal will be to "write a literacy design unit", and further specified, "It [the evaluation] made me think about what really went into the lesson and trying to come up with better ways to do things, new content, and new approaches."

Participant P#2 commented on developing self-directed goals for the first time since being an Instructional I teacher. While a goal had not yet been developed, Participant P#2 mentioned that the plan would be, "to go back to the evaluation…and see

what my supervisor identified." Similarly, when developing professional goals based on the evaluation, Participant P#8 explained:

> I'll look at the criticism or comments. Even with the experienced teachers…I think they'll be the first to tell you they're not perfect. I just think looking to grow is important and if there's anything that can help me on the evaluation, then I absolutely use it.

Participant P#10 believed the reflective nature of the evaluation influenced this year's creation of professional goals. This year Participant P#10 planned to "focus on the content using formative assessments." Participant P#10 commented that the goal will impact the classroom by "gauging the students' understanding." Professionally speaking, Participant P#10 stated that the focus on formative assessment will "build the skills in what I am doing to make sure the kids are getting those skills rather than just at the moment within the classroom setting."

Participant P#9 discussed developing professional goals for the department using what was learned during the pilot study. Participant P#9 stated:

> I would like to have teachers use Danielson's Framework and the new evaluation system… The Framework was created to help teachers become better by using it as a tool to communicate and improve our instruction, not just as an evaluation tool. I would like to have teachers observing teachers using the evaluation system so that we could just focus on what we could do to improve our teaching. I just want to have teachers observe me and me observe them and talk about what we're doing right and what we're doing wrong and how we can fix it so that when we

do get observed, it won't be an issue. We'll be comfortable with it and know what to expect.

Summary

This chapter analyzed the results from survey and interview responses to identify teachers' perceptions of the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model. Research Question One focused on the impact of the Pennsylvania standards-based teacher evaluation process on teachers' instructional practice. Research Question Two examined teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on teacher collaboration. Research Question Three surveyed teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on their professional growth.

Based on the data collected, the majority of participants indicated that the teacher evaluation model had an impact on instruction. Participants agreed in both the survey and interview statements that the teacher evaluation was able to determine the quality of teaching and provided meaningful feedback. The majority of respondents also commented that they were using student data to drive their instruction and were more likely to vary their instructional techniques due to the teacher evaluation. Respondents were also in agreement that the evaluation rubrics were a beneficial guide for developing lessons. However, there was mixed reporting on two points: whether the teacher evaluation directly changed teachers' instructional strategies and if the teacher evaluation required teachers to increase their time to plan lessons and prepare for classes. The Likert survey and open-ended results were evenly divided on these topics. Nevertheless, the interviews revealed that the Danielson Framework did influence teacher instruction.

Teachers reported that the Framework made them more reflective and focused on their instruction. Overall, the majority of respondents agreed that the teacher evaluation process impacted their instructional practices.

Response data also identified teachers' perceptions of the impact of the Pennsylvania standards-based teacher evaluation process on teacher collaboration. The majority of participants agreed that teachers meet and discuss ways to improve their teaching, have professional conversations regarding building level data, and collaboratively review and share student work to improve instructional practices. Respondents commented that they collaborate specifically about student data and instructional strategies due to the implementation of the teacher evaluation process. However, the data collected also suggest that respondents collaborate with their colleagues regardless of the teacher evaluation process. Furthermore, the participants disagreed that they meet more with their peers due to the teacher evaluation process. Based on the results collected, the majority of participants believe that the teacher evaluation has had no impact on collaboration with their peers.

Teacher response data from both the survey and interviews maintained that the Pennsylvania standards-based teacher evaluation had an impact on professional growth. The majority of participants were in agreement that the teacher evaluation contained comments and suggestions to establish goals for professional growth. Respondents identified that their evaluation has been used to direct their professional development activities. Additionally, participants commented that they focused their professional growth on student achievement that directly affected classroom instruction to enhance

classroom learning. A summary of the study and its results, as well as its relationship to other research, is discussed in Chapter Five.

Chapter Five – Discussion

Summary of the Study

This qualitative study explored teacher perceptions of the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model. The study focused specifically on the Pennsylvania Teacher Effectiveness evaluation model's influence on teacher instruction, collaboration, and professional growth.

The researcher examined the beliefs and perceptions of 35 teachers from four school districts of different demographics. Two of the school districts selected were suburban districts in southeastern Pennsylvania. One school district was a rural district located in Eastern-Central Pennsylvania. Another was an Intermediate Unit in Pennsylvania. Data were collected using an online Surveymonkey.com survey instrument, which included 25 Likert-type items and three open-ended response items. Additionally, 13 of the survey respondents from two of the four districts participated in interviews with the researcher to expand the survey responses.

The data collected from the survey were evaluated in relation to the research questions of the study, placed into tables, and reported in summary form. Participants' interview responses were also analyzed to identify trends, attitudes, and practices of teachers in relation to the impact of the teacher evaluation on instruction, collaboration, and professional growth. The survey and open-ended responses were categorized and analyzed to determine emerging themes. The qualitative data provided insight and depth into understanding the teachers' perceptions.

Summary of the Results

This study was designed to describe and analyze the data collected from Likert-scale survey responses, open-ended survey questions, and individual interview responses

from teachers. The responses were categorized and analyzed for key themes and trends related to the three research questions.

Research Question One examined the impact of the Pennsylvania Teacher Effectiveness evaluation process on teacher instruction. Analysis of the data related to Research Question One concluded that Pennsylvania's implementation of the Danielson Framework as part of the Teacher Effectiveness program has influenced the way teachers instruct. Furthermore, teacher responses, statements, and comments from the survey and interviews revealed consistent findings related to teaching practices, skills, and knowledge that positively impact student learning.

The survey responses identified that the evaluation process was an effective indicator for determining the quality of teaching and provided information for feedback and planning instruction. Teachers surveyed overwhelmingly agreed (92%) that the teacher evaluation encouraged teachers to use a variety of instructional techniques. Overall, 68% of the participants stated that they were more likely to use student data to plan their instruction and that the teacher evaluation process has impacted their instructional practices. A slight majority (54%) of teachers indicated that the teacher evaluation provided an accurate picture of their teaching ability. The results were evenly divided with 49% agreeing and 51% disagreeing that the teacher evaluation changed teachers' instructional strategies and required teachers to increase their time to plan lessons and prepare for classes.

The findings confirm a majority of teachers perceived that the evaluation process supported the utilization of differentiated instructional strategies. This marked a change in instructional practice for approximately half of the sample, who indicated that the

process had caused them to alter their teaching methods. For the other half of the sample, the new system did not necessitate a change in teaching practice: it appears that they already employ a range of instructional strategies. The findings indicate that the Teacher Effectiveness model supports teachers to select appropriate instructional strategies based on student data, lesson reflection, and feedback.

Open-ended responses and interview information from teachers supported the findings of the survey. Teachers identified that the Danielson Framework embedded in the teacher evaluation influenced their teaching and instruction. All 13 participants interviewed indicated that the Danielson Framework focused their instruction by providing guidelines for lesson planning and reflection. The majority (70%) of the 13 teachers interviewed stated they use building data for instructional purposes. Teachers specified that teacher evaluation encouraged them to use student data from formative and summative assessments for instructional purposes. Teacher responses from the open-ended questions and interviews indicated that they changed their instruction because of feedback they received during the evaluation process.

Consistent throughout the survey and interviews, teachers believed that the adoption of the Danielson Framework made a positive impact on their instruction. Teachers shared that the new evaluation has enhanced their skills and made them more effective instructionally. Teachers used data to plan and adjust instruction to meet the needs of the students. Linking the results of the data shows that the Pennsylvania Teacher Effectiveness model has helped teachers to become more reflective and more effective classroom practitioners.

Research Question Two sought to identify the impact of the teacher evaluation process on collaboration. The Pennsylvania Teacher Effectiveness evaluation process focuses on quality collaboration, reflection, and discussion. Teacher responses to the survey, open-ended responses, and interview questions revealed that the majority of teachers meet and discuss ways to improve instruction through professional conversations and those conversations focus on student achievement data and classroom instruction.

Teacher responses from the data clearly indicated that collaboration is a regular means of professional practice. The analysis of the survey data showed that the vast majority (92%) of the 35 teachers agreed that teacher learning is supported through collaborative practices such as professional learning communities and building-level committees. Additionally, a large majority (80%) of the teachers agreed that they meet regularly to discuss methods to improve their instruction. Furthermore, 74% of teachers agreed that they collaboratively review and share student work for the purpose of improving classroom instruction. A majority of participants (60%) also agreed that they have had more conversations with their administrators to improve their teaching methods. However, a majority (62%) of participants disagreed that they are meeting more with their peers due to the teacher evaluation. A slight majority (54%) of teachers stated that the teacher evaluation has had no impact on collaboration with their peers.

The Teacher Effectiveness model may not directly impact teacher collaboration. However, there is an indirect relationship. The collected data indicated the presence of a culture of collaboration. The analysis of the teachers' responses reflects that professional development and collaboration are ongoing and linked. Study findings suggest that

teachers are regularly engaged in the examination of teaching practices and student learning.

Correspondingly, open-ended responses mirrored the survey results. Half of the open-ended responses specified that teachers thought the Pennsylvania standards-based teacher evaluation promoted collaboration with their colleagues. Those participants indicated that they collaborate specifically about student data and instructional strategies due to the implementation of the teacher evaluation process. The other half of the responses stated that the teacher evaluation process was not a factor in promoting collaboration with their colleagues. Teachers clarified that they regularly had collaborated with their colleagues prior to the implementation of the teacher evaluation.

Findings from the interviews were consistent when compared to the survey and open-ended responses. Unanimously, teachers indicated that collaboration was part of their regular professional practice. Teachers are committed to professional learning through collaborative inquiry. A slight majority (54%) of the responses indicated that discussions concerning building-level data were taking place because of the teacher evaluation process. Furthermore, teachers described their collaboration as taking place with the purposes of discussing student achievement data, aligning curriculum, and assessment development.

The Pennsylvania Teacher Effectiveness model was designed to encourage collaborative inquiry. However, establishing whether teacher collaboration was directly impacted based on the PA standards-based teacher evaluation could not be clearly determined. As stated earlier, a slight majority of teachers surveyed believed that the evaluation had no impact (54%) and disagreed (62%) that they were collaborating more

because of the evaluation process. The open-ended statements and interviews confirm these positions. Teachers did not perceive that the evaluation assisted or encouraged collaborative inquiry. The teachers' responses indicated that collaboration was already an established practice in the school districts of this study. The response to the second research question supported collaboration as an on-going professional practice. However, the study results from both the open-ended responses and the interviews were evenly divided as to whether or not there was an impact on collaboration as part of the evaluation process. Regardless of the impact of the Pennsylvania Teacher Effectiveness evaluation process, the results of this study showed that teachers are engaging in collaborative inquiry to analyze and improve their classroom instruction.

Research Question Three addressed teachers' perceptions of the impact of the Pennsylvania teacher evaluation process on professional growth. Response data from the survey, open-ended responses, and interviews supported that the Pennsylvania standards-based teacher evaluation had an impact on teachers' professional growth. The majority of participants were in agreement that the teacher evaluation contained comments and suggestions to guide goals for professional growth. Teachers identified that the evaluation's feedback directed them to activities that focused on student achievement and classroom instruction.

The majority of participants in the survey agreed that the Pennsylvania standards-based teacher evaluation promotes comments and suggestions for improvement and growth. Participants also agreed that professional growth impacts student achievement, directly affects classroom instruction, and improves instructional strategies used in the

classroom. Lastly, the majority of participants agreed that the evaluations had directed their professional development.

In analyzing the participants' open-ended statements, it was clear that the evaluation focused professional development towards enhancing classroom learning. Teachers agreed (92%) that they chose professional growth options that directly affect classroom instruction in large part due to the evaluation process. Additionally, teachers specified that the evaluation helped guide and focus their professional development on topics such as formative assessments, teaching strategies, and reflection.

Interview responses were comparable to those identified in the Likert-scale and open-ended questions. Participants remarked that the teacher evaluation contained comments and suggestions to establish professional growth goals that are linked with classroom instruction. Teachers specified that the building data provide additional information to develop their professional goals. During the interview, teachers commented that the information gathered supports their participating in building committees, such as the formative assessment and grading committees, to enhance their professional growth.

The results of this study confirm that teachers recognized that the Pennsylvania standards-based teacher evaluation process had an impact on classroom instruction and teacher professional growth. While the results of the study were divided as to the impact on teacher collaboration, teachers in the study were regularly utilizing collaborative inquiry as professional growth. Teachers' responses clearly indicated that the Danielson Framework, as part of the evaluation process, had provided teachers with information for lesson reflection and feedback. Furthermore, teacher statements revealed that the

evaluation process has encouraged teachers to employ student data for instructional planning, which has resulted in teachers' varying their instruction. The study also concluded that the comments and feedback from the teacher evaluation have directed teachers' professional development towards specific goals. Teacher statements illustrated that they are choosing professional growth options to address areas specified in the evaluation. Clearly, the evaluation process provided a shared understanding of effective teaching practices and promoted professional conversations that impact student learning. The results of the study found that the Pennsylvania standards-based evaluation model's design has created the conditions for teachers to reflect, collaborate, and grow to improve classroom instruction and student achievement.

Limitations Found in the Study

While some limitations to the study were identified in Chapter One, additional limitations emerged during the study. Subjects in this study had previously participated in the pilot phase of Pennsylvania's implementation of the Teacher Effectiveness model. A total of 293 school districts in Pennsylvania participated in the pilot phase. Each district identified a pilot group of volunteer teachers from their staff to participate in the Pennsylvania Teacher Effectiveness pilot. The number of available participants varied based on the district's pilot group. The reasoning for a teacher to participate varied as well. Teachers who were involved in the pilot were volunteers and, therefore, made the choice to be evaluated in this fashion. Consideration must account for the possibility that teachers who participated had a particular interest or bias toward the topic of the study. The potential for individual bias may have skewed the overall results of the study.

Another limitation was the lack of sample size. Four school districts participated in the study; 117 potential teachers were eligible to participate in the study. Only 35 participated in the survey and 13 participated in the interview process. In addition, no teachers from School Districts B and D volunteered to participate in the interview process. The limited number of participants and the lack of interviewees from School Districts B and D impact the generalizability of the study.

Finally, the 2013-14 school year was the first full implementation of the Teacher Effectiveness Evaluation. Teachers have had a limited exposure to the process. Therefore, their perceptions and insights can be considered initial and novice.

Relationship to Other Research

Researchers have agreed that in order to improve teacher instruction and increase student achievement, the current teacher evaluation system needed to be reformed (Danielson, 2011a; Kimball et al., 2004; Marshall, 2012). It is believed that teacher evaluations should be designed to determine the overall quality of a teacher and identify areas for specific professional growth (Danielson & McGreal, 2000; MET, 2013). Darling-Hammond et al. (2012) suggested a need for a more comprehensive teacher evaluation system that must provide teachers with meaningful feedback, use multiple measures of data, offer professional growth, and integrate with a system-wide emphasis on quality teaching. The results of this research add to the existing literature regarding the teacher evaluation process and its impact on instruction, collaboration, and professional growth. The findings of the study were consistent with the literature reviewing the impact of a teacher evaluation system as presented in Chapter Two.

Danielson (2007) remarked that for educators to be highly effective there is a need to create feedback systems that will give educators specific information on their teaching and how this affects student performance. According to Danielson (2011), the Framework's value is its design to provide a foundation for professional conversation, feedback, and reflection. Linking those purposes together, the Danielson Framework assists teachers in becoming more thoughtful practitioners (Danielson, 2011). The results of this study support Danielson's remarks. All 13 teachers interviewed stated that the Danielson Framework provided meaningful feedback that was used for reflection and the improvement of classroom instruction. A majority of the teachers (77%) surveyed agreed that the evaluation process provided useful feedback for planning and instruction. Additionally, results from the survey and interviews specified that teachers are having professional conversations regarding building level student data and collaboratively review and share student work to improve instructional practices.

Borman and Kimball's (2004) study of the Washoe County School District and Milanowski's (2004) study of the Cincinnati Public Schools system collectively analyzed Danielson's Framework of Teaching. Both the Cincinnati Public Schools' and the Washoe County School District's studies indicated that standards-based evaluation systems can be designed to measure teacher quality. Additionally, both studies concluded that standards-based evaluation systems can be designed to measure teacher quality through domains, rubrics, and sources of evidence such as student achievement data, as an overall measure of teaching quality.

The results of this study validate both Borman and Kimball's (2004) and Milanowski's (2004) studies. As evidenced by the results of this study, teachers

recognized the Danielson Framework as a valuable instrument for feedback and reflection. The results also confirmed that the evaluation provided a more accurate picture of teacher performance. Teachers surveyed in this study (68%) utilized student achievement data for planning purposes because of the teacher evaluation process. In the interviews completed in this study, teachers acknowledged using feedback from their evaluations to enhance their lessons and vary their instructional techniques. The results of this study validate the research as specified by the teachers' perceptions of the effectiveness of the Danielson Framework and its on impact instruction.

The perceptions of teachers in this study support the research that a standards-based evaluation design, with feedback and reflection on multiple measures, can impact teacher instruction (MET, 2013). The MET project study focused on 3,000 teachers and over 23,000 hours of lesson videos to determine how evaluation methods could best be utilized to help teachers be most effective. The study identified that multiple measures such as observations, observation rubrics, and student performance data can provide an accurate picture of teaching effectiveness. Results obtained in this study support the MET project's findings. Teachers perceived the teacher evaluation as an effective model for determining the quality of teaching, providing feedback and reflection, and using student data for planning purposes. Additionally, 92% of the participating teachers acknowledged that the evaluation required them to utilize a variety of instructional strategies.

Effective instructional collaborative inquiry focuses on improving classroom practices in order to improve student achievement (Nelson et al., 2010). Various structures have been identified to support teachers' professional growth in collaborative

settings, including professional learning communities, lesson study, communities of practice, and peer observation (Nelson & Slavit, 2008). Similarly, the results of this study support the statements that collaborative teams, structured through professional development opportunities, have helped teachers expand their pedagogy and utilize differentiated learning strategies that positively impact student performance (DuFour et al., 2004).

The results of this study found that teachers are teaming, working in collaborative groups, and having collegial conversations regarding student achievement and instruction. Further, the results from the study confirm other research, in that collaborative inquiry supports individual professional development (Nelson et al., 2010). The study results compare to the results of Robert's (2010) research regarding creating a culture of collaboration. In both studies, teachers reported that they collaborate specifically about student data and instructional strategies. Key themes from this study showed that many teachers participate in a culture of collaboration. The vast majority of teachers (92%) indicated that teacher learning is supported through collaboratively practices such as professional learning communities. Teachers frequently referenced their participation in committee work and department level meetings as opportunities to work collaborative with their peers. A majority of the teachers surveyed (74%) reported that they collaborate and share student data that directly impacted classroom instruction and student achievement. While the results are not conclusive as to whether the teacher evaluation has had a direct influence on teacher collaboration, it is appropriate to state that teachers are practicing collaborative inquiry. Unanimously, all teachers indicated that collaboration was part of their regular practice.

Congruent with the research, the results of this study found that teacher collaboration worked in conjunction with professional growth to impact classroom instruction. Focused work in collaborative teams, structured through professional development opportunities, can help teachers expand their pedagogy, differentiate learning strategies, and positively impact the classroom instruction (DuFour et al., 2004). The themes of this research study's findings support the formation of professional learning communities among group members. These communities provide flexibility in response to the individual teacher's needs, and offer opportunities for professional development based on the outcomes of the teacher evaluation. Survey and interview responses indicated that the teacher evaluation has directed teachers towards participation in professional growth that focuses on student achievement. Teacher evaluations provided valued feedback to assist teachers in developing quality professional growth options (Danielson, 2010). This study concurs with other research as teachers postulated that the evaluation included suggestions to establish goals for professional growth, and theyperceived this feedback as an enhancement to professional learning opportunities.

Finally, research has suggested that a rigorous evaluation model, based on multiple measurements, provides quality targeted feedback (Darling-Hammond et al., 2012; Nelson et al., 2008). It is important to improve student learning, but it must simultaneously promote teacher development (Papay, 2012). The results of this study support that teacher evaluation has impacted teacher instruction and professional growth. Teachers are using the evaluation process to adjust their instructional practice in an effort to improve student achievement. Additionally, teachers are looking to collaborative inquiry as a means to an end to improve the quality of their own instructional practices.

Professional growth is perceived to occur through peer conversations which include a review of building-level data and student performance. The result of this study indicate that professional growth efforts directly impact student achievement and success. Grading and student assessment committees, lesson studies, and peer observations were forms of collaboration described by subjects in this study which have been viewed positively in the literature (DuFour et al., 2004; Glaser, 2005; Marshall, 2009).

Recommendation for Further Research

This study focused on teacher perceptions of the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model on instruction, collaboration, and professional growth. During the time of the research study, the participants had one year of experience with the Pennsylvania's Teacher Effectiveness evaluation model. In the 2013-14 year, every district in Pennsylvania was required to use the Teacher Effectiveness model. This study has prompted recommendations for further research. First, it would be beneficial to conduct a longitudinal study to gain further information about the effectiveness of the evaluation model by continuing the study after teachers had several years of experience. A longitudinal study would gauge the evolution and long-term impact of the teacher effectiveness model.

Second, a study could be developed to examine administrators' perspectives of the impact of the teacher effectiveness model on teachers' instruction, collaboration, and professional growth. An administrative comparison of the study would provide a broader picture of the impact of the teacher effectiveness model.

Third, the use of student achievement data to drive classroom instruction was identified as a theme of influence from the teacher evaluation model. A focused study on

a teacher's impact on the academic progress rates of students could be examined by utilizing data from the Pennsylvania Value Added Assessment System. The study could analyze teachers' use of PVAAS data for instructional purposes.

Conclusion

The purpose of this qualitative study was to explore teacher perceptions of the impact of Pennsylvania's standards-based, Teacher Effectiveness evaluation model. The study explored teachers' perceptions of how the evaluation process influenced teacher instruction, collaboration, and professional growth. The schools selected for the study were from three school districts and an Intermediate Unit that had participated in the pilot phase of the Pennsylvania's Standards-Based Teacher Effectiveness program. Conclusions were made based on the data collected through the use of an online Surveymonkey.com survey with Likert-scale and open-ended questions and face to face interviews.

The results of the study revealed that teachers believed that the teacher evaluation process has impacted their instructional practices. A common theme emerged that teachers valued the information provided by the Danielson Framework. Throughout the survey and interviews, teachers described using the feedback gained from the evaluation as a means for lesson reflection. Teachers identified that the evaluation process provided an accurate picture of the quality of teaching and instructional practices. Moreover, the research findings concluded that the teacher evaluation encouraged teachers to use student data for developing lessons and varying their classroom instructional techniques.

Academic growth is one of the primary goals of education. In order for students to achieve in the classroom, teachers must provide quality instruction and educational

practices. Routine evaluation of teachers has offered crucial feedback to teachers as they continue to develop as educators. Teachers have reported that the evaluation process provides an accurate picture of instruction and fosters collaboration among staff. Use of student test results has been added to Pennsylvania testing. The teacher evaluation model encourages use of student data to develop lessons and vary classroom instructional techniques.

Furthermore, the research findings concluded that teachers are having professional conversations to discuss ways to improve their teaching. Collaborative inquiry conversations focus on student work and building level student data to improve instructional practices. However, it was not clear that collaboration was directly impacted based on the PA standards-based teacher evaluation. Teacher responses were evenly divided as to whether collaboration was directly impacted by the evaluation process. While the results cannot conclude that the teacher evaluation process directly influenced teacher collaboration, the survey and interview data indicated that collaboration was a consistent professional practice.

The research data from both the survey and interviews concluded that the teacher evaluation had an impact on teachers' professional growth. The findings suggest that teacher evaluations contain comments and suggestions to establish goals and direct their professional development activities. However, a key theme emerged tying professional growth and collaborative inquiry. The research data suggest that teachers are regularly collaborating with their colleagues as part of their professional growth. Teachers discussed the work completed in committees such as student achievement data, aligning curriculum, and assessment development. Teachers commonly referred to work being

completed as a committee member with dual purposes of collaborative inquiry and professional growth. The findings support that teachers choose professional growth options that directly impact student achievement. Additionally teachers commented that they focused their professional growth on student achievement and instruction to enhance classroom learning.

Teacher quality is one of the most important factors in determining student success. The findings described in this study serve to support an evaluation model that engages teachers in productive conversations about instructional practice. In considering the various ways of measuring teacher effectiveness, a teacher evaluation system should be engaging, concentrating on the importance of student learning, lesson reflection, and instruction. Additionally, educators should consider an evaluation system that utilizes common criteria so evaluators and teachers can better understand how to develop instruction in their schools. Evaluating teacher effectiveness should ultimately lead to improved instruction. Schools will benefit when educational leaders align professional learning opportunities that support the improvement of teachers. These opportunities need to link formal professional development with individual and building goals that allow for shared expertise and collaboration. This study has connected the benefits of a teacher evaluation process that encourages reflection, teacher collaboration, and professional growth to student achievement.

The results of the study support the research that a comprehensive multiple measures observation and evaluation model portrays a better, more accurate picture of teachers' performance than the traditional, classroom observation and evaluation format (DuFour et al., 2004; MET, 2013; Marshall, 2009). Furthermore, the standards-based

evaluation design improves teacher quality through clear standards and measures of accountability (Donaldson, 2009). The findings of the study conclude that the Pennsylvania standards-based evaluation model creates a comprehensive evaluation process that cultivates the teachers' ability to reflect, collaborate, and refine instructional skills. Development of these skills among teachers is crucial, as educators strive to provide the best educational experience to the students in our schools.

References

Amrein-Beardsley, A. (2008, March). Methodological concerns about the education value-added assessment system. *Educational Researcher*, *37*(2), 65-75.

Anderman, E. M., Anderman, L. H., Yough, M. S., & Gimbert, B. G. (2010). Value-added models of assessment: Implications for motivation and accountability. *Educational Psychologist*, *45*(2), 123-137.

Aument, R. (2011). A new teacher evaluation system for Pennsylvania. *Commonwealth Foundation*. Retrieved from www.commonwealthfoundation.org

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., & Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.

Baratz-Snowden, J. (2009). *Fixing tenure: A proposal for assuring teacher effectiveness and due process*. Washington, DC: Center for American Progress.

Borman, G. D., & Kimball, S. M. (2004). *Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps?* (TC-04-03). Retrieved from Consortium for Policy Research in Education: http://cpre.wceruw.org/papers/teacher_equity_aera04.pdf

Center for Education Policy and Practice. (2010). *Reinventing educator evaluation: Connecting professional practice with student learning*. Retrieved from http://www.massteacher.org

Chait, R. (2009). *Ensuring effective teachers for all students*. Washington, DC: Center for American Progress.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long term impacts of teachers: Teacher value-added and student outcomes in adulthood*. Cambridge, MA: National Bureau of Economic Research.

Colorado Department of Education. (2013). *State model evaluation system for teachers*. Retrieved from http://www.cde.state.co.us/EducatorEffectiveness/SMES-Teacher.asp

Creswell, J. W. (2009, April). Mapping the field of mixed methods research. *Journal of Mixed Methods Research, 3*, 95-108. http://dx.doi.org/10.1177/155868908330883

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

Danielson, C. (2010, December). Evaluations that help teachers learn. *Educational Leadership*, *68*(4), 35-39.

Danielson, C. (2011a, October). It's your evaluation: Collaborating to improve teacher practice. *NJEA Review*, *85*(2), 24-27.

Danielson, C. (2011b). *The Framework for Teaching evaluation instrument*. Retrieved from http://www.danielsongroup.org

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

Darling-Hammond, L., Cook, C., Jaquith, A., & Hamilton, M. (2012). *Creating a comprehensive system for evaluating and support effective teaching*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

David, J. L. (2009). What research says about collaborative inquiry. *Educational Leadership*, *66*(4), 87-88.

Davis, M. L., Williams, S. C., & Griffin, H. (2003, March 20). Teacher education reform: A search for common ground. *Current Issues in Education, 6*(6). Retrieved from http://cie.ed.asu.edu/volume6/number6/

Donaldson, M. L. (2009). *So long Lake Wobegon? Using teacher evaluation to raise teacher quality*. Washington, DC: Center for American Progress.

Donaldson, M. L., & Peske, H. G. (2010). *Supporting effective teaching through teacher evaluation: A study of teacher evaluation in five charter schools*. Retrieved from Center for American Progress: http://www.americanprocess.org

DuFour, R. (2011, February). Work together but only if you want to. *Phi Delta Kappan, 95*(5), 57-61.

DuFour, R., DuFour, R., Eaker, R., & Karhanek, G. (2004). *Whatever it takes: How professional learning communities respond when kids don't learn*. Bloomington, IN: National Educational Service.

DuFour, R., & Marzano, R. J. (2009, February). High leverage strategies for principal leadership. *Educational Leadership, 66*(5), 62-68.

Educator Effectiveness Project Overview. (2013). Retrieved from http://www.portal.state.pa.us/portal/server.pt/community/EEP/20903/page/1171617

Fullan, M. (2006, November). Leading professional learning. *The School Administrator, 63*(10), 10-14.

Fullan, M. (2007, Summer). Change the term for teacher learning. *National Staff Development Council, 28*(3), 35-36.

Furtwengler, C. B. (1995, February 15). State actions for personnel evaluation: Analysis of reform policies, 1983-1992. *Education Policy Analysis Archives, 3*(4), 1-27.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Cranberry, NJ: Merrill.

Given, H., Kuh, L., LeeKeenan, D., Mardell, B., Redditt, S., & Twombly, S. (2010). Changing school culture: Using documentation to support collaborative inquiry. *Theory into Practice, 49*, 36-46. http://dx.doi.org/10.1080/00405840903435733

Glaser, J. (2005). *Leading through collaboration: Guiding groups to productive solutions*. Thousand Oaks, CA: Corwin Press.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

Green, P. C., Baker, B. D., & Oluwole, J. (2012). The legal and policy implications of value-added teacher assessment policies. *B.Y.U Education and Law Journal, 57*, 1-29.

Hanna, J. (2005, June 1). The Elementary and Secondary Education Act: 40 years later. *Ed Magazine*. Retrieved from http://www.gse.harvard.edu/news/2005/0819_esea.html

Hanover Research. (2012). *Best practices for including multiple measures in teacher evaluations*. Retrieved from http://www.hanoverresearch.com/wp-content/uploads/2012/05/Best-Practices-for-Including-Multiple-Measures-in-Teacher-Evaluations-Membership.pdf

Hazi, H. M., & Rucinski, D. A. (2009, Summer). Teacher evaluation as policy target: Viable reform venue or just another tap dance? *Educational Research Service, 27*(3), 31-40. Retrieved from http://depts.hre.wvu.edu/hhazi/spec_su_09_teacheval.pdf

Hightower, A. M., Lloyd, R. C., Wittenstein, R., Sellers, K., & Swanson, C. B. (2011). *Improving student learning by supporting quality teaching: Key issues, effective strategies*. Retrieved from http://www.edweek.org/media/eperc_qualityteaching_12.11.pdf

Johnson, S. M. (2011). *Why teachers must have an effective evaluation system*. Washington, DC: American Federation of Teachers.

Kane, T. J., & Cantrell, S. (2012). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Seattle, WA: Bill and Melinda Gates Foundation.

Kimball, S. M. (2002). *A case study of the Anoka-Hennepin teacher performance appraisal system*. Madison, WI: Consortium for Policy Research in Education.

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

Kohut, G. F., Burnap, C., & Yon, M. G. (2007). Peer observation of teaching: Perceptions of the observer and the observed. *College Teaching, 55*(1), 19-25.

Koops, J. B., & Winsor, K. A. (2005, Fall). Creating a professional learning culture through faculty evaluation. *The Journal of Education, 186*(3), 61-70.

Kress, S., Zechmann, S., & Schmitten, J. M. (2011). When performance matters: The past, present, and future of consequential accountability in public education. *Harvard Journal on Legislation*, *48*(1), 185-234.

Kupermintz, H. (2003). Teacher effects and effectiveness: A validity investigation of the Tennessee Value Added Assessment system. *Educational Evaluation and Policy Analysis*, *25*(5), 287-298.

Lane, S., & Horner, C. (2011). *Pennsylvania teacher and principal evaluation pilot final report*. Retrieved from http://teampa.com/wp-content/uploads/2012/04/SuzanneLaneFinalReport.pdf

Leana, C. R. (2011, Fall). The missing link in school reform. *Stanford Social Innovation Review*, 30-35.

Learning Point Associates. (2007). *Implementing the No Child Left Behind Act: Teacher quality improves student achievement* (ED-01-CO-0011). Washington, DC: Government Printing Office.

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

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage.

Marshall, K. (2005, June). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, *86*(10), 727-735.

Marshall, K. (2006). Teacher evaluation rubrics: The why and the how. *EDge Magazine*. Retrieved from ecologyofeducation.net/wsite/wp.../edge-rubrics-jan-1-25-06.doc

Marshall, K. (2009). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap*. San Francisco, CA: Jossey-Bass.

Marshall, K. (2012). Fine tuning teacher evaluation. *Educational Leadership*, *70*(3), 50-53.

Marzano, R. J. (2012, June). The two purposes of teacher evaluation. *Educational Leadership*, *70*(3), 14-19.

Marzano, R., Toth, M., & Schooling, P. (2012). *Examining the role of teacher evaluation in student achievement*. Retrieved from http://ok.gov/sde/sites/ok.gov.sde/files/TLE-MarzanoWhitePaper.pdf

McClellan, C., Atkinson, M., & Danielson, C. (2011). *Teacher evaluator training and certification: Lessons learned from the Measures of Effective Teaching Project*. San Francisco, CA: Teachscape.

McGuinn, P. (2012). *The state of teacher evaluation reform: State education agency capacity and the implementation of new teacher-evaluation systems*. Washington, DC: Center for American Progress.

McTighe, J. (2007). Making the most of professional learning communities. *The Learning Principal*, *3*(8), 4-8.

Measures of Effective Teaching Project. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project's three year study*. Seattle, WA: Bill and Melinda Gates Foundation.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, *79*(4), 33-53.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.

National Commission on Teaching & America's Future. (1996). *What matters most: Teaching for America's future*. Retrieved from http://nctaf.org/wp-content/uploads/2012/01/WhatMattersMost.pdf

Nelson, T. H., Deuel, A., Slavit, D., & Kennedy, A. (2010). Leading deep conversations in collaborative inquiry groups. *The Clearing House, 83*, 175-179. http://dx.doi.org/10.1080/000986503505498

Nelson, T., & Slavit, D. (2008, Winter). Supported teacher collaborative inquiry. *Teacher Education Quarterly*, 99-116.

Nelson, T. H., Slavit, D., Perkins, M., & Hathorn, T. (2008). A culture of collaborative inquiry: Learning to develop and support professional learning communities. *Teachers College Record*, *110*(6), 1269-1303.

New Teacher Project. (2007). *Teacher hiring, transfer and assignment in Chicago public schools*. Retrieved from http://tntp.org/assets/documents/TNTPExecSumm-Chicago.pdf?files/TNTPExecSumm-Chicago.pdf

Papay, J. P. (2012, Spring). Refocusing the debate: Assessing the purpose and tools of teacher evaluation. *Harvard Education Review*, *82*(1), 123-141.

Pearson Education. (2008). *Policy report: Value-added assessment systems*. San Antonio, TX: Pearson Education.

Pennsylvania Department of Education (2012). *Standards Aligned System*. Retrieved from http://www.pdesas.org/Standard/CommonCore

Pennsylvania Department of Education. (2013a). *Guide to PVAAS public reporting.* Retrieved from https://pvaas.sas.com

Pennsylvania Department of Education. (2013b). *Highly Qualified Teacher Guidelines.* Retrieved from http://www.pde.state.pa.us/portal/server.pt/community/highly_qualified_teacher_requirements/8631/highly_qualified_teacher_guidelines/506867

Pennsylvania Department of Education. (2013c). *Measuring Educator Effectiveness*. Retrieved from http://www.education.state.pa.us

Pennsylvania Department of Education. (2013d). *Pennsylvania Value-Added Assessment System (PVAAS) school year 12-13 pilot teacher specific reporting.* Retrieved from https://pvaas.sas.com/

Pennsylvania State Education Association. (2011). *Pennsylvania's new teacher evaluation system*. Retrieved from http://www.psea.org

Reeves, D. B. (2007). *Ahead of the curve: The power of assessment to transform teaching and learning*. Bloomington, IN: Solution Tree.

Reeves, D. B., & Flach, T. (2011, August). Data: Meaningful analysis can rescue schools from drowning in data. *Journal of Staff Development, 32*(4), 34-40.

Reform Support Network. (2012). *Race to the Top at a glance: Evaluation of teacher effectiveness: State requirements for classroom observations* (GS-23f-8182H). Washington, DC: Government Printing Office.

Roberts, M. (2010). *Improving student achievement through professional learning communities* (Doctoral dissertation). Retrieved from http://www.ncsa.org/sites/default/files/media/dissertations/RobertsMindy.pdf

Rockoff, J. E., & Speroni, C. (2010, May). Subjective and objective evaluations of teacher effectiveness. *American Economic Review, 100*(2), 261-66. Retrieved from http://www0.gsb.columbia.edu/faculty/jrockoff/papers/rockoff_speroni_subjective_evals_AEA_PP_final.pdf

Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS); Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299-311.

Sanders, W. L., & Wright, S. P. (2008). *A response to Amrein-Beardsley (2008): Methodological concerns about the education value-added assessment system*. Retrieved from http://www.sas.com/resources/asset/Sanders_Wright_response_to_Amrein-Beardsley_4_14_2008.pdf

Saphier, J. (2012). *Walkthroughs, rubrics, and teacher evaluations: Do they improve teaching*? Retrieved from http://www.rbteach.com

Sass, T. R. (2008). *The stability of value-added measures of teacher quality and implications for teacher compensation policy*. Retrieved from http://www.urban.org/UploadedPDF/1001266_stabilityofvalue.pdf

Schmoker, M. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: ASCD.

Servage, L. (2009). Who is the "professional" in a professional learning community? An exploration of teacher professionalism in collaborative professional development settings. *Canadian Journal of Education, 32*(1), 149-171.

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. http://dx.doi.org/10.1177/0022487111404241

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Washington, DC.: Education Sector Report.

Tomalis, R. (2012). *Path to prosperity*. Retrieved from www.education.state.pa.us

Towe, P. B. (2012). *An investigation of the role of a teacher evaluation system and its influence on practice and professional growth in four urban high schools* (Doctoral dissertation, Seton Hall University). Retrieved from http://scholarship.shu.edu/cgi/viewcontent.cgi?article=2792&context=dissertation

Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: ASCD.

United States Department of Education. (1998). *Goals 2000: Educate America Act*. Retrieved from http://www.ed.gov/PDFDocs/g2kfinal.pdf

United States Department of Education. (2004a). *Elementary and Secondary Education Act of 1965*. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/pg1.html

United States Department of Education. (2004b). *New No Child Left Behind flexibility: Highly qualified teachers*. Washington, DC: Government Printing Office.

United States Department of Education. (2009). *Race to the Top executive summary*. Retrieved from http://www2.ed.gov/programs/racetothetop/executive-summary.pdf.

United States Department of Education. (2010). *ESEA blueprint for reform*. Washington, DC: Evaluation and Policy Development.

Veir, C. A., & Dagley, D. L. (2002). *Legal issues in teacher evaluation legislation: A study of state statutory provisions*. Provo, UT: Brigham Young University Law School.

Wiggins, G., & McTighe, J. (2007). *Understanding by design: A framework for effecting curricular development and assessment* (2nd ed.). Alexandria, VA: ASCD.

Wood, D. R. (2007). Professional learning communities: Teachers, knowledge, and knowing. *Theory into Practice, 46*(4), 281-290

Appendix A

Teacher Evaluation Questionnaire

1. Including the current year, how many years have you been teaching?

- 1 to 3 years
- 4 to 5 years
- 6 to 10 years
- 11 to 15 years
- 16 or more years

2. Indicate your level of education.

- Bachelor's
- Bachelor's plus 15
- Master's
- Master's plus 15
- Master's plus 30
- Doctorate

3. Indicate the number of years you have been in your current position.

- 1 to 4 years
- 5 to 9 years
- 10 to 14 years
- 15 to 19 years
- 20 to 25 years
- 26 to more years

Please use the rating provided on the following questions to describe yourself and the nature of your most recent teacher evaluation experience with the PA standards-based teacher evaluation process. Please consider the entire evaluation process including professional growth, meeting with colleagues and administrators, and planning for instruction and feedback.

4. The PA standards-based teacher evaluation process is an effective evaluation model in determining the quality of my teaching.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

5. The PA standards-based teacher evaluation process has provided useful feedback on my planning and instruction.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

6. The PA standards-based teacher evaluation provides a more accurate picture of my teaching ability.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

7. Since the implementation of the PA standards-based teacher evaluation, I have changed the instructional strategies I use in my classroom due to the implementation.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

8. Since the implementation of the PA standards-based teacher evaluation, the time required for me to plan lessons and prepare for classes has increased.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

9. Since the implementation of the PA standards-based teacher evaluation process, I am more likely to use student achievement data to plan my instruction.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

10. The PA standards-based teacher evaluation rubric provides guidelines for developing future lessons.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

11. The PA standards-based teacher evaluation requires me to utilize a variety of instructional strategies, such as high level questioning skills.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

12. The PA standards-based teacher evaluation process has had no impact on my instructional practices.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

For the following statements, refer to your experiences with the impact of the PA standards-based teacher evaluation process on collaboration.

13. Teachers in the building have professional conversations regarding building level student achievement data such as: PSSA/Keystone scores, PVAAS growth, AP participation, and SAT results.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

14. Teachers in my building collaboratively review student work to share and improve instructional practices.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

15. I am collaborating more with other teachers because of the PA standards-based teacher evaluation process.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

16. In the past year I have had more conversation about ways to improve my teaching methods with administrators.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

17. Teacher learning is supported through a combination of strategies (e.g., team planning, professional learning communities, peer observations, examination of student work).

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

18. Teachers in my building meet to discuss ways to improve their teaching and learning.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

19. The PA standards-based teacher evaluation has had no impact on collaboration with my peers.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

For the following statements, refer to your experiences with the impact of the PA standards-based teacher evaluation process on your professional growth.

20. In general, I believe that the professional development that I participate in improves the instructional strategies I use in my classroom.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

21. Through my career, my professional growth work directly impacts student achievement and success.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

22. The implementation of the PA standards-based teacher evaluation has helped me establish goals for my professional growth.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

23. My evaluations have been used to direct my professional development activities.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

24. I choose a professional growth option that directly affects my classroom instruction.

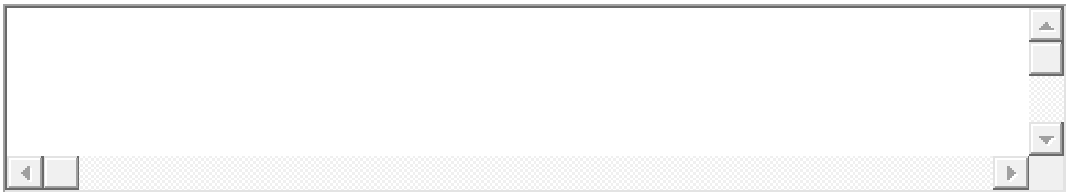
- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

25. With the PA standards-based teacher evaluation process, my teacher evaluations usually contain a good balance between encouraging positive comments and suggestions for improvement and growth.

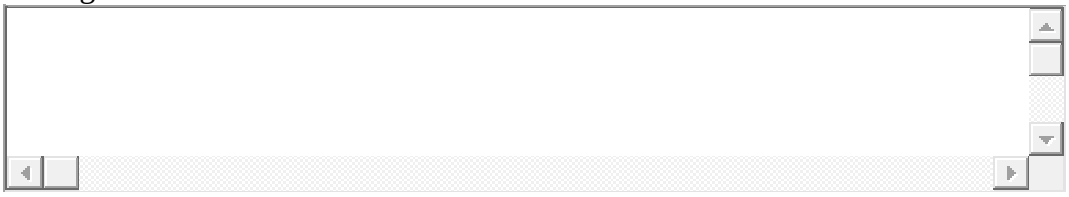
- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

Open ended responses. Please answer the following questions while reflecting on your experience with the PA standards-based teacher effectiveness evaluation process.

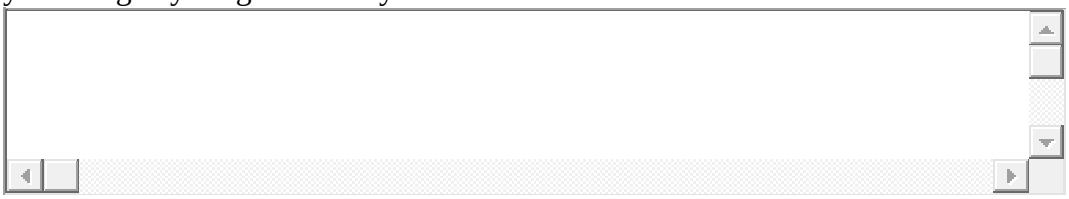
26. How has the PA standards-based teacher evaluation impacted your classroom instruction?



27. How has the teacher evaluation process promoted your collaboration with your colleagues?



28. How has your professional development been affected by the evaluation system? Are you doing anything differently?



Appendix B

Interview Questions

1. Reflect on the Danielson Framework as part of the standards-based teacher effectiveness model. How has it influenced your teaching?

2. Describe the use of building data in the evaluation process.
    - How have they impacted your instruction?
    - How have they altered your professional collaboration?
    - To what extent have they influenced your professional growth?

3. How will your professional goals for next year be influenced by this year's evaluation?

4. Describe specific changes that you have made to your instructional practices from the feedback received since the implementation of the standards-based teacher effectiveness model.

5. Describe any differences in your collegial interactions because of the standards-based evaluation system.

Appendix C

IMMACULATA UNIVERSITY RESEARCH ETHICS REVIEW BOARD REQUEST FOR PROTOCOL REVIEW--REVIEWER'S COMMENTS FORM (R1297)

Name of Researcher: Marshall Hoffritz

Project Title: Teacher Perceptions of Pennsylvania's Teacher Effectiveness Model

Reviewer's Comments:

Your proposal is approved. You may begin your research or collect your data. PLEASE NOTE THAT THIS APPROVAL IS VALID FOR ONE YEAR **(365 days)** FROM DATE OF SIGNING.

Reviewer's Recommendations:

_____Exempt	__X__ **Approved**
_____Expedited	_____ Conditionally Approve
_____Full Review	_____ Do Not Approved

Thomas J. O'Brien August 26, 2013