STARS

Electronic Theses and Dissertations, 2004-2019

2015

A Study of the Impact of Brevard Public School's Peer Coaching Model on Student Achievement Outcomes and Teacher Evaluation Results

Debra Pace
University of Central Florida

Part of the Education Commons

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

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

A STUDY OF THE IMPACT OF BREVARD PUBLIC SCHOOL'S
PEER COACHING MODEL ON STUDENT ACHIEVEMENT OUTCOMES
AND TEACHER EVALUATION RESULTS

by

DEBRA P. PACE
B.A. Auburn University, 1985
M.Ed. University of Central Florida, 1989

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

Spring Term
2015

Major Professor: Barbara A. Murray

ABSTRACT

The focus of this research was to evaluate the effectiveness of a sustained
professional development initiative, Peer Coaching, on improving teacher performance
and student achievement. Developed in a large school district on the east coast of central
Florida, Brevard's Peer Coaching Model (BPCM) was implemented during the summer
following the 2011-2012 school year to support teachers and administrators with
implementation of a new teacher evaluation system designed to promote continuous
improvement in teaching and learning. Teams of highly effective teachers were chosen
from each school to participate in nine days of training and follow-up over the course of
the study, in order to encourage improvement in their own instructional practice and the
practice of their peers through greater understanding of the evaluation framework and
observation rubrics, enhanced collaboration, and peer observation and coaching.

Quantitative data were obtained from Brevard Public Schools Office of Testing
and Accountability for professional practices evaluation scores and value-added results.
Professional practices scores are determined by trained and certified school
administrators, assigning up to three points across seven dimensions for a total possible
of 21 points. The school district assigns all teachers a three-year aggregated VAM score,
based on results from Florida's standardized test for reading and mathematics, FCAT,
and BPS teachers earn a teacher-aggregated VAM (TAV), a non-FCAT teacher-
aggregated VAM (NFTAV), or a school-aggregated VAM (SAV) depending on their
grade level and content area assignment. Results for teachers who participated in a
minimum of six days of BPCM training, before, during, and after the training's

implementation were analyzed and compared with the results of teachers who did not participate in BPCM training. The sample consisted of 174 BPS teachers, BPCM participants, similar in demographics and years of experience to the comparison population of teachers, non-BPCM participants.

Findings indicated that BPCM participants demonstrated a significant increase in professional practices scores during and after the training, with the most significant growth occurring after year one. In addition, BPCM participants earned significantly higher professional practices scores compared to teachers in the non-BPCM population, before, during, and after the professional learning experience. Differences in value-added results, however, were not statistically significant. Although both professional practices scores and value-added scores improved for BPCM and non-BPCM teachers during and following the training, changes in student outcomes were not statistically significant. These findings replicated previous findings that demonstrated a positive impact on instructional practices following implementation of a peer coaching professional learning model but limited impact on student achievement.

Implications for practice and recommendations for future research were provided for Brevard Public Schools and other school districts considering development and implementation of high quality professional learning to promote improvements in teaching and learning. Professional development models represent a significant investment of resources requiring careful planning for effective program evaluation to ensure that the desired outcomes of changes in practice and increases in student achievement are recognized.

To my mother and father, who have modeled the core values of a strong work ethic, a

commitment to personal excellence, and a love of learning throughout my life.

To my husband and children, who have demonstrated patience, support, and

encouragement when the days and the nights grew long and we all grew weary.

And to Granny Pete, who loved me, believed in me, and pushed me to give everyone and

everything my very best effort.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AIR: American Institutes for Research

BPCM: Brevard Peer Coaching Model

BEST: Brevard's Essential Strategies for Teaching

BPS: Brevard Public Schools

CET: Clinical Educator Training

ELL: English Language Learners

FCAT: Florida Comprehensive Achievement Test

FEAPs: Florida Educator Accomplished Practices

FLDOE: Florida Department of Education

IPPAS: Instructional Personnel Performance Appraisal System

NFTAV: Non-FCAT Teacher Value-added Score

NSDC: National Staff Development Council

PAR: Peer Assistance Review

SAV: School Value-added Score

SAT: Stanford Achievement Test

SGIC: Student Growth Implementation Committee

SPSS: Statistical Package for Social Sciences

SRI: Scholastic Reading Inventory

TAV: Teacher Value-added Score

TPLP: Teachers Perceptions of Professional Learning

TVAAS: Tennessee Value-added Assessment System

VAM: Value-added Model

CHAPTER 1
INTRODUCTION

Background of the Study

In 2011-2012, Brevard Public School District (BPS), a large school district
located on the East Coast of Central Florida serving more than 70,000 students with over
9,000 employees and 84 schools (excluding charter schools), implemented a new
evaluation system for all instructional personnel in compliance with revisions to Florida
State Statute 1012.34 (2012) and requirements inherent with the acceptance of the
Federal Race to the Top Grant (American Recovery and Reinvestment Act, 2009).
Required elements included a focus on research-based instructional practice, a cycle of
continuous improvement, a multi-metric system, and the inclusion of student
achievement growth and performance in the final evaluation rating.

Prior to 2011-2012, BPS, on its own initiative, spent two years developing and
implementing a consistent instructional model based on current educational research,
including the state-adopted works of Marzano (2007) and Danielson (2007). Six
fundamental components of effective instructional design, planning, assessment, and
evaluation were developed for Brevard's Effective Strategies for Teaching [B.E.S.T.]
(2009). School district staff worked hand-in-hand with school administrative teams and
prepared teachers to build an understanding of the components of effective instruction
with the underlying goal of improving student achievement.

For the summer of 2012, a professional learning model to develop and implement
peer coaching in every school was created by the BPS Human Resources Division and

the Office of Leading and Learning to support implementation of the school district's Instructional Personnel Performance Appraisal System [IPPAS] (Brevard Public Schools, 2011) and to build teacher-leadership at each school using the Brevard Peer Coaching Model (BPCM). Participants in BPCM included three teachers from each school chosen by the principal based on a highly effective evaluation in the previous year and a record of completion of Clinical Educator Training (CET). BPCM participants worked together in teams over two days to improve their understanding and identification of elements of exemplary instructional practice found in the IPPAS rubrics (2011) and to develop their skills in providing meaningful feedback to fellow teachers to enhance efforts to improve performance. Two additional days of BPCM follow-up training were provided for peer coaches the following school year, 2012-2013, in October and in January to reinforce the training concepts, provide additional practice opportunities, and confirm expectations for providing support to their fellow teachers. In the following summer of 2013, these same peer coaches participated in additional BPCM training designed to develop their capacity to train other teachers at their sites in BPCM peer coaching skills. Follow-up and support for BPCM implementation continued in year two with training from a professional learning consultant (Abrams, 2009), and specific sessions on high-yield strategies found in the IPPAS rubrics. These included differentiated instruction, implementation of technology, higher order questioning, and the inquiry method. The Office of Leading and Learning provided funds to pay teachers chosen to attend summer training and to cover substitute teacher expenses for BPCM participants to attend the follow-up training held

during the school year and to observe colleagues' classrooms and provide coaching feedback (J. Respess, personal communication, March 24, 2014).

Previous researchers of peer coaching models have focused their studies on teacher and administrator perceptions about the peer coaching experience and its impact on instructional practice. They have typically relied on observations, questionnaires, and surveys to gather data from both teachers and leaders, reporting generally positive findings (Hill & Rapp, 2012; Johnson & Fiarman, 2012; Kohler, Crilley, Shearer, & Good, 1997; Sparks & Bruder, 1987). Each of the studies examined utilized a small sample size, limiting the generalizability of the results; and limited attention was directed toward measuring student performance for teachers who participated in the peer coaching professional learning program.

Bowman and McCormick (2000) conducted a study at a large Midwestern university to analyze the difference between a peer coaching model and a traditional supervision model for pre-service teachers. Extensive training on developing a common language of effective practice was provided for both the control group and the experimental group along with training on providing timely, specific feedback about effective practice for the study participants. The authors reported positive results for collegiality, technical feedback, analysis of application, adaptation to students, and personal facilitation, citing commonly reported research findings about the impact of professional collaboration on improving instructional practice. They concluded that the increased frequency of feedback for those working with peer coaches, as well as the increased collegiality, helped pre-service teachers incorporate new skills and strategies

3

into their toolboxes. They advocated for continued study of the use of peer coaching in field experiences for pre-service teachers.

Murray, Ma, and Mazur (2008) conducted a small study after the implementation of a peer coaching program in the Appalachian Mountains, examining both observational data and student achievement in mathematics. The results indicated that although study participants welcomed the collaborative feedback, there was no significant impact on student performance. The authors submitted several recommendations for future study into the impact of peer coaching:

- Clearly define peer coaching and identify roles and expectations of participants for consistent implementation.

- Study the results for a longer period of time (the study looked at data after a single semester).

- Provide additional planning time for teachers participating in the project to reflect deeply on their work.

Although Johnson and Fiarman (2012) did not review specific student achievement data in their study of a peer assistance and review process conducted in seven school districts across the country, they indicated that both principals and teachers reported improvements in instructional practice system-wide. The authors further reported that teachers chosen to serve as peer consultants described their work as both "challenging and rewarding" (p. 22). The researchers offered several recommendations for future study:

- Carefully select the teachers who will be trained as peer review consultants. To be effective, these teachers must be perceived as highly effective teachers in the eyes of their peers.

- Specifically identify guidelines to direct the work of the peer review consultants.

- Develop and implement clear and specific expectations for effective instructional practice through evaluation rubrics, and provide training to both administrators and peer review consultants in use of the rubric to provide specific feedback to teachers on performance.

- Provide adequate training to peer review consultants on the change process, building trust, managing a caseload, efficient record-keeping, and effective feedback, along with other essential coaching skills that are separate and unique from training they've received about how to be an effective teacher.

- Emphasize both prongs of a successful peer review program: evaluation and assistance.

- Effective peer review consultants have an evaluative role in moving poor performers out of the profession, but they play a valuable role in helping all teachers to improve their practice through specific feedback and peer support.

The recommendations and results of this and additional prior research on the topic of peer coaching and effective professional development guided the development of the research design for this study.

Statement of the Problem

To date, limited research has been conducted to evaluate the effect of professional preparation for peer coaches on improving student outcomes and teacher practice over an extended period of time. Rather, researchers have primarily examined teacher perceptions about peer coaching as a professional learning model with limited investigation into measurable changes in instructional practice and impact on student achievement. Because this training model represented a significant investment of school district resources, raising the question of return on investment, the study was vital to ensure that time and money were spent on efforts that have a significant positive impact on student performance.

Purpose of the Study

The purpose of this study was to analyze whether one public school district's investment in professional learning for peer coaches resulted in improved instructional practice for the participants and their peers and increased levels of student achievement as measured by Florida's value-added model, [VAM] (2013).

Significance of the Study

In the spring of 2003, the Florida Department of Education (FDOE) developed and implemented the Professional Development Evaluation Protocol in response to Florida Statute 1012.98. This protocol established the purpose of professional development systems as "[to] increase student achievement, enhance classroom instructional strategies that promote rigor and relevance throughout the curriculum, and prepare students for continuing education and the workforce." (FDOE, 2010a). The standards were revised in 2010 to align with legislative changes and new developments in educational research and practice, but the importance of the evaluation of professional development in terms of both participant perception and impact on instructional practice and student outcomes remained unchanged; at a school district level, standard 3.4, Evaluating, specified seven actions that should be taken to ensure that professional development resources, both time and money, were directed towards learning activities that make a difference in teacher and student performance.

These professional development protocol standards are consistent with 21st century research (Guskey, 2002; Haslam, 2010). However, findings from the FDOE's Professional Development System Evaluation Protocol, Third Cycle/Second Year Technical Report for the 2011-2012 School Year (2012) indicated that "Districts need continued improvement and assistance in evaluating the impact of professional learning" (p. 4). Of four strands included in the FDOE Protocol Standards (planning, learning, implementing, and evaluating), evaluation earned the lowest average score across the 17 Florida school districts reviewed in 2011-2012.

The need for increased attention to teacher implementation of new learning and its impact on student performance are areas for growth at national, state, and local levels. This research was conducted to improve current understanding and practice in this regard. This study added to the body of knowledge on the effectiveness of peer coaching on improving teacher and student performance and provided meaningful data to guide future decisions about resource allocation and investment in professional development.

Definition of Terms

Brevard Public Schools (BPS): A large school district in Brevard County, Florida.

The Brevard Peer Coaching Model (BPCM): A process through which two or more professional colleagues agree to work together to observe teaching and learning; exchange specific feedback about what was observed; and reflect on the feedback in order to improve instructional practice and increase student achievement.

Clinical Educator Training (CET): Required for all instructional personnel in Florida who serve as mentors or supervisors for teacher preparation students during field experiences and internships. The program includes training to develop clinical skills for the analyzing teacher, including how to provide effective feedback on instructional performance and develop professional improvement plans as a part of reflective professional practice.

Florida Educator Accomplished Practices (FEAPs): Florida's established standards for effective professional educators as established by State Board Rule 6A-5.065 (2010); required elements in state-approved teacher evaluation systems.

Instructional Personnel Performance Appraisal System (IPPAS): Brevard Public Schools' state-approved teacher evaluation system; comprised of the evaluation of professional practices, professional growth plan development and implementation, collaboration and mutual accountability, and student achievement outcomes/value-added growth measures.

Professional Practices Scores: Represent the summative evaluation rating assigned to teachers by the supervising administrator (principal, assistant principal, or director), using the IPPAS rubrics to differentiate Distinguished, Proficient, Professional Support Needed, and Unsatisfactory levels of performance in formal and informal classroom observations. The professional practices score reflects a teacher's overall performance on seven instructional dimensions and represents 21 points of a teacher's final summative evaluation rating.

Student Growth Implementation Committee (SGIC): Comprised of 27 members, including teachers, school administrators, district administrators, post-secondary stakeholders, parents, and business owners, and tasked with planning for the sales tax referendum and other key initiatives.

Value-added Scores (VAM): Represent the difference between students' expected achievement on a standardized assessment and the actual performance of the students who share similar characteristics; when aggregated, the value-added model represents a teacher's contribution to the students' learning. For BPS teachers, VAM equates to 35 points in the final summative evaluation rating in the baseline and year one of the study, and 45 points in year two of the study.

Teacher Value-added Scores (TAV): Represent the difference between an individual teacher's students' expected achievement on the Florida Comprehensive Assessment Test (FCAT) and the actual performance of students who share similar characteristics. For teachers in this study, TAV applies to teachers of reading/language arts and/or math, Grades 4-10 and represents their students' performance, aggregated over a three year period to reduce the impact of measurement error.

School Value-added Scores (SAV): Represent the difference between a school's students' expected achievement on the Florida Comprehensive Assessment Test (FCAT), reading and mathematics, and the actual performance of students who share similar characteristics, aggregated over a three-year period. For teachers in this study, SAV applies to teachers assigned to students in Grades K–2 and Grades 11–12 who are not annually assessed by the Florida Comprehensive Assessment Test (FCAT) or students in Grade 3 who are assessed using the Florida Comprehensive Assessment Test (FCAT) but do not have baseline results to establish predictive scores.

Non-FCAT Teacher Value-added Scores (NFTAV): Represent the difference between an individual teacher's students expected achievement on the Florida Comprehensive Assessment Test (FCAT) and the actual performance of students who share similar characteristics. For teachers in this study, NFTAV applies to teachers in Grades 4-10 assigned to content areas other than reading/language arts

and/or mathematics and represents their students' performance, aggregated over a three-year period to reduce the impact of measurement error.

<div align="center">Conceptual Framework</div>

The conceptual model underlying Brevard Public Schools' Peer Coaching Model (BPCM) initiative was grounded in the work of Deming, an American statistician who achieved recognition for his work with the Japanese automobile industry following World War II. He lived from 1900–1993, serving in his lifetime as a consultant to the U.S. Secretaries of War and Defense, an advisor for the U.S. Bureau of Budget, and faculty member at both Yale University and New York University (Petersen, 1999). Deming defined systems, emphasizing (a) the importance of interaction between the elements of an organization in determining the organization's level of performance and (b) the significance of a focus on quality and the customer at every level of the organization (Gruska, 2000). His primary work, however, focused on leadership that promoted collaboration and participation of all stakeholders to achieve a shared vision. He used the analogy of an orchestra conductor to describe effective participatory leadership: "An orchestra is judged by listeners, not so much by illustrious players, but by the way they work together. The conductor, as manager, begets cooperation between the players, as a system, every player to support the others." (*System of profound knowledge*, 2014, para. 3). To achieve success, organizations implement structures to support continuous improvement, employees learning from each other, teamwork, and

distributive leadership, defined in Deming's 14 points for the transformation of management.

Since Deming's death, other authors have built on his work to further clarify the meaning of participatory leadership. Rook and Torbert (2005) offered the following definition:

> The participatory leadership paradigm is based on respect and engagement. It constructively focuses energy in every human to human encounter. A more advanced, more democratic and more effective model of leadership, it harnesses diversity, builds community, and creates shared responsibility for action. It deepens individual and collective learning yielding real development and growth. (p. 66)

According to Rook and Torbert (2005), participatory leadership is an effective strategy to channel organizational creativity and passion toward a common goal, creating and sustaining successful organizational development and change. Leaders functioning at this level establish collegial networks for mutual mentoring and build learning communities that support peer-to-peer professional growth and development.

This conceptual model of participatory leadership has also been evident in other adult learning frameworks, including the work of Drago-Severson and Cuban cited by Burke, Marx, and Lowenstein (2012) in their study of leadership development. Burke et al. (2012) argued that "Leaders who understand their own development as learners, acting in social organizational systems, will recognize each participant as a learner whose individual development can be a key component to building the leadership capacity of the

larger system" (p. 114). Leaders with a vision of improved student performance establish

structures whereby educators influence other educators to improve instructional practice

in collaborative professional learning communities.

Other conditions used to describe the participatory leadership conceptual

framework include more effective utilization of the level of expertise in an organization,

improving levels of commitment within an organization, and targeting issues related to

organizational fit within an entity. Participatory leadership focuses on teamwork, sharing

of differing opinions and beliefs, and continuous learning through collaboration in order

to more effectively achieve an organization's strategic goals and initiatives (Kezar,

2001).

Brevard's Peer Coaching Model (BPCM) was based on the premise of

participatory leadership. It was founded on the principle that the power to improve

teaching and learning lies within educators working together in a system focused on

continuous improvement through teamwork, collaboration, and effective feedback.

<u>Research Questions and Hypotheses</u>

1. To what extent do the VAM scores of those teachers who participated in peer

   coaching training (BPCM) differ from the VAM scores of the population of

   Brevard Public School teachers who did not participate in BPCM?

   H_{01} There is no statistically significant difference in the VAM scores of those

   teachers who participated in peer coaching training (BPCM) from the VAM

scores of the population of Brevard Public School teachers who did not participate in BPCM.

2. To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM?

   H_{02}: There is no statistically significant difference in the professional practices scores of those teachers who participated in peer coaching training (BPCM) from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM.

3. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

   H_{03}: There is no statistically significant difference in the VAM scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

4. To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

14

$H_{04:}$ There is no statistically significant difference in the professional

practices scores of those teachers who participated in peer coaching training

(BPCM) from prior to training, at the end of the first year after initial training,

and at the end of the second year after additional follow-up training.

<u>Delimitations</u>

This study was delimited by the following factors:

1. Data collection and analysis were delimited to teachers from only one large

    Florida school district, Brevard Public Schools, Brevard County, Florida.

2. Data collection and analysis were further delimited to teachers from Brevard

    Public Schools, Brevard County, Florida, who were selected by their

    principals to participate in the school district-provided Peer Coaching

    Training.

3. This study was delimited to teachers who, having completing at least six days

    of peer coaching training and follow-up, earned a VAM (Value-added model)

    score and professional practices scores in three consecutive school years,

    2011-2014.

4. This study was delimited to application of the state-mandated model for

    inclusion of student achievement data from standardized tests in teacher

    evaluation.

<u>Limitations</u>

This study was limited by the extent to which the following parameters were applied:

1. The student data assigned to each teacher for the purpose of this study was based solely on the Florida Comprehensive Achievement Test (FCAT), a standardized assessment unique to the state of Florida.

2. Changes in school administrators between the Fall of 2011 and Spring of 2014 were beyond the researcher's control.

3. School administrators rated their own teachers higher on observation of practice than outside observers, in accordance to current research (Ensuring Fair and Reliable Measures of Effective Teaching, 2013).

4. Non-FCAT teacher aggregated value-added model scores (NFTAV) and school-aggregated value-added model scores (SAV) assigned to teachers of non-FCAT tested grade levels or content areas complied with the state model for attribution of student achievement data in teacher evaluation but did not reflect student performance in the teachers' areas of expertise or classroom experience.

<u>Assumptions</u>

This study included the following assumptions:

1. Administrators received the required training offered by Brevard Public Schools in the use of the IPPAS rubrics to evaluate the professional practices of teachers in the study.

2. Administrators understood the differences among the rubric classifications for the observation of professional practices, including Distinguished, Proficient, Professional Support Needed, and Unsatisfactory.

3. Administrators evaluating the professional practices of the teachers in the study utilized the evaluation rubric without bias and recorded the results accurately.

4. Teachers chosen to participate in the peer coaching professional learning model did so willingly and with the intent of improving their instructional practice and the practice of their peers.

5. Teachers chosen to participate in the peer coaching professional development model made a commitment to complete the two-year training regimen with the intent of improving their instructional practice and increasing their student outcomes.

6. The state and the school district accurately attributed the student data assigned to each teacher during the 2011-2012, 2012-2013, and 2013-2014 school year.

<center>Methodology</center>

<center>*Research Design*</center>

A quasi-experimental design was chosen for this study in order to determine if there were differences in the VAM scores and professional practices scores (a) between the two sample populations, those teachers from Brevard Public Schools who participated in the peer coaching training (BPCM) and those teachers from Brevard Public Schools who did not and (b) between VAM and professional practices scores of the convenience sample chosen to participate in BPCM prior to training, at the end of year one, and at the end of year two of the study. The study was intended to improve on previous research efforts in three specific areas: size of the sample, use of a value-added approach to evaluate growth in student achievement, and three complete years of student achievement and professional practices data to evaluate training impact on both teacher performance and student learning.

The dependent variables in the study were observation of professional practices, as measured by administrator evaluations, and student achievement results on FCAT, as reported by VAM, the state's value-added measure. Although administrator evaluations of professional practices are subjective in nature, advances in the use of rubrics have contributed to more clearly defining effective instructional practice and training for administrators. Rubrics plus the use of video clips, instructional rounds, and expert consultants have resulted in improved inter-rater reliability. In addition, Brevard Public

<center>18</center>

Schools used three years of aggregated FCAT data to calculate VAM for teachers and schools in order to minimize the impact of the standard error of measurement.

Population

Principals selected three teachers from their schools to participate in BPCM, establishing the convenience test sample. Candidates were required to have highly effective annual evaluations and have completed state-approved Clinical Educator Training (CET) in advance; the original static training group included 285 participants, a group of teachers diverse in gender, years of experience, and level of expertise (elementary, middle, or high school). The study's sample was further refined by the requirement that teachers attend a minimum of six of the nine days of peer coaching training offered between the summer of 2012 and the spring of 2014, reducing the sample to 174 participants. Results for teachers who dropped out of BPCM during the course of the study were excluded from the comparative analysis between results for BPCM participants and results for the population of BPS teachers to eliminate potential bias in the scores from partial completion of the training.

Data Collection

The researcher proposed to analyze and report aggregate professional practices and FCAT VAM results for members of both the sample selected to participate in BPCM over a three-year period and the comparison group, the population of BPS teachers not selected for BPCM, as follows: (a) 2011-2012: Baseline results, prior to BPCM training;

(b) 2012-2013: year one results, following initial BPCM training and implementation; and (c) 2013-2014: year two results, following additional BPCM training and monitoring. Both the sample and the population were delimited to teachers with three years of VAM data in order to limit researcher bias.

Data Analysis

Quantitative data for professional practices and student achievement (VAM) were analyzed and reported for range, mean, standard deviation, skewness, and kurtosis for each year (baseline, year one, and year two). Mean results from the convenience sample were compared to determine if a statistically significant difference existed using a paired samples t test. Mean results of the convenience sample were also compared to the mean results of the population of teachers in Brevard Public Schools, using an independent samples t-test, to determine if a statistically significant difference existed. Table 1 provides a summary of the variables, data sources, and methods of analysis used to answer each of the four research questions.

Table 1

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

Research Questions	Variables	Data Source	Analysis
1. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from the VAM scores of the population of teachers in Brevard Public Schools?	Dependent: FCAT VAM scores, BPCM participants Independent: FCAT VAM scores, Population of BPS teachers	VAM scores	Independent samples *t* test
2. To what extent do the Professional Practices scores of those teachers who participated in peer coaching training (BPCM) differ from the Professional Practices scores of the population of teachers in Brevard Public Schools?	Dependent: Professional Practices scores, BPCM participants Independent: Professional Practices scores, Population of BPS teachers	Professional Practices scores	Independent samples *t* test
3. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?	Dependent: FCAT VAM scores, years 1 and 2 Independent: FCAT VAM scores, prior year	VAM Scores	Paired samples t-test
4. To what extent do the Professional Practices scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?	Dependent: Professional Practices scores, years 1 and 2 Independent: Professional Practices scores, prior year	Professional Practices Scores	Paired samples t-test

Summary

The researcher strived to find results that duplicated those of previous studies regarding the positive effects of a peer coaching professional development model in increasing collaboration among educators and improving instructional practice. By building both human and social capital, peer coaching invokes previously proven tenets of effective professional development, including attention to making it ongoing and connected to practice, with modeling and supported demonstration. This study proposed to further explore changes in instructional practice and student achievement outcomes, as measured by the state-adopted formula for VAM, in order to effectively evaluate the return on investment of a school district's peer coaching professional development model.

Organization of the Study

This research study is presented in five chapters. Chapter 1 includes the background of the study, the statement of the problem, purpose of the study, definition of terms, conceptual model, research questions and hypotheses, limitations, delimitations, methodology, and the organization of the study. Chapter 2 presents a review of the literature relevant to the study, including a review of the models of effective professional learning, peer coaching as a professional learning model designed to improve teaching and learning, and measures used to evaluate teacher performance, including teacher observation using Marzano's (2007) and Danielson's (2007) instructional practice frameworks and Florida's approach to the value-added model as a measure of teacher

impact on student achievement. Chapter 3 describes the methodology used for this

research study including an introduction, the research design, the selection of study

participants, instrumentation, data collection, data analysis, and a summary. Chapter 4

presents the study's findings including demographic information, the results of data

analysis for the four research questions, additional data analysis, and a summary.

Chapter 5 provides a summary of the entire study, discussion of the findings, implications

of the findings for practice, recommendations for further research, and conclusions.

CHAPTER 2
REVIEW OF LITERATURE AND RELATED RESEARCH

Introduction

Continuous learning is recognized as a critical attribute of professionals in many

career fields. Doctors, lawyers, engineers, accountants, and teachers are required to be

licensed in their fields and must demonstrate continued learning in order to stay current in

the latest advancements in knowledge, technology, and skill and renew professional

licensure (Professional Career Development, 2014). The work of Deming on

participatory leadership, this study's conceptual framework, also places considerable

emphasis on the importance of continuous learning within an effective organization,

recognizing the value of professionals learning with and from each other (Kezar, 2001).

This chapter presents the rationale for conducting further research on the impact

of a school district's peer coaching model on teacher and student performance. The

review of literature has been organized around (a) professional learning models and their

roots in adult learning theory, (b) peer coaching as a high quality professional learning

model designed to improve teaching and learning, and (c) measures used to evaluate

teacher performance, including teacher observation using Marzano's (2007) and

Danielson's (2007) instructional practice frameworks and Florida's approach to the

value-added model as a measure of teacher impact on student achievement.

Multiple sources were used to search the literature, including ERIC, ProQuest,

dissertations and theses, Internet sources, and references including journals, periodicals,

books, published reports, and professional interviews. Appendix A contains a summary

of the search parameters used including key terms and scholars identified through the research.

<center>Principles of Effective Professional Learning</center>

Effective professional learning is grounded in the research related to adult learning. Knowles (1980) identified the following unique characteristics of the adult learner: a desire to be self-directing, a richer bank of experience on which to draw to support new learning, a need to see the purpose and application of new learning in order to create motivation and ensure readiness to learn, and an orientation focused on performance as the most productive means of developing competence. Merriam and Caffarella (1999) further clarified the need of adult learners to operate in a learning organization that provides strong systemic support and motivation for quality adult learning. Learning organizations are characterized by an acceptance of change as a way of life and work; adults routinely examine problems of practice and work collaboratively to discover solutions through inquiry and action-research (Merriam & Caffarella, 1999).

There has been some debate in the research as to the use of the terms "professional development" and "professional learning." The National Staff Development Council (NSDC) first published guiding standards for quality professional development systems in 2001. The standards, built on the known principles of adult learning, were developed with representation from a diverse group of leading organizations from education, government, and business, including the National Education Association, the American Federation of Teachers, the American Association

<center>25</center>

of School Administrators, the National School Board Association, the Council of Chief State School Officers, and the U. S. Department of Education. The standards were categorized by context, process, or content and provided clear guidelines to help school districts build high quality professional development systems with targeted focus on research-based elements, including learning communities, leadership, use of data to plan and monitor professional learning, collaboration, and equity (Hirsch, 2007). The NSDC standards defined critical attributes of effective professional learning which included (a) a comprehensive approach focusing on improving instructional effectiveness to increase student learning, (b) a focus on continuous improvement, (c) employment of skilled facilitators working with learning teams over a sustained period of time, and (d) use of coaching to support implementation of new knowledge and skills in the practice of teaching (NSDC, 2001, p. 1). The terms, professional development and professional learning, were used interchangeably in the original work of the NSDC, and the standards were widely adopted by states, universities, and school districts.

The NSDC later changed its name to Learning Forward, the Professional Learning Association, and collaborated with over 40 agencies to revise and rename the standards in 2011, calling them Standards for Professional Learning, guided by the following principle: "Increasing the effectiveness of professional learning is the leverage point with the greatest potential for strengthening and refining the day-to-day performance of educators" (Standards for, 2011, p. 1). The revised standards emphasized the importance of teachers (a) having influence over the development of the content, context, and evaluation of their own professional learning, (b) being empowered as teacher-leaders

dedicated to professional growth in the continuous effort to provide high quality learning

opportunities for all children. Day & Sachs (2004) provided a definition of professional

development that incorporated the concept of professional learning and was useful in

guiding this study:

> . . . all natural learning experiences and those conscious and planned
>
> activities which are intended to be of direct or indirect benefit to the
>
> individual, group or school and which contribute . . . to the quality of
>
> education in the classroom. It is the process by which, alone or with
>
> others, teachers review, renew and extend their commitment as change
>
> agents to the moral purposes of teaching and by which they acquire and
>
> develop critically the knowledge, skills and emotional intelligence
>
> essential to good professional thinking, planning and practice with
>
> children, young people and colleagues through each phase of their
>
> teaching lives. (p. 34)

One issue revealed in the literature about principles of effective professional

learning focused on the manner in which adult learning impacts student learning and the

fact that determining a direct link between professional development and student

achievement has proven difficult (Borko, 2004; Gusky, 2002; Joyce & Showers, 2002;

Yoon et al., 2007). Effective professional learning models are built on a sequential series

of actions, starting with professional development activities that are focused, tightly

structured, and directly tied to curriculum content that is current and relevant; followed

by support for the translation of the professional development into new knowledge and

improved skills through modeling, practice, and coaching; implementation of the new knowledge and skills in classroom instruction; resulting in increases in student learning (Yoon et al., 2007). Many factors along the way, including availability of resources, quality of the adult learning experiences, level of teacher motivation, and organizational acceptance of change, influence final outcomes. "Given these requirements, it is unsurprising that few rigorous studies address the effect of professional development on student achievement" (Yoon et al., 2007, p. 5).

Despite the research evidence demonstrating the moderate positive impact that high quality professional development has on student achievement (21 percentile points on average) according to Yoon et al. (2001), the United States has historically underfunded and underemphasized quality professional learning experiences for teachers (Darling-Hammond, 2010). In contrast to the ongoing, job-embedded professional development activities found in top-ranked nations such as Finland, Singapore, and Sweden (Darling-Hammond, 2010), professional development for American educators has typically been short in duration, focused on teacher behaviors and skills, and without follow-up or support for implementation of new knowledge (Dubois, 2012). Yoon et al. (2007) noted that high quality learning experiences proven to impact teacher practice and student achievement have a duration of more than 14 hours. Darling-Hammond, Wei, Andree, Richardson, and Orphanos (2009) noted that such high quality learning experiences are specifically connected to student learning and curriculum content, align with school improvement goals and priorities, and provide time and structured opportunity for teachers to collaborate throughout the implementation stage. Coaching

during the implementation stage also increases application of new skills in the classroom, along with training experiences that include description, modeling, practice, and follow-up (Joyce & Showers, 2002).

Hawley and Valli (1999) conducted a widely noted metasynthesis of the relevant research on effective professional development in the late 20th century, developing nine guiding principles evident throughout the literature and underlying the NSDC standards developed in 2001 and revised in 2011. These authors observed that quality professional learning experiences for educators should include the following: (a) a focus on the standards students will be expected to learn, (b) an analysis of the gap between what students currently know and what they need to know and be able to do, (c) an opportunity for teachers to have input into what they are going to learn and how they will engage in the learning process, (d) a job-embedded approach with opportunities for collaboration and problem-solving; systemic support and emphasis on continual learning, (e) a variety of data elements to assess improvements in teaching and learning, (f) an understanding of the underlying theory and principles behind the new knowledge and skills, and (g) a focus on systemic change to improve student performance Hawley & Valli, 1999).

Getting It Right, a professional learning model illustrating the nine guiding principles, was implemented with success in Australia (Ingvarson, 2005). The reform strategy focused on improving elementary mathematics instruction through comprehensive, sustained professional learning with the support of a quality curriculum base and job-embedded coaching offered by specialist teachers at each school site. The specialist teachers were identified by their historic performance in providing effective

teaching and learning experiences for students, and they were trained at the district level for consistency of implementation. In addition, school principals were provided training in effective use of the specialist teachers to support implementation of the new curriculum tool. Ingvarson, in studying the model's implementation, examined data collected through observations, interviews, and surveys, concluding that both teachers and principals believed the model was having a significant impact on improving instructional practice.

Yates (2007) analyzed survey responses of 395 elementary and secondary teachers, following their participation in professional learning opportunities between July, 2003 and April, 2004, to 21 statements about the principles of quality professional development. Results of the analysis supported the importance of planning for professional development that is longer in duration and focused on application of new content and skills in order to change teacher practice. Longer courses that incorporated elements of collegiality and applicability had a greater perceived influence on feelings of professional renewal (Yates, 2007).

A more recent study examined the impact of a multi-year professional development model targeting elementary school teachers and their development and implementation of high quality curriculum units for science instruction for English language learners (ELL) (Lee & Maerten-Rivera, 2012). Teachers initially attended a three-day workshop in the summer, followed by one-day workshops held during the school year and in subsequent summers, representing a total of 14 days of workshop training. The workshops focused on improving teachers' level of scientific knowledge

and understanding of effective strategies for English language development in the content area. Pre-and post-tests were used to measure the impact on ELL students in six elementary schools in a large urban school district in the southeastern portion of the United States, and almost 200 teachers participated in the project. A questionnaire was used to capture qualitative data from teachers about implementation of the new instructional strategies and approaches, and administrator observation data were also incorporated in the final analysis. Analysis of the data confirmed that multiyear professional development models generally result in changes in teacher practice after year one and typically stabilize over subsequent years. The researchers concluded that changes in teacher knowledge and instructional practice were greatest after year one of the professional development experience, but the improvements were not statistically significant after year one or over three years of the professional development intervention. Although the study's design included some critical features of high quality professional learning, e.g., a substantial degree of training exposure, it was limited by the failure to plan for a school leadership role in supporting implementation of the new practices, collegial support through peer coaching and collaboration, and a lack of randomization and attrition due to student and teacher mobility in the large urban district (Killion, 2014).

Horn and Little (2010) conducted a study analyzing professional conversations as a tool of professional learning to improve instructional practice. They found that conversations supported the importance of collegiality and applicability of content to improve teacher practice. They also determined that context, skill, and prior experience

in working collaboratively to resolve problems of practice influenced the outcomes for two separate teacher work groups in a large urban high school. Although one group of teachers had success in sharing student work samples and engaging in peer observation and productive feedback conversations, the other group stayed focused on curriculum development and planning and rarely addressed approaches to teaching the curriculum effectively. Professional learning communities with shared norms and a common focus on improving student performance "may be necessary but not sufficient for collaboration to yield opportunities for professional learning" (Horn & Little, 2010, p. 212).

Little (2006) conducted research on the role of school and community on the quality of professional learning experiences and their impact on teacher effectiveness and student learning. Historically, professional development for American teachers has typically occurred outside the school setting in formal, structured experiences that were short in duration and limited in impact on teacher practice or student learning (Borko, 2004; Darling-Hammond, 2010; Yoon et al., 2007). In contrast, Little (2006) concluded that schools are more effective in providing high quality learning experiences for all students when learning for all, teachers and students, is a priority. Four key goals were identified in Little's research: "making headway on the school's central goals, priorities, or problems; building the knowledge, skill, and disposition to teach to high standards; cultivating strong professional community conducive to learning and improvement; and sustaining teachers' commitment to teaching" (p. 2). Little also identified two reasons to support the argument that professional learning belongs in the school setting. First, the school site is where the work happens, and schools have unique needs related to levels of

staff training, quality, and experience as well as student demographics and community

priorities. Second, costs are high when schools fail to provide adequate learning

opportunities for students. Most impacted by these costs are the school administration,

the school faculty, the students, and the local community. Thus, Little (2006)

emphasized the importance of professional learning communities with shared values

including ownership of the responsibility for student learning and a collaborative focus

on continuous improvement of instructional practice and student outcomes. In learning-

centered schools, teacher learning as well as student learning is valued and respected, and

structures are established to promote and encourage productive professional dialogue.

The values of collaboration and school cultures that promote peer learning as a

tool for professional development are embedded in the body of research on high quality

professional learning models. Jackson and Bruegmann (2009) conducted a longitudinal

analysis of elementary school student performance in North Carolina using a value-added

approach. They concluded that students in teachers' classrooms with high performing

colleagues performed a standard deviation above their peers on state mathematics and

reading assessments. Data were collected from 1995 to 2006 for students in third through

fifth grades who had the same teacher for the entire school year. This resulted in 1.37

million data elements for analysis. Students in classrooms with less experienced teachers

had the most significant gains in performance, and the effect of peer quality was

consistent over the length of the study.

Louis & Marks (1998) also advocated for the professionalization of teaching.

They supported restructuring of school environments to provide time and purpose for

teachers to work together in productive groups with a collective focus on improving student learning. In their analysis, they collected data through surveys, classroom observations, authentic assessment tasks, student work samples, and case studies at eight elementary, eight middle, and eight high schools. They concluded that in schools with strong professional learning communities, social support for high levels of learning and student achievement results on authentic assessment tasks were both significantly higher than in schools lacking a professional culture of collaboration and a common commitment to improving student learning (Louis & Marks, 1998).

Similar results were found in a more recent study sponsored by Learning Forward's Center for Results. The study was conducted in 12 districts involved in the Bill & Melinda Gates Foundation's Supporting Effective Teaching Knowledge Development Initiative (Curtis, 2013). Teachers teaching teachers or supporting the learning of their peers through coaching and collaboration had positive impacts on school culture and student learning. However, other researchers have observed that effective peer learning models depend on key conditions; teacher leaders need support and direction from district and school leaders, a clear vision of what peer learning is supposed to accomplish that is widely communicated and understood, training in observation and feedback skills, and monitoring of impact on instructional practice and student learning (Curtis, 2013; Duff & Islas, 2013).

In summary, although members of the education profession in American schools have not historically been engaged in the meaningful, long-term professional learning that produces adaptations in teacher practice and improvement in student learning, recent

research and current trends provide direction for an improvement in the future. True

collegiality is characterized by teachers talking both formally and informally about

teaching and student learning, teachers sharing their knowledge about their craft, teachers

observing one another and providing helpful feedback to one another, and teachers

supporting one another in the effort to get better every day (Barth, 2006). "By building

permanent structures for collegial relationships, schools organize themselves for

improvement in multiple areas" (Joyce & Showers, 2002, p. 89). These principles, along

with the underlying framework of participatory leadership, provided the foundation for

the development and implementation of the BPCM, the professional learning model in

this study.

<div align="center">Peer Coaching as a Professional Learning Model</div>

Peer coaching, as an effective model for promoting improvement in teacher

practice, has a solid base in educational research (Bowman & McCormick, 2000; Hill &

Rapp, 2012; Johnson & Fiarman; 2012; Kohler et al., 1997; Leana, 2011; Murray et al.,

2008; Sparks & Bruder, 1987). However, the majority of studies conducted on peer

coaching have focused on the positive relationships, impact on school culture, and

increase in implementation of professional learning in teacher practice as a result of peer

coaching. Studies indicating significant impact on student achievement have been

limited (Joyce & Showers, 2002; Knight, 2007; Kretlow & Bartholomew, 2010).

Several common characteristics have been associated with effective peer coaching

professional learning models in the professional literature. Bartolo (2012) called for

patience, practice, and follow-up, with allowances for time to allow the peer coach to develop a comfort level with the process and the skills necessary to provide specific feedback that will improve instructional practice. Hall & Simeral (2008) emphasized the importance of positive relationships between effective peer coaches and teachers. These researchers found that to build trust, peer coaches recognized the individual strengths of the teachers with whom they were working, they asked questions to stimulate reflective practice, and they validated the insights and suggestions teachers expressed during the coaching process. Beyond building positive relationships, peer coaches also analyzed assessment results, collected additional evidence through peer observation, provided constructive feedback, collaborated on improvement goals, and suggested resources to support teacher growth (Hall & Simeral, 2008). Effective coaching was contingent upon strong questioning skills. Good peer coaches asked questions that prompted higher-level thinking and self-reflection; generated thoughtful, in-depth responses; and encouraged teachers to examine values and beliefs as well as past practice, current practice, and desired improvements in practice (Bearwald, 2011). Other key behaviors increasing the effectiveness of peer coaches included modeling, offering recommendations, affirming teacher efforts, and praising teachers when changes in practice occurred (Collet, 2012).

Several other key principles were identified consistently in the literature on effective peer coaching models (Demonte, 2013; Killion, Harrison, Bryan, & Clifton, 2014; Knight, 2011; Marzano & Simms, 2013; Tomlinson, 2011). First, equality was a significant concern. According to the researchers, effective peer coaches do not assume a position of power or authority when building an effective peer coaching partnership; they

work together with teachers as equal colleagues in the quest to improve practice. Second, teachers must have a choice and a voice in the improvement effort. The teacher establishes the improvement goals with the support of the peer coach, not under the direction of the peer coach or a supervisor. Third, teachers and peer coaches exchange candid, honest feedback in an atmosphere of mutual respect, trust, and humility. An open and honest dialogue was identified as a precursor to building effective peer coaching relationships. Fourth, practice was a key component; the goal of peer coaching is implementation of new and improved instructional strategies that positively impact student learning. This occurs when opportunities for modeling, practice, and feedback are built into the implementation plan, resulting in consistent application of new knowledge and skills in daily instruction. Finally, researchers agreed that the level of effectiveness of peer coaching is dependent on the level of respect the peer coaches hold as high quality teachers and the level of training they have received in observation and feedback skills. When these fundamental characteristics of effective peer coaching training and implementation were in place, teachers receiving support from a peer coach practiced new strategies more often, used their new strategies more appropriately and more consistently in the long-term. They discussed their new strategies with their students to identify the purpose of changing methodology. They also better understood how to share their adaptations in pedagogy with their colleagues to increase their influence as positive teacher leaders and ambassadors of effective practice (Joyce & Shower, 2002; Marzano & Simms, 2013).

Cross (2012) reported positive results on the implementation of a coaching model in 2005 in a small, semirural school district outside of Denver, Colorado. The school district was struggling with a 34% teacher turnover rate and low student achievement. A full-time achievement coach was selected for every school, and a half-time coach was dedicated at the district level to provide training and support for the coaches and school-based administrators to guide implementation and improve professional learning. The coaches first dedicated time to improving their own instructional skills, working in a professional learning community to engage in research, data analysis, communication, goal setting, and collaboration about how to develop trust with teachers in their schools and guide teacher learning to improve practice. Five years after full implementation of the peer coaching model in 2005, the teacher turnover rate had decreased by 21%.

> With each coaching interaction, teachers and coaches hone their skills and increase their appreciation of the power of and need for high-quality professional learning that makes a difference for students. Teachers . . . felt the satisfaction of increasing student achievement with new energy, worthy risks, strategies for thinking skills, and increased efficacy. (Cross, 2012, p. 41)

Another sustained implementation of a peer coaching model was conducted in nine Title I elementary schools by Gallimore, Ermeling, Saunders, and Goldenberg (2009). Grade-level teacher teams engaged in professional learning using a collaborative inquiry-based protocol to analyze problems of student learning, engage in common planning and peer observation with feedback, and use evidence from their collective work to change instructional practices. Data gathered from surveys and observations

indicated improvements in teacher practice and increases in student learning when teachers shared the same content, were guided by a trained peer coach, and operated in a stable school environment with administrator support for the initiative (Gallimore et al., 2009). In a similar study conducted in elementary schools in Ann Arbor, Michigan, Sparks and Bruder (1987) concluded that "An entire school, with proper support, can implement peer coaching that results in greater communication about and experimentation with teaching techniques" (p. 57).

Peer coaching appears to have a positive impact on collaborative practice and culture beyond the classroom teacher. Although direct teaching of students is a secondary responsibility of most school librarians, effective instruction is a job responsibility; however, librarians often work in isolation with limited formal training in the art of teaching (Sinkinson, 2011). A study conducted at the University of Colorado in Boulder in 2009 included eight elementary school librarians who volunteered to participate in a peer coaching professional development opportunity. After participating in an initial workshop introducing them to peer coaching and clarifying the role and expectations of the study, the librarians alternated functioned as inviting teachers and peer coaches. Inviting teachers allowed the peer coaches to observe their instruction to a class of students in the library, after a pre-observation meeting, followed by a post-observation meeting to provide opportunity for feedback and reflection; peer coaches were tasked with collecting data during observation of a fellow librarian teaching about a specific skill or strategy mutually identified prior to the observation and using probing questions to encourage reflection in the post-observation meeting. Each participant in the

study engaged in two observation cycles as a peer coach and two sessions working with a peer as an inviting teacher. The impact on student learning was not analyzed. Rather, participants were surveyed about feelings of comfort and confidence in the process. All participants in the study reported feeling comfortable in their roles as peer coaches, and 62% reported believed they had been successful in offering specific feedback that encouraged their peers' professional reflection. However, 25% of the participants expressed feeling uncomfortable in the role of inviting teacher, and they remained hesitant to invite others into observe their practice. Long-term impact on teaching or learning was not analyzed (Sinkinson, 2011).

The pattern demonstrated at the University of Colorado was consistent with that demonstrated ten years earlier by librarians in a California university. Two faculty members at California State University explored the discomfort that school librarians described in their roles as inviting teachers by initiating a peer observation process themselves (Eisenbach & Curry, 1999). The study participants were careful to distinguish between observation as a measurement tool for evaluative purposes and observation as a systematic process to stimulate reflection and professional growth through the collection of data and sharing of feedback. As a result of their peer coaching partnership, the university faculty members overcame their initial feelings of vulnerability, identified elements of their instructional practice to adapt and improve, and gained confidence and validation as a result of de-privatizing their practice (Eisenbach & Curry, 1999).

Peer coaching as a method to improve instructional practice and increase positive feelings about collaboration among teachers was also studied in exceptional student education classrooms by researchers at Texas A & M University (Hasbrouck & Christen, 1997). With the increase of inclusion of exceptional education students in regular education classrooms where teachers have limited training and experience in managing student behaviors and supporting individual learning needs, consulting teachers have been used successfully as peer coaches to strengthen teacher practice and improve learning opportunities for all children. Teachers participating in the study indicated they believed their teacher effectiveness had improved. They indicated that they wished to continue working with a peer coach and that they valued the increased opportunity for professional collaboration. No analysis of impact on student learning was reported (Hasbrouck & Christen, 1997).

Peer coaching has consistently been found to increase implementation of new strategies and skills and new curriculum concepts in classrooms of all levels and schools of varying sizes while adding to teachers' feelings of self-efficacy, collaboration, collegiality, and empowerment (Brown, 2010; Bruce & Ross, 2008; Collet, 2012; Cushman, 2013; Garet et al., 2008, Koch, 2014). Teachers in a K-12 private school who participated in peer coaching described a transformed professional school culture (Koch, 2014). Preschool teachers were more willing and successful in adopting a new pre-kindergarten mathematics curriculum when they experienced the support of a peer coach (Brown, 2010). Middle school teachers from nine schools in Tennessee, supported in their implementation of new standards and assessments by peer coaches, improved their

teaching practices in literacy and numeracy. They demonstrated higher levels of self-efficacy, increased their professional self-reflection, and produced a statistically significant difference in student achievement in mathematics (Cushman, 2013).

Adoption of a new reading program occurred more rapidly when teachers were supported by peer coaching, following their participation in a professional development seminar; however, there was no significant difference in student achievement between the teachers who attended the seminar without coaching and the teachers who were engaged in the implementation process with a peer coach following the training (Garet et al., 2008). Though Collet (2012) wrote "Teachers benefit when they are supported in the process of changing their practices" (p. 27), clear links to improved student achievement have not yet been established.

An important distinction between coaching and evaluation is another strand found in the professional literature relevant to peer coaching as an effective professional learning model. Evaluation plays a critical role in educational organizations. Standards are established and administrators are trained to observe performance, compare it to the established standards, and use the comparative data to make important human capital decisions about employment, salary, and professional advancement opportunities. In contrast, coaching is about continuous improvement, measuring teachers' performance against the established standards and guiding and supporting individual efforts to become more competent at designated tasks (Tschannen-Moran & Tschannen-Moran, 2011). Peer Assistance Review (PAR) programs are similar to peer coaching models but different in key ways. First established in Toledo, Ohio, in the early 1980s, PAR

programs were found in 41 American states at the time of the present study. These

programs have relied on trained teacher-mentors, much like trained peer coaches, to

observe colleagues and provide specific feedback to improve instructional practice.

However, PAR feedback has served a dual purpose. Although the information provided

to teachers has been intended to support professional growth and development, it has also

been used as evidence in summative evaluations used to make key human capital

decisions (Darling-Hammond, 2013). Darling-Hammond (2013) observed that PAR

programs support growth in teacher leadership, professional learning, and collaboration.

A cornerstone of effective peer coaching models, however, is trust between teachers and

coaches. The work of both peer coaches and peer evaluators depends on a professional

culture where practice is de-privatized and observations are seen as constructive

opportunities to learn and grow. Teachers from Hillsborough County Public Schools,

Florida, and other districts with PAR programs have expressed apprehension about

whether the role of the peer evaluator is to help or to judge (Marietta, 2011). BPCM was

developed and implemented with a concentrated focus on professional growth and

improvement through peer-to-peer collaboration, separate from the summative evaluation

process.

In summary, peer coaching as a professional learning model has been shown to

deepen acquisition and implementation of new knowledge, strategies, and skills in

teaching. School systems that have made the move from good to great and great to

excellent have concentrated efforts on establishing professional cultures that include

coaching, peer-led learning, and collaborative practice (Mourshed, Chijioke, & Barber,

2010). When coaches spend their time planning lessons with teachers, reviewing student work samples, observing teachers, and facilitating post-observation feedback sessions using effective questioning techniques, they become important forces for positive improvements in teaching and learning (Fullan & Knight, 2011). The present study sought to validate one school district's investment of resources in a peer coaching learning model built on this body of research.

<p style="text-align:center"><u>Measures of Teacher Effectiveness</u></p>

Although research on high quality adult learning to promote professional growth drove the development of the subject of this study, Brevard Public School's Peer Coaching Model (BPCM), the political process drove the implementation of new standards-based teacher evaluation systems in Florida in 2011-2012. Florida's participation in the federal Race to the Top grant program was contingent upon the development and implementation of research-based, multi-metric evaluation systems for educators including measurement of student growth and a meaningful observation and feedback process focused on continuous improvement (U. S. Department of Education, 2009). This requirement was signed into Florida law with Senate Bill 736, also known as the Student Success Act, mandating that all districts in Florida, even those opting out of participation in the state's Race to the Top program, would implement new state-approved evaluation systems in the 2011-2012 school year (Florida Government, 2011).

The Florida Department of Education (FDOE), tasked with interpretation and implementation of the new law, designated the Marzano and Danielson evaluation

<p style="text-align:center">44</p>

frameworks as the approved state models (FDOE, 2011) and provided training and support for district implementation with Race to the Top dollars. Districts had limited time to prepare for the obligatory shift in evaluation methods, processes, and procedures. Similar actions were taken across the United States as states and school districts scrambled to respond to the federal grant requirements. At the time of this study, new evaluation systems were the measuring sticks for teacher performance in Florida and in many states across the country.

Classroom observation has long been a tool used for evaluation of performance in schools. Strengths of the observation process include an action-orientation that is contextual. Thus, school and student characteristics are typically factors considered in the evaluation outcome. However, as noted in the New Teacher Project's Widget Effect Study (2009), principals have historically identified the majority of teachers as highly effective or effective, lacking the tools and the skills to more accurately distinguish levels of teacher performance using traditional observation and evaluation methods. Known limitations include (a) observer bias; (b) limited evidence of consistency in performance if the evaluative observation is a once-a-year, 20-minute episode; and (c) a lack of clear criteria, descriptors, and shared understanding of what represents effective instructional practice (Darling-Hammond, Wise, & Pease, 1983). The development and implementation of new observation tools and evaluation models aim for improved validity and reliability of the teacher evaluation process.

Observation Using the Danielson and Marzano Evaluation Models

Danielson first published her *Framework for Teaching* in 1996, but the second edition, published in 2007, has been the focus of implementation efforts in some districts in Florida and in other states (Danielson, 2007). Danielson (2007) has cited several purposes for use of the framework. These include providing a guide for novice teachers, supporting veteran teachers in reflection and collaboration about instructional practice, and providing a structure for both teachers and administrators to focus improvement efforts. The framework recognizes the complexity of teaching, the comprehensive nature of the job, and the work required before an effective lesson is delivered. The framework includes four domains: (a) planning and preparation, (b) the classroom environment, (c) instruction, and (d) professional responsibilities (Danielson, 2007, p. 22). The domains are further categorized into 22 components and 76 elements, thereby organizing teaching tasks into a structure that can be used by supervisors, coaches, or teachers to analyze instructional practice for the purposes of formative observation and feedback and summative evaluation.

Danielson's observation framework was evaluated for validity and reliability in the Met Project, *Gathering Effective Feedback* (2012). Researchers found a positive relationship between teacher observation scores and student achievement in reading and mathematics resulting from trained use of the Danielson frameworks. The researchers further found that reliability improved with short, frequent observations conducted by multiple observers. According to Griffin (2013), Danielson emphasized the role of school culture in effective implementation of her evaluation model:

Site administrators as instructional leaders must appreciate the role of school

culture, a professional culture, a culture of professional inquiry. They must define

teaching as not just what you do with your kids for six hours a day but also about

building a professional culture in which everybody is still learning. (Griffin,

2013, p. 29)

Griffin also observed that Danielson further highlighted the framework's intentional

effort to increase and improve professional conversation between principals and teachers

around the shared understanding of what represents effective teacher practice (Griffin,

2013.

Marzano, Frontier, and Livingston (2011), using the Marzano Art and Science of

Teaching Framework, stated the purpose of supervision very clearly: "The purpose of

supervision should be the enhancement of teachers' pedagogical skills, with the ultimate

goal of enhancing student achievement" (p. 2). Like Danielson, Marzano et al. (2011)

identified four domains in the Art and Science Teaching Framework: classroom

strategies and behavior; planning and preparing; reflecting on teaching; and collegiality

and professionalism (p. 4).

This framework is further delineated using a total of 60 elements, 41 of them in

Domain 1 related to classroom strategies and behaviors, organized into nine design

questions and three lesson segments: "routine events, addressing content, and enacted on

the spot" (Marzano Art and Science of Teaching Framework, 2011, p. 1). Unique to

Marzano's framework is the explanation of the relationship among the domains.

According to Marzano et al. (2011), the elements of Classroom Strategies and Behaviors,

47

Domain 1, have a direct impact on student achievement. The activities described in

Planning and Preparing, Domain 2, directly impact what happens in the classroom, and

the tasks in Reflecting on Teaching, Domain 3, directly impact the effectiveness of the

planning and preparation process. The activities described in Domain 4, Collegiality and

Professionalism, do not have a direct impact on student achievement but represent the

context of the teaching act (pp. 5-6). Dufour and Marzano (2009) emphasized a shift in

focus from supervision to capacity building for principals regarding formal teacher

evaluation, and they further emphasized the critical role of building collaborative teams

of teachers examining student learning and the impact of teacher practice on student

achievement, elements of teaching specifically addressed in Marzano's Domain 4.

Coulter (2013) conducted a small qualitative study in Washington State analyzing

teacher and administrator perspectives about the three state-approved teacher appraisal

systems (Danielson, Marzano, and CEL 5D+) in the state of Washington. Six teachers

and six principals from six school districts were interviewed, and responses of principals

and teachers were similar across the three instructional frameworks. Respondents found

the evaluation tools helpful in encouraging professional development in teachers but too

time-consuming to be used within the current school structures. Respondents also

questioned the ability of principals to fairly evaluate teachers using the new tools.

Sartain, Stoelinga, and Brown (2011) conducted a two-year study in Chicago on

implementation of the Danielson Framework for Teaching, concluding that though the

new evaluation system was an improvement over the previous model, there was a lack of

capacity for coaching and feedback and a lack of consistent support and buy-in from

school administrators. The researchers found that the classroom observation scores were predictive of student achievement and that principals and teachers were engaged in more collaborative conversations about improving instructional practice, but consistency across schools presented a challenge.

Maslow and Kelley (2012) conducted a study on the implementation of the Danielson evaluation framework in four Midwest high schools with large, diverse student populations. Through interviews with teachers and principals, the researchers concluded that the ability of school leaders to manage the safety of the school environment through effective student discipline procedures significantly impacted the implementation of the new evaluation tools. In two of the four schools studied, school leadership team members assigned management of student conduct a higher priority than teacher observation and feedback and reported often resorting to crisis management. As a result, the Danielson evaluation framework was not an effective tool for improving teacher or student performance. Time that must be dedicated to observation and feedback for effective use of the Danielson frameworks was instead spent on managing the safety and security of the school environment. In the other two schools, a collaborative professional culture was previously established that supported the dedication of time for consistent, systemic use of observation and evaluation to provide meaningful feedback to teachers with a focus on professional growth. The authors concluded:

Under the right conditions, evaluation can provide meaningful formative feedback to individual teachers and useful data to inform human resource management and school improvement. The 'right conditions' include an

underlying organizational culture focused on advancing student learning, a

shared belief that effective teaching can produce high levels of learning

for all students, a collaborative professional environment that promotes

conversations about teaching and learning, and a safe and orderly school

environment so that adults can focus on improving teaching and learning

rather than being overwhelmed by student safety and discipline issues

(Maslow & Kelly, 2012, p. 628).

A team from the RAND Corporation and the American Institutes for Research

(AIR), supported by the Bill & Melinda Gates Foundation, examined year one

implementation of the Danielson and Marzano frameworks in the following four sites:

Hillsborough County Public Schools in Florida, Memphis City Schools in Tennessee,

Pittsburgh Public Schools in Pennsylvania, and The College-Ready Promise charter

schools in California (Stecher, Garet, Holtzman, & Hamilton, 2012). The researchers

distributed surveys to 4,444 teachers (81% response) and 1,193 school leaders (76%

response) regarding their perceptions of the implementation of the new evaluation

systems. Overall, teachers and school administrators reported benefits from the change,

including a more clear understanding of what was being observed, a recognition that

observers were trained and qualified, and an acknowledgement that feedback from

observations was useful in identifying areas and strategies for improvement in teaching

practice. Teachers and administrators complimented the specificity of the observation

frameworks and the development of a common language about effective teaching.

However, concerns were raised by respondents from both groups of stakeholders and at all four sites about the amount of time required to engage fully in the new evaluation process, the inter-rater reliability and consistency across observers and schools, and the increased stress experienced as a result of the new system. There was less consistency in the survey results regarding the influence of the observation and evaluation process on professional development in year one of the project. Over 75% of school administrators reported a clear link between teacher evaluation and their opportunities for professional development. Responses from teachers varied from 25% to 45% across the four sites in the study regarding a clear relationship between their evaluation results and their professional learning opportunities. Both groups reported an increase in collegiality and collaboration among teachers and between school administrators and teachers, but the increased workload had generated increased pressure and tension.

Gallagher-Fishbaugh (2011), while President of the Utah Education Association, provided anecdotal data regarding the effective implementation of a teacher evaluation system focused on improving schools based on the research of Danielson and Marzano in the Granite School District in Utah. The school district defined improvement of every teacher as the project's primary goal. After a school observation, the author identified the following key elements contributing to the program's success and the improved relationships between teachers and administrators at the school and district level: (a) the training, time and support given to principals; (b) the reliability of the observation data; (c) the immediate feedback given to teachers; (d) a school culture accustomed to frequent

formal and informal observations; (e) district specialists to help improve teaching practice; and (f) the collaborative partnership between the teachers' association and district administration in facilitating the evaluation process.

Scanella & McCarthy (2014) examined implementation of a new evaluation system similar to those of Danielson and Marzano in Red Bank Borough Schools, New Jersey. The superintendent shared that the research-based model had added to the professional growth and development of everyone involved, resulting in an improved professional culture across the district, but not without paradigm shifts. Establishment of a strong evaluation advisory committee with stakeholders from all representative groups helped to develop clear pathways for communication and increased trust between teachers and administrators, and comprehensive training provided to both teachers and administrators created shared understanding of the instructional frameworks and the district's focus on building teacher capacity. Successful implementation required ongoing problem-solving, adaptations in resource allocation, and mutual accountability for the program's success.

Mielke & Frontier (2012) emphasized the importance of keeping professional growth and improvement in practice at the forefront in the implementation of new evaluation systems, basing their work on self-assessment with teams of teachers in districts using either the Marzano or Danielson models. The authors concluded that the frameworks, when implemented with fidelity, empower teachers, serving as tools for them to use in deliberate efforts to improve their instructional practice. The authors found that in effective implementation both administrators and teachers viewed the

frameworks as supportive of a formative evaluation process. Effective implementation also required school cultures that regarded a teacher's need for improvement as an opportunity for professional growth.

Brevard Public Schools teachers and administrators have been surveyed annually regarding implementation of the school district's evaluation system which is a hybrid model built on the research base of the Marzano (2007) and Danielson (2007) frameworks. Annual climate survey results related to implementation of the evaluation system from 2011-2012, 2012-2013, and 2013-2014, are reported in Appendix B. These surveys have been conducted in the spring of each year by Brevard Public Schools as a part of the strategic planning process. Overall, teachers and principals responded favorably to questions about more frequent and more specific professional conversations about the improvement of teaching and learning. Responses were mixed in regard to the functionality of the teacher appraisal system to consistently and reliably evaluate and improve teacher effectiveness. Teachers also reported increased professional dialogue with peers about the art of teaching but were unconvinced that these conversations would result in improvements in teaching and/or increases in student achievement (J. Respess, personal communication, May 16, 2014).

In summary, relatively new, research-based evaluation tools from Marzano (2007) and Danielson (2007) have provided teachers and instructional leaders with a common language to guide conversations about high quality teaching and learning in many school districts in the United States. Such interaction has been spurred by the mandates in the federal Race to the Top initiative (U. S. Department of Education, 2009) as well as

changes which followed in state legislation, e.g., Florida Statute 1012.34 (2012). These

evaluation systems were designed to have the dual purpose of (a) encouraging and

promoting improvement in teacher practice and (b) providing tools to consistently and

reliably distinguish between levels of performance in order to make sound human capital

decisions (Marzano, 2012). Hall (2013) noted positive results reported in the literature

thus far included increased professional dialogue about instructional practice among

teachers and between teachers and administrators and more clarity about what constitutes

effective teaching. However, implementation of the required changes has been

inconsistent and the challenge has been to determine a statistically significant impact on

teacher practice or student achievement. The present study was designed to examine the

impact of (a) a professional learning model, BPCM, on professional practice measured

using evaluation tools built on the Marzano (2007) and Danielson (2007) frameworks,

and (b) the impact, if any, on student performance.

The Value-added Approach

Although administrator observations have historically been the primary source of

teacher evaluation data, and data have been able to be improved by the standards-based

observation frameworks of Marzano (2007) and Danielson (2007), accountability for

student performance has continued to be a significant policy concern. With Race to the

Top (U. S. Department of Education, 2009) and subsequent changes in state legislation,

including Florida State Statute, 1012.34 (2012), student achievement now constitutes

50% of overall teacher and principal evaluation scores in Florida, as measured by the

state's value-added model (Florida Department of Education, 2011). The goal of a value-added formula is to distinguish the impact of non-school characteristics that affect student learning, e.g., family demographics, student factors, and peer influence, from school and teacher effects on student growth. As defined by the Institute of Education Sciences, "Value-added is the difference between expected and actual growth, used to measure changes in performance attributable to teacher, classroom, or school effects" (*Growth models*, 2012, p. 4). Florida's model was developed by a team of 27 teachers, school and district administrators, university professors, business community members, and parents, supported by experts from the American Institutes for Research (AIR). A covariate adjustment model accounting for student, classroom, and school characteristics was chosen and later approved by the Florida State Board of Education. Specific student characteristics included as factors in the model include:

- Up to two prior years of achievement scores (the strongest predictor of student growth)

- The number of subject-relevant courses in which the student is enrolled

- Students with Disabilities (SWD) status

- Gifted status

- Attendance

- Mobility (number of transitions)

- Difference from modal age in grade (as an indicator of retention)

(Florida Department of Education, 2011, p. 14).

A unique feature of Florida's model is its trajectory approach. Predictions about a student's projected performance include a factor designed to close the achievement gap within a three- to four-year period (*Growth models*, 2012). Other advantages of a value-added model, according to the Florida Department of Education, include the following:

- Teachers teach classes of students who enter with different levels of proficiency and possibly different student characteristics

- Value-added models 'level the playing field' by accounting for differences in the proficiency and characteristics of students assigned to teachers

- Value-added models are designed to mitigate the influence of differences among the entering classes so that schools and teachers do not have advantages or disadvantages simply as a result of the students who attend a school or are assigned to a class (Florida Department of Education, 2011, p. 8).

Proponents of value-added models in educator evaluation systems have defended the long-term impact of teacher effectiveness on a student's current and future performance and the historic failure of principals to accurately distinguish between levels of teacher performance using observation alone (Wesiberg et al., 2009). Chetty, Friedman, and Rockoff (2012) analyzed over 2.5 million standardized student test records along with adult outcomes including college acceptance and attendance, teenage birth rates, earnings, and retirement savings for the same former students. They found that students consistently assigned to teachers with higher value-added scores were more likely to attend and complete college, have higher salaries, and contribute more to

retirement savings. Students assigned to teachers with a standard-deviation increase in

value-added scores demonstrated on average three additional months of learning gains

(Chetty et al., 2012).

Glazerman et al. (2010) expressed the opinion that "Value-added evaluations are

as reliable as those used for high stakes decisions in many other fields" (p. 7). Though

not complete or perfect measures of teacher effectiveness, their correlation with student

performance has been stronger than relationships between teacher experience, teacher

certification, seniority, and observation scores and student learning gains (Glazerman et

al., 2010).

Concerns evident in the research about the use of value-added models in educator

evaluation have centered around variability of results across different measures, the

impact of non-school effects on student learning that may or may not be accounted for in

the value-added model, the lack of random assignment of students to teachers'

classrooms, and the role of standard measurement error in high stakes human capital

decisions (Baker et al., 2010; Baker, Oluwole, & Green, 2013; Harris, 2010; Hershbert &

Robertson-Kraft, 2010; Martineau, 2006; Rothman, 2010; Sparks, 2012). A teacher's

value-added score has been shown to fluctuate widely from year to year with between

10% and 20% of teachers in the top 20% of performance in one year being in the bottom

20% the following year (Rothman, 2012). Another criticism comes from the changing

nature of assessment. For value-added models to be reliable, there must be a direct

correlation between assessments and student data from year to year, yet with the advent

of common core, standardized tests across the national landscape have experienced

dramatic changes (Martineau, 2006; Sparks, 2012). Lipscob, Chiang, & Gill (2012) spoke to the issue of VAM value and fairness:

> The value of VAMS depends in significant part on the validity of the underlying student assessments in capturing what students ought to be learning and the capacity of the tests to allow VAMs to capture meaningful distinctions in achievement. . . a VAM's fairness depends on whether or not the method successfully removes influences outside the educator's control (p. iv).

Papay (2011) conducted a comprehensive study analyzing value-added results from over 25,000 Grades 3-5 student records and over 600 teachers in the Houston Independent School District and found significant variability in teacher value-added estimates across different measures. An examination of reading test results from the Stanford Achievement Test (SAT), the state standardized assessment, and the Scholastic Reading Inventory (SRI) illustrated differences in teacher value-added scores. Some of the scores would have resulted in a reduction in pay for performance of $2,000 or more for some teachers initially rated high on one assessment, based on the Houston Independent School District's pay for performance plan. Some teachers who scored in the lowest quartile on one measure would have earned bonuses if the pay for performance decisions were determined by results from one of the other measures.

Timing of test administration also played a role in the variation in value-added results. Some growth measures were calculated on a March test administration, and another set of growth measures was calculated on a May test administration. Finally, measuring student growth from one school year to the next neglected to consider the

impact of summer learning loss, known to be more significant for some students than others. Papay (2011) concluded that although there was a moderately high correlation between teacher effect and student growth, value-added measures would be most reliable for use in school and program evaluation and improvement, rather than for high-stakes human capital decisions such as teacher retention and pay for performance.

Kupermintz's 2003 review of the Tennessee Value-added Assessment System (TVAAS) raised similar concerns about the validity of using value-added estimates of teacher performance for high stakes human capital decisions. Although the business model of accountability for measuring school and teacher performance has strong appeal in the contemporary world, isolating student learning gains and attributing them directly to schools and teachers without taking into account the multitude of other variables that impact student growth raises questions of fairness and equity. TVAAS has been used to rank teachers within school systems, based on student performance, ignoring the variability in performance between systems. Teachers rated low in a high performing school system could be rated high in a lower performing system; the reverse was also true. In addition, the number of student results used to calculate a teacher's value-added score determined the overall validity. Kuppermintz found that measurement error played a significant role if a teacher had less than three years of student performance data to use in the calculation, and teachers in schools with a high rate of mobility and teachers assigned to smaller classes had a greater likelihood of having a score closer to the mean. Though a low performing teacher may potentially benefit from this regression, a high performing teacher's actual performance may not be evident in the final result. TVAAS

accounts for student aptitude, motivation, and demographic effects by using prior year test performance as a factor in the value-added formula. There have been, however, differentiated results in schools with low free and reduced lunch rates where the average gain was 103% compared to schools with high rates of students receiving free and reduced lunch where the average gain was 95%. Kuppermintz concluded that approximately one-third of the teachers deserving recognition for outstanding performance could be excluded due to factors known to impact student learning beyond the teacher's control.

Hill, Capitula, and Umland (2010) conducted a study comparing value-added measures for a sample of middle school mathematics teachers from a midsized district with a diverse student population in the southwestern United States to evidence gathered from standards-based formal observations, in-depth interviews, and surveys related to content knowledge and pedagogy. They found a moderate correlation between student growth measures of teacher performance and other measures of teacher quality. High performing teachers were rated consistently on both student growth measures and observations; however, seven of eight teachers rated in need of improvement on a formal observation tool had relatively high value-added results. An examination beyond the data found that students appeared to have benefited from participation in after-school tutoring and in-school test preparation sessions. Special education teachers were consistently rated low in value-added results, and teachers in schools with more affluent student populations or those working with gifted students earned higher average value-added results, even when prior year test performance and other student characteristics were

included in the value-added formula. Though value-added scores were determined

insufficient as a single measure to improve teacher quality or identify teachers for either

performance pay or removal, they were considered to be a valuable tool when used in

combination with a high quality observational tool (Hill et al., 2010).

Other researchers have found weak to moderate correlation between administrator

evaluation of teachers using classroom observation tools and value-added results.

Analysis of results from Brevard Public Schools teacher evaluation results in 2011-2012,

the first year of implementation of a teacher appraisal system, indicated a weak

correlation coefficient of .209 (Mela, 2013). A similar analysis of teacher evaluation

results from a larger urban school district in central Florida yielded a moderate

correlation coefficient of .50 for middle school reading teachers and .47 for high school

reading teachers (Fritz, 2014). In an older study, which combined both quantitative

analysis of observation and value-added scores with confidential principal interviews

regarding administrator impression of teacher effectiveness from 30 schools in a mid-

sized Florida school district, weak correlation coefficients of .276 for mathematics and

.168 for reading were found when overall principal evaluation and value-added measures

were compared for each teacher (Harris, Ingle, & Rutledge, 2014). The authors found

that personal characteristics of teachers, including strong communication skills, ability to

work with others, and caring, played a significant role in how principals judged teacher

performance. Kimball and Milanowski (2009) also found considerable variation between

districts and schools in the relationship between observation ratings and value-added

results. They questioned the reliability of using evaluation scores and student

achievement results to make data-driven decisions about professional development or human capital management. Finally, it was determined in an analysis of the use of value-added results in teacher effectiveness rankings for teachers in the Los Angeles Unified School District sponsored by the National Education Policy Center, and originally published in the *Los Angeles Times*, that there were significant issues with false positives and negatives in scores of between 14% and 22% of the teachers included in the study (Briggs & Domingue, 2011). However, though significant questions were raised about the accuracy of the value-added ratings for individual teachers, it was concluded that the formula used in Los Angeles provided useful information for district and school analysis and represented an improvement over evaluation systems that relied on observation only without some incorporation of student achievement outcomes in measuring teacher performance.

In summary, evaluation of teacher performance is a complex task and an important one in the eyes of the public. There are issues to consider when incorporating value-added results in teacher, school, and district evaluation, but to ignore student outcomes neglects critical elements of educator effectiveness. To mitigate the potential misuse of value-added measures, researchers have recommended a balanced approach to educator evaluation, using multiple sources of data to evaluate teacher effectiveness. Valid and reliable value-added models use two to three years of student performance data, use all available student data, including students with incomplete records, consider the effects from multiple teachers into an individual student's learning, and use assessments that are aligned to the standards. In addition, evaluation systems should

clearly specify standards for effective teaching, use valid and reliable tools, and provide

sufficient training for teachers and evaluators to understand and use the system with

fidelity (Hershbert & Roberston-Kraft, 2010).

Summary

Ideally, professional learning and teacher evaluation are closely aligned with the

shared purpose of improving instructional practice to increase student achievement.

Much professional development in schools has not been guided by the research

advocating the importance of a longer duration with collaboration and support for

implementation built into the design, and teacher evaluation has been hampered by

observations characterized by non-specific feedback and limited distinction between

levels of performance. Changes in recent years show promise for improvements in the

design of professional learning. These changes include implementation of standards-

based evaluation systems and observation frameworks that clearly define effective

teaching as well as the measurement of student achievement using value-added models

that incorporate factors beyond the control of teachers and schools. The present study

was conducted to evaluate the impact of a professional learning model built on the

current research, BPCM, on professional practice and student growth. The findings of

prior researchers described in this review of the literature were used to guide the

investigation.

CHAPTER 3
METHODOLOGY

Introduction

The Florida Department of Education's Professional Development System

Evaluation Protocol is based on a system of 65 standards organized around four strands

based on Learning Forward's publication, *Standards for Staff Learning* (2011*)*.

Bergquist (2012) identified evaluation as the weakest in implementation of the protocol's

four strands in the third cycle of the protocol's implementation in 2011-2012. Specific

standards in the protocol call for evaluation of professional development activities at the

educator, school, and district levels based on subsequent changes in educator practice and

student outcomes (Standards 1.4.2, 1.4.3, 2.4.2, 2.4.3, 3.4.2, and 3.4.3). Specific results

for Brevard Public Schools in the areas of "implementing" and "follow up" in 2011-2012

earned a mean score 0.5 points lower than scores in the areas of "planning" and

"learning" (Bergquist, 2011). These results informed this study's purpose and

designated outcome measures. This chapter outlines the methodology used in the study.

It includes descriptions of the research design, the participants, instrumentation, and data

collection and analyses procedures used to address the research questions and test the

study's hypotheses.

Purpose of the Study

The purpose of this study was to analyze whether a school district's significant

investment of time and resources in a two-year peer coaching professional learning model

resulted in improved instructional practice and increased levels of student achievement as measured by Florida's value-added model [VAM] (2013).

Research Design

A quasi-experimental design was chosen for this study in order to determine if there were statistically significant differences in the educator evaluation scores and the student achievement results between the study's sample, those teachers from Brevard Public Schools who participated consistently in the peer coaching training (BPCM); and the population, all teachers from Brevard Public Schools who earned professional practices and VAM scores during the course of the study but did not participate in BPCM. In addition, the study was conducted to determine if a significant difference existed in the evaluation scores and student achievement results of BPCM teachers prior to the training, at the end of year one of the training, and at the end of year two of the training. The focus was on comparing (a) 2011-2012 results, prior to BPCM training and in the first year of the evaluation, with (b) 2012-2013 results, after the first year of BPCM training and in the second year of the evaluation system's implementation, and (c) 2013-2014 results, after two years of BPCM training and in the third year of the evaluation system's implementation.

Population

The participants for BPCM were selected by their principals to attend the initial training developed and provided by the BPS Office of Leadership and Learning in the

summer of 2012. Candidates were required to have highly effective annual evaluations in the 2011-2012 school year and to have completed the state-approved Clinical Educator Training (CET) in advance. Candidates were also required to complete the full three days of initial BPCM training to be considered a peer coach at their schools. The original training group included 285 participants who had an average experience level of 13.75 years. Two days of follow-up training were provided, one day each semester, during 2012-2013. Two additional days of training were offered in the summer of 2013, and two more days of follow-up training continued in the 2013-2014 school year. To remain in the study's sample for data collection and analysis, teachers must have attended a minimum of six of the nine days of BPCM training offered in 2012-2014. As a result of subject mortality and the researcher's assumption that peer coaches would use the full two-year training regimen in order to demonstrate changes in practice and improved student outcomes, the study sample size was reduced to 174 participants. Although reliance on a convenience sample limited the study's generalizability, the study's sample was very similar in demographics and experience to the 2011-2012 population of teachers in Brevard Public Schools. The study sample contained a larger percentage of elementary school teachers (70%) than represented in the population (57%), reflecting the school district's structure of smaller, more numerous elementary schools compared to secondary middle and high schools. Every school participating in the study was invited to send three participants to the initial BPCM training, producing the skew in teacher participation by school level. Table 2 contains the complete demographics for the study sample and the population.

Table 2

Comparison of Study's Sample and Population Demographics

	Sample		Population	
Descriptors	*f*	*%*	*f*	*%*
Size	174	4.8	3,435	95.2
Gender				
Female	156	90.0	2,851	83.0
Male	18	10.0	584	17.0
Race				
White	161	92.0	3,091	90.0
Black	6	4.0	171	4.9
Hispanic	7	4.0	173	5.1
Average Years of Experience		13.3		11.2
Level				
Elementary	121	70.0	1,958	57.0
Middle	24	14.0	515	15.0
High	29	16.0	962	28.0

Method of Data Collection

The 2011-2012 school year represented the initial year of a new evaluation system in Brevard Public Schools, as required by FL Senate Bill 736, known as the Student Success Act (2011), and Florida's participation in the nationally funded Race to the Top initiative (U. S. Department of Education, 2010). As a result of these measures, teacher evaluation was required to be multi-metric, including both evaluation of professional practice by trained administrators based on the Florida Educator Accomplished Practices (FEAPs) and student achievement outcomes on Florida's standardized assessment, the

Florida Comprehensive Assessment Test (FCAT), as measured by the state's value-added model (VAM).

Approval to conduct the study was sought from the Institutional Review Board at the University of Central Florida (Appendix C). Approval for the researcher's collection of summary data for the study's population and individual data for the study's sample was obtained from the Brevard Public Schools Office of Accountability, Testing and Evaluation (Appendix D). Quantitative data representing evaluation scores and VAM results, as well as demographic data for the population and members of the convenience sample for 2011-2012, 2012-13, and 2013-2014, were obtained with cooperation from the school district's Division of Human Resources. These data were compiled in Microsoft Excel and imported for analysis into the software program, the Statistical Package for Social Sciences (SPSS).

The professional practices evaluation data was collected from administrators and reviewed with teachers using the BPS Instructional Personnel Evaluation Instrument (Appendix E). These data were stored in an electronic database for school district analysis and required state reporting. Aggregated VAM scores for teachers and schools were reported annually by the Florida Department of Education and incorporated into the school district database for similar analysis and reporting.

Instrumentation

Brevard Public Schools developed and implemented its Instructional Personnel Performance Appraisal System (IPPAS) in 2011-2012 (Brevard Public Schools, 2011).

A primary component of the system included the evaluation of an educator's professional practices by a trained administrator, across seven dimensions of the Florida Educators Accomplished Practices (2010) based on the extensive research on effective classroom practice of Danielson (2006, 2007), Marzano (2007), Marzano, Marzano, & Pickering (2003), and Marzano, Pickering, and Pollock (2001). The seven dimensions are identified in Appendix F, and rubrics further defining each dimension are found in Appendix G. Levels of performance for each element are defined across the following 4-point Likert-type scale:

A. **Distinguished (3 points):** Indicates performance that consistently exceeds the requirements of the position and the level of performance commensurate with the experience of the teacher.

B. **Proficient (2 points):** Indicates performance that consistently meets the requirements of the position and the level of performance commensurate with the experience of the teacher.

C. **Developing or Professional Support Needed (1 point):** Indicates performance that requires additional attention to ensure an accepted level of proficiency. Further, this performance is not characteristic of the requirements for the position and experience of the teacher.

D. **Unsatisfactory (0 points):** Indicates performance that does not meet the minimum requirements of the position and the level of performance commensurate with the experience of the teacher (Brevard Public Schools, 2011, p. 2).

Administrators collect formative observation data and other pieces of evidence of a teacher's level of performance across the seven dimensions through classroom walk-throughs, informal observations, formal observations, professional learning community participation, peer observations, parent conferences, and other professional activities. The evaluation process concludes with an administrator's summative evaluation using the BPS Instructional Personnel Evaluation Instrument (Appendix E). Points are assigned for each element in a dimension, summed, and averaged to achieve a mean score for each dimension. Mean scores for each dimension are then added to represent a teacher's total Evaluation of Professional Practices score with a potential value of up to 21 points in the summative evaluation rating.

Because 2011-2012 represented the initial year of implementation of the new evaluation system, and no field-testing was conducted in advance of implementation, validity of the instrument has yet to be fully determined. The validity is dependent upon the research base. A study following the first year of implementation of IPPAS revealed a weak correlation (.231) between the Evaluation of Professional Practices total scores and VAM scores for the population of BPS teachers in Grades 4-10 teaching FCAT-assessed content (Mela, 2013).

Consistency across administrators in use of the study's instrumentation, or inter-rater reliability, in the evaluation of professional practices is a second key consideration. Appraisal systems developed with specific guidelines and clear definitions improve inter-rater reliability, as does extensive training for raters (Holland, 2011; Thornton, 2012). The rubric design of IPPAS was intended to provide clear guidelines and definitions of

distinguished levels of educator performance, and additional actions were taken by the BPS Office of Leadership and Learning under the leadership of Dr. Jane Respess to continuously improve the skills of administrators to accurately and consistently evaluate teacher practice. Specifically, according to Respess, multiple trainings included: (a) an initial overview of the appraisal process for both administrators and teachers in the summer of 2011; (b) observation training, scoring practice, and observer certification in the fall of 2011 for all principals and assistant principals; (c) ongoing training for teachers in analyzing the observation rubrics; and (d) recertification for all principals and assistant principals using classroom videos to practice and certify reliable scoring in the summer of 2013. New administrators have been provided training and must pass a certification test (scoring of a classroom video) with a 75% level of accuracy before they are permitted to observe and evaluate teacher practice. Ongoing inter-rater reliability training using calibrated classroom videos at monthly leadership team meetings has been provided for all observers. During the 2013-2014 school year, instructional rounds for principals and assistant principals, guided by trained lead observers, were added to the training regimen to further strengthen inter-rater reliability in the observation and evaluation process (J. Respess, personal communication, March 24, 2014).

A second critical component of teacher evaluation for the purposes of this study consisted of student achievement outcomes, as measured by Florida's Value-added Growth Measure (VAM). The state-approved teacher evaluation system for educators from Brevard Public Schools indicated the school district's intention to fully comply with both Race to the Top requirements and the legal mandates in the Student Success Act

(2011), specifically s.1012.34(3), F.S., which required school district plans to utilize "indicators of student learning growth assessed annually and measured by statewide assessment" (para.7) in teacher evaluation results. Rather than relying on a strict proficiency-based model, which obviously held potential disadvantages for schools with more challenging student populations, Florida appointed a Student Growth Implementation Committee (SGIC) in 2010, comprised of 27 teachers, parents, business members, and school and district administrators. This group was assigned to work with the American Institutes for Research (AIR) to develop and guide implementation of a valid and reliable value-added formula for use in Florida to measure the impact of an individual teacher on student learning (Florida Department of Education, 2011b).

The model recommended by the SGIC and later approved by Florida's Commissioner of Education was a covariate model that incorporates weighted criteria of variables impacting student performance. These include: two years of student performance on the state's standardized assessment test (FCAT), student attendance, student mobility, class size, and student status as gifted, student with disability, or English language learner (*Florida's value-added model*, 2013). The Institute of Education Sciences (IES) Statewide Longitudinal Data Systems [SLDS] (July 2012) examined Florida's VAM formula and commented on its trajectory approach. In this approach, the model not only recognizes and rewards growth in student performance above a predicted score, but it particularly rewards growth that is on target to close the achievement gap in a three to four year period (Growth Models, 2012). Using Florida's covariate approach that considers characteristics of the student, the classroom, and the

school (*Florida's value-added model*, 2013), a positive VAM score represents aggregate student performance above what was predicted; A negative VAM score represents aggregate student performance lower than what was predicted,

According to an analyst in the BPS Office of Accountability, Testing, and Evaluation, the department carefully analyzed 2010-2011 VAM data received from the U. S. Department of Education using frequency distribution to develop fair and reliable conversion tables in order to assign points for levels of student performance in the teacher evaluation system (J. Carr, personal communication, March 24, 2014). However, for the purposes of this study, the researcher relied on the three-year aggregated VAM (mathematics and reading) scores received directly from the Florida Department of Education, validated by the American Institutes for Research, to calculate differences in results over the length of the study. Three types of VAM scores were considered: TAV, SAV, and NFTAV. Teachers of students in Grades 4-10 assigned to reading/language arts and/or mathematics earned the aggregated VAM scores for their assigned students (TAV). Teachers of students in Grades K-3 and 11-12 earned the aggregated VAM scores for all students in their assigned schools (SAV). Teachers of students in Grades 4-10 assigned to content other than reading/language arts and/or mathematics earned the aggregated VAM scores for their assigned students (NFTAV). Teachers with less than two years of student data also earned the school aggregated VAM (SAV) to limit the potential impact of standard error in the calculation. Table 3 reflects the distribution of VAM types in both the sample and the population.

Table 3

Distribution of Value-added Model (VAM) Types in Study's Sample and Population

Value-added Scores	Sample	Population
Teacher (TAV)	46.6%	28.5%
School (SAV)	37.9%	36.3%
Non-FCAT Teacher (NFTAV)	15.5%	35.2%

Methods of Data Analysis

This study was guided by an interest in determining if a sustained professional development model, specifically peer coaching, improved instructional practice and student performance in teacher-participants from Brevard Public Schools. The following research questions governed the selection of statistical measures and analytical procedures to be used:

1. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from the VAM scores of the population of Brevard Public School teachers who did not participate in BPCM?

   H_{01} There is no statistically significant difference in the VAM scores of those teachers who participated in peer coaching training (BPCM) from the VAM scores of the population of Brevard Public School teachers who did not participate in BPCM.

2. To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from the professional

practices scores of the population of Brevard Public School teachers who did not participate in BPCM?

H_{02}: There is no statistically significant difference in the professional practices scores of those teachers who participated in peer coaching training (BPCM) from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM.

3. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

H_{03}: There is no statistically significant difference in the VAM scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

4. To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

H_{04}: There is no statistically significant difference in the professional practices scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

Quantitative data for professional practices and student achievement (VAM)

scores were collected, analyzed, and reported for the descriptive statistical measures of

range, mean, standard deviation, skewness, and kurtosis for each year of the study (prior

to training-2011-12, after year one of training-2012-2013, and after year two of training-

2013-2014). The mean and standard deviation of the professional practices and VAM

scores for the sample and the population were compared using an independent samples t

test to determine whether or not there was a significant difference in the results and to

respond to Research Questions 1 and 2. Although use of a convenience sample limited

the generalizability of this study, the sample of 174 participants met the required size for

use of the independent samples t test to draw reasonable conclusions about significance

when the standard deviation for the population is a known factor (Steinberg, 2011).

Mean results and the standard deviation for the sample were compared to determine if a

statistically significant difference existed using a paired samples t-test and to respond to

Research Questions 3 and 4. The researcher first evaluated differences between baseline

and year one scores, but the primary analysis focused on whether or not there was a

significant difference between baseline and year two results (Lomax, 2007). The

statistical software package, SPSS, was used to perform all of the statistical analyses

associated with the four research questions.

<u>Summary</u>

This chapter further solidified the purpose and methodology used for this-

comparative study, including a restatement of the research questions and null hypotheses.

76

A discussion of the research design and a description of the study's sample and population were presented. Methods used in quantitative data collection were explained, and instrumentation used to collect data for the study was identified. A rationale for the methods of data analysis chosen to report results, draw conclusions, and make recommendations for future study were also provided. The following chapter is dedicated to the reporting of the study's results.

Introduction

The purpose of this study was to analyze whether one public school district's

investment in professional learning for peer coaches, Brevard's Peer Coaching Model

(BPCM), resulted in improved instructional practice for the participants and their peers

and increased levels of student achievement as measured by Florida's value-added model,

VAM. Three years of data were analyzed as follows: (a) professional practices

evaluation results as measured by Brevard Public Schools' Instructional Personnel

Performance Appraisal System (IPPAS); and (b) aggregated value-added scores as

determined using Florida's value-added model, using appropriate statistical measures

based on the number of study participants and distribution of the data. Independent

samples *t* tests and the paired samples *t* tests were used to compare scores of BPS

teachers in the study's sample and population. To minimize bias, the study's sample and

population were delimited to teachers in Brevard Public Schools with three years of

professional practices and VAM data (2011-2012, 2012-2013, and 2013-2014). This

chapter provides descriptive statistics for the study's sample and population followed by

data analysis results for the four research questions which guided this study.

Descriptive Statistics for the Sample and Population

During the 2011-2012 school year, the baseline year for this three-year study,

there were 84 schools operated by the Brevard Public School District which included 56

elementary schools and 28 middle/high schools. Charter schools, adult education centers, alternative learning schools, and virtual schools were excluded from the study. The study's sample and population were delimited to teachers working in one of the 84 schools who earned a professional practices observation score and a value-added score during each academic year of the project (2011-2012, 2012-2013, 2013-2014). The following criteria further delimited the study's sample: (a) teachers earned a highly effective evaluation in the previous year, (b) teachers had previously completed Clinical Educator Training (CET), and (c) teachers completed a minimum of six of the nine days of BPCM training offered by school district personnel. Teachers who did not meet all three criteria were excluded from the study. The study's comparison population was delimited to teachers from Brevard Public Schools with three years of evaluation data, professional practices observation scores and student achievement data as measured by Florida's aggregated value-added model for each year of the project (2011-2012, 2012-2013, 2013-2014).

The initial number of teachers who participated in one or more days of BPCM training was 285; however, over the course of the study, the sample narrowed to 174 teachers. The initial population of Brevard Public School teachers who did not participate in BPCM training was 4,735; similarly, the population narrowed to 3,435 teachers over the course of the study due to the identified delimitations. The majority of teachers in the sample and the population were elementary (n = 123, 74%, and n= 1,958, 57%, respectively), reflecting the larger number of elementary schools in the BPS community. Both the sample and population teachers were primarily female (n = 156,

90% and n = 2,851, 83%), White (n = 161, 92%, and n= 3,091, 90%), had comparable

years of experience (m =13.25 for teachers in the sample, m = 11.15 for teachers in the

population).

A majority of teachers (46.6%) in the study's sample earned a Teacher-

Aggregated VAM score (TAV), 37.9% earned a School Value-added score (SAV) and

15.5% earned a Non-FCAT Teacher Value-added score (NFTAV). VAM results for the

population were more evenly distributed across TAV (28.5%), SAV (36.3%), and

NFTAV (35.2%) scores. For both the sample and the population, the teacher-aggregated

VAM score (TAV) is the value-added score earned by teachers assigned to teach

reading/language arts and/or mathematics to students in Grades 4-10. The non-FCAT

teacher value-added score (NFTAV) is the value-added score earned by teachers assigned

to students in Grades 4-10 and to content areas other than reading/language arts and/or

mathematics. The school-aggregated VAM score (SAV) is the value-added scored

earned by teachers assigned to students in Grades K-3 or 11-12 and other instructional

personnel assigned to non-classroom roles. All instructional personnel in Brevard Public

Schools earn a value-added score that represents student performance in reading/language

arts and/or mathematics aggregated over a three-year period to reduce the impact of

measurement error. Figure 1 illustrates the VAM types for the sample and the
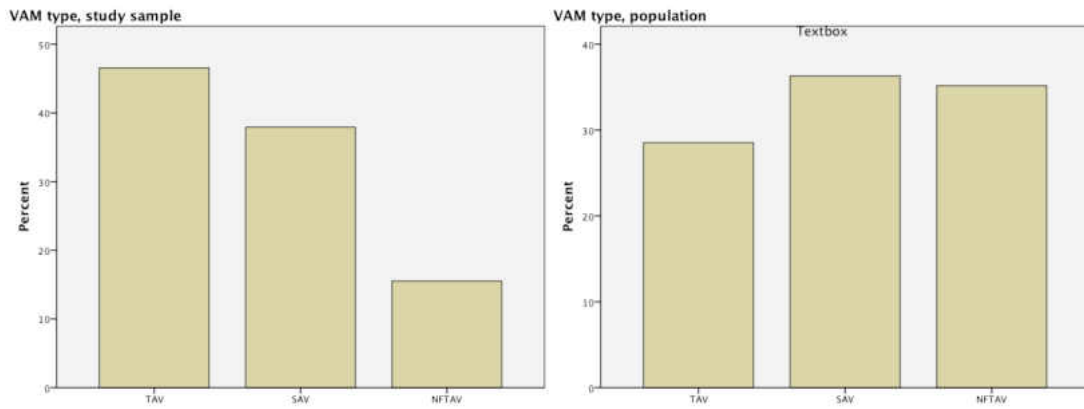
population.

Figure 1. Comparison of Study Sample and Population: Value-added Model Types

Presentation of Data Analysis

The presentation of the analysis of data has been organized around the four

research questions which guided the study. Descriptive statistics are presented in tabular

form and discussed.

Research Question and Hypothesis 1

To what extent do the VAM scores of those teachers who participated in peer

coaching training (BPCM) differ from the VAM scores of the population of

Brevard Public School teachers who did not participate in BPCM?

H_{01} There is no statistically significant difference in the VAM scores of those

teachers who participated in peer coaching training (BPCM) from the VAM

scores of the population of Brevard Public School teachers who did not participate

in BPCM.

Value-Added Scores (VAM) represented the difference between students'

expected achievement on a standardized assessment and the actual performance of the

students who shared similar characteristics. When aggregated, the value-added model

represented a teacher's contribution to the students' learning. For BPS teachers, VAM

equated to 35 points in the final summative evaluation in the baseline and year one of the

study (2011-2012, 2012-2013), and 45 points in year two (2013-2014) of the study. For

the purposes of the study, the raw VAM scores, rather than the scaled scores, were used

for more precise analysis and comparison. BPS teachers in both the sample and

population of this study earned a TAV (teacher-aggregated VAM), a SAV (school-

aggregated VAM), or a NFTAV (a non-FCAT teacher aggregated VAM), based on three

years of reading and mathematics aggregated data from Florida's state assessment test,

FCAT, for each year in the study.

From the initial year of this study, 2011-2012, prior to BCPM, through the final

year of the study, 2013-2014, the value-added mean scores improved for both the sample

and the population. The range of value-added scores also narrowed for both teacher

groups. For the 2013-2014 school year, year two of the BPCM implementation and year

three of the study, the sample's VAM scores ranged from -.67 to 1.13, had a mean of .11

(SD = .24), and a variance of .06. The population's VAM scores ranged from -1.17 to

1.40 and had a lower mean of .07 (SD = .23) and a variance of .05.

Tables 4 and 5 contain the complete descriptive statistics for the value-added scores of

the sample and population.

Table 4

Value-added Scores for Sample

Year	Range	Mean Statistic	Standard Deviation	Skewness	Kurtosis
2011-2012	5.58	.08	.41	-3.93	48.54
2012-2013	5.61	.08	.43	-3.30	38.15
2013-2014	1.80	.11	.24	1.08	4.71

Table 5

Value-added Scores for Population

Year	Range	Mean Statistic	Standard Deviation	Skewness	Kurtosis
2011-2012	3.01	.06	.22	.21	7.73
2012-2013	4.35	.07	.28	.07	6.13
2013-2014	2.56	.07	.23	.65	3.31

An independent samples *t* test was conducted on the value-added scores to evaluate whether the mean score of the sample, BPCM participants, was significantly different than the mean score of the population, non-BPCM teachers in Brevard Public Schools with three years of aggregated student data. Because the size of the sample and the population differed, equal variances were not assumed in the analysis of the results.

The 2013-2014 sample's mean value-added score ($M = .11$,, $SD = .24$) was not significantly greater than the mean value-added score for the population ($M = .07$, $SD = .24$), $t(190.075) = -1.669$, $p = .097$, at a 95% confidence level. That is to say, though

Brevard Public School teachers who participated in the BPCM professional learning

series had higher levels of impact on student learning than teachers who did not

participate in BPCM, as reflected in the average value-added scores, the impact on

student learning was not statistically significant. Table 6 contains complete results of the

independent samples *t* test.

Table 6

Independent Samples t-Test: Overall Value-added Results of BPCM and non-BPCM
Participants

Descriptor	Mean	Standard Deviation	*t*	df	Sig. (2-tailed)
Non-BPCM	.07	.23			
BPCM	.11	.24			
Equal variances not assumed			-1.669	190.075	.097

Note. BPCM = Brevard Peer Coaching Model

Further examinations of the data were conducted using independent *t* tests on the

2013-2014 value-added scores, separated by VAM type (TAV, SAV, and NFTAV). The

difference in results between BPCM and non-BPCM participants using the teacher-

aggregated VAM formula (TAV) remained insignificant ($t(93.679) = -.447$, $p = .097$, at a

95% confidence level. The same held true for the NFTAV participants in the study

($t(27.362) = -.046$, $p = .964$, at a 95% confidence level). In contrast, the difference

between school-aggregated VAM scores between teachers in the sample with BPCM

training and teachers in the population without BPCM training was significant (t(84.340) = -4.125, p = .000, at a 95% confidence level). However, though the mean VAM score for BPCM teachers was higher than the mean VAM score for non-BPCM teachers in all three categories, the difference was insignificant except for teachers earning school-aggregated VAM scores (SAV). Because analysis of scores for two of the three VAM types produced less than statistically significant results, and the difference between mean scores of the population and sample for all VAM types was also not statistically significant, the null hypothesis was accepted. Tables 7, 8, and 9 present the complete results of the analysis.

Table 7

Independent Samples t-Test: Value-added Results for BPCM and non-BPCM Teacher Value-added Scores (TAV)

Descriptor	Mean	Standard. Deviation	t	df	Sig. (2-tailed)
Non BPCM	.08	.32			
BPCM	.10	.32			
Equal variances not assumed			-.447	93.679	.097

Note. BPCM = Brevard Peer Coaching Model

Table 8

Independent Samples t-Test: Value-added Results for BPCM and non-BPCM Non-FCAT Teacher Value-added Scores (NFTAV)

Descriptor	Mean	Standard Deviation	t	df	Sig. (2-tailed)
Not BPCM	.07	.22			
BPCM	.07	.20			
Equal variances not assumed			-.046	27.362	.964

Note. BPCM = Brevard Peer Coaching Model

Table 9

Independent Samples t-Test: Value-added Results for BPCM and non-BPCM School Value-added Scores (SAV)

Descriptor	Mean	Standard Deviation	t	df	Sig. (2-tailed)
Non BPCM	.08	.15			
BPCM	.13	.09			
Equal variances not assumed			-4.125	84.349	.000

Note. BPCM = Brevard Peer Coaching Model

Further examinations of the data were conducted, removing outlier scores, to confirm appropriate use of the independent samples *t*-test with slightly skewed data and size differences between the sample and the population. Three scores were removed from the BPCM sample 2013-2014 data set, one score less than -.5 and two scores greater

than 1.01. A total of 13 scores were removed from the non-BPCM population's 2013-2014 data set, using the same parameters of < -0.5 and > 1.01. The difference in mean VAM scores between BPCM and Not BPCM teachers remained insignificant ($t(190.32)$ = 1.660, p = .099, at the 95% confidence interval). Because this was a client-based study and no significant difference was observed in the results when comparing all scores to scores with outliers removed, no VAM scores were removed from further analysis in the study.

Research Question and Hypothesis 2

To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM?

H_{02}: There is no statistically significant difference in the professional practices scores of those teachers who participated in peer coaching training (BPCM) from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM.

The professional practices scores of this study's sample and population represented the summative rating assigned to teachers by a supervising administrator (principals, assistant principal, or director). The IPPAS rubrics were used to differentiate Distinguished, Proficient, Professional Support Needed, and Unsatisfactory levels of performance in formal and informal classroom observations. The professional practices

87

score reflected a teacher's level of performance on seven instructional dimensions (Instructional Design and Lesson Planning, Learning Environment, Instructional Delivery and Facilitation, Assessment, Professional Responsibilities and Ethical Conduct, Relationships with Students, and Relations with Parents and Community), and a teacher's rating (Distinguished, Proficient, Professional Support Needed, or Unsatisfactory). Each dimension contributed up to three points toward the total possible (21 of 100) for the professional practices component of the teacher's final summative evaluation.

The professional practices mean scores improved for both the sample and the population between years one and two of the study, after the initial BPCM training and implementation. However, in the final year of the study, 2013-2014, the professional practices mean scores showed a slight decrease for both the sample and the population. For the 2013-2014 school year, year two of the BPCM implementation and year three of the study, the sample's professional practices scores ranged from 13.00 to 21.00 and had a mean of 19.73 (SD = .1.28) and a variance of 1.64. The population's professional practices scores ranged from 5.47 to 21.00 and had a lower mean ($M = 18.82, SD = 1.84$) with a variance of 3.42. Tables 10 and 11 contain the complete descriptive statistics for the sample and population's professional practices scores.

Table 10

Professional Practices Scores for Sample

Year	Range	Mean Statistic	Standard Deviation	Skewness	Kurtosis
2011-2012	6.10	19.37	1.25	-1.18	1.57
2012-2013	7.00	19.79	1.25	-1.93	4.81
2013-2014	8.00	19.73	1.28	-2.13	6.40

Table 11

Professional Practices Scores for Population

Year	Range	Mean Statistic	Standard Deviation	Skewness	Kurtosis
2011-2012	14.88	18.22	1.84	-.63	.19
2012-2013	12.32	18.83	1.82	-1.13	1.20
2013-2014	15.53	18.82	1.85	-1.55	3.92

An independent samples *t* test was conducted on the professional practices scores to evaluate whether the mean of the sample, BPCM participants, was significantly different than the mean score of the population, non-BPCM teachers in Brevard Public Schools with three years of performance data. Because the size of the sample and the population differed, equal variances were not assumed in the analysis of the results.

The 2013-2014 sample's mean professional practices score of 19.73 ($SD = 1.28$) was significantly greater than the mean professional practices score for the population ($M = 18.82$, $SD = 1.85$), $t(211.497) = -8.884$, $p = .000$, with a 95% confidence level. Over

the course of the study, BPS teachers who participated in BPCM demonstrated significantly stronger results in professional practices observation scores than their colleagues who did not participate in BPCM. Thus, the null hypothesis was rejected. Table 12 contains the complete results of the independent samples *t* test.

Table 12

Independent Samples t-Test, Overall Professional Practices Results for BPCM and Non-BPCM Participants

Descriptor	Mean	Standard Deviation	*t*	df	Sig. (2-tailed)
Non BPCM	18.82	1.90			
BPCM	19.73	1.28			
Equal variances not assumed			-8.884	211.497	.000

Note. BPCM = Brevard Peer Coaching Model

Further examinations of the data were conducted, removing outlier scores, to confirm appropriate use of the independent samples *t* test with slightly skewed data and unequal populations. Four outlier scores were removed from the 2013-2014 non-BPCM data set, professional practices scores less than 10. No scores met that parameter in the BPCM 2013-2014 data set; thus, no scores were removed as outliers. The difference between professional practices scores of the BPCM teachers, all results, and non-BPCM teachers, outliers removed, remained significant (*t (217.653) = .000*, at a 95% confidence level). Data also remained positively skewed, with and without the outlier scores. The

most frequent score for both BPCM and non-BPCM participants was 21, of a possible 21 points, and the majority of teachers in both groups (92.5%) earned ratings reflecting distinguished or proficient performance (18 points or higher). Because this was a client-based study and no significant difference was observed in the results when comparing all scores to scores with outliers removed, no scores were removed from further analysis in the study.

Research Question and Hypothesis 3

To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

H_{03}: There is no statistically significant difference in the VAM scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

The mean value-added (VAM) score for the study's sample improved from .08 ($SD = .40$) in 2011-2012, the baseline year of the study, to .11 ($SD = .24$) in 2013-2014, year two of the professional learning model's implementation and year three of the study. The 2012-2013 mean score ($M = .03$, $SD = .43$) was also higher than the 2011-2012 mean VAM score but slightly lower than the 2013-2014 mean VAM score. The mean range in scores across all three years of the study was 4.33, with a mean variance of 0.37. The

data were negatively skewed in years one and two of the study and positively skewed in year three, with a positive kurtosis in all three years.

The paired-samples *t* test was conducted to evaluate whether there was a significant difference between the baseline VAM scores and scores after year one and year two of BPCM. The results indicated that the mean VAM score for 2012-2013 (*M* = .03, *SD* = .43), though higher, was not significantly different than the mean 2011-2012 VAM score (*M* = .08, *SD* = .40, t(173) = -.564, at a 95% confidence interval). The evaluation of results between years one and two of BPCM implementation also indicated that the mean VAM score for 2013-2014 (*M* = .11 *SD* = .24) was not significantly different than the 2012-2013 results (*t*(173) = -.869, at a 95% confidence interval). Furthermore, though the mean VAM score improved from 2011-2012 to 2013-2014, the difference between the two scores was not statistically significant (*t*(173) = -1.142, at a 95% confidence interval) and yielded an effect size less than the minimum for small effect of .2 (*d* = .09). Thus, although the VAM scores increased for teachers who participated in the BPCM professional learning experience, there was not a statistically significant improvement in student achievement over the course of the study. Table 13 contains the results of the paired samples *t*-test.

Table 13

Paired Sample t-Test: Results for Value-added Scores for the Sample

Paired Sample	Mean	Standard Deviation	95% Confidence Interval		t	df	Sig. (2-tailed)
			Lower	Upper			
2011-2012 to 2012-2013	-.01	.16	-.03	-.02	-.564	173	.574
2012-2013 to 2013-2014	.02	.03	-.07	.03	-.869	173	.386
2011-2012 to 2013-2014	.03	.03	-.08	.02	1.1-1.142	173	.255

Further examinations of the data were conducted using the paired samples *t* tests on the sample teachers' value-added scores, separated by VAM type (TAV, SAV, and NFTAV). A comparison of the scores using the paired samples *t* test between years one and two of the study, years two and three of the study, and years one and three of the study, after the completion of the BPCM professional learning experience, yielded similar results. Though teachers who participated in BPCM training earned higher mean value-added scores over the course of the three years, the increase was not statistically significantly different. Thus, the null hypothesis was accepted. Tables 14, 15, and 16 contain the results of the analyses.

Table 14

Paired Sample t-Test: Results for Teacher Value-added (TAV) Scores

Paired Sample	Mean	Standard Deviation	*t*	df	Sig. (2-tailed)
2011-2012 to 2012-2013	.01	.14	.902	80	.370
2012-2013 to 2013-2014	-.06	.46	-1.103	80	.273
2011-2012 to 2013-2014	.04	.48	-.806	80	.423

Table 15

Paired Sample t-Test: Results for School Value-added (SAV) Scores

Paired Sample	Mean	Standard Deviation	*t*	df	Sig. (2-tailed)
2011-2012 to 2012-2013	-.01	.09	-.551	65	.583
2012-2013 to 2013-2014	.00	.07	.445	65	.658
2011-2012 to 2013-2014	-.00	.09	-.236	65	.814

Table 16

Paired Sample t-Test: Results for Non-FCAT Teacher Value-added (NFTAV) Scores

Paired Sample	Mean	Standard Deviation	t	df	Sig. (2-tailed)
2011-2012 to 2012-2013	-.07	.27	-1.342	26	.191
2012-2013 to 2013-2014	.02	.33	.241	26	.811
2011-2012 to 2013-2014	-.05	.21	-1.372	26	.182

Research Question and Hypothesis 4

To what extent do the professional practices scores of those teachers who

participated in peer coaching training (BPCM) differ from prior to training, at the

end of the first year after initial training, and at the end of the second year after

additional follow-up training?

$H_{04:}$ There is no statistically significant difference in the professional practices

scores of those teachers who participated in peer coaching training (BPCM) from

prior to training, at the end of the first year after initial training, and at the end of

the second year after additional follow-up training.

The mean professional practices score for the study's sample (*M* = 19.37, *SD* =

.09) improved from 2011-2012, the baseline year of the study, to 2013-2014, year two of

the professional learning model's implementation and year three of the study (*M* = 19.73,

SD = .10). The mean range in scores across all three years of the study was 7.03, with a

mean variance of -1.75. The data were negatively skewed in all three years with a positive kurtosis.

The paired-samples *t* test was conducted to evaluate whether there was a significant difference between the professional practices scores prior to BPCM training and scores after year one and year two of BPCM. The results indicated that the mean professional practices score for 2012-2013 ($M = 19.79$, $SD = 1.25$), after year one of BPCM implementation, was significantly different than the mean professional practices results in the baseline year ($M = 19.37$, $SD = 1.25$, $t(173) = -4.273$, at a 95% confidence interval). The evaluation of results between years one and two of BPCM implementation indicated that the mean professional practices score for 2013-2014 ($M = 19.73$, $SD = 1.28$) was actually slightly lower and not significantly different than the 2012-2013 results ($t(173) = .666$, at a 95% confidence interval). However, the difference between the mean baseline professional practices score from 2011-2012, prior to BPCM, and the mean score from 2013-2014, following two years of BPCM implementation, was significant ($t(173) = -3.544$, at a 95% confidence level), with a small effect size ($d - .27$). Although growth in the mean professional practices scores was greatest in year one of the initial implementation of BPCM and decreased in year two, the null hypothesis was rejected based on the overall statistically significant improvements in teacher effectiveness as measured by observation of professional practices for this study's sample. Table 17 contains the complete results of the paired sample t-test results for the professional practices scores for the sample.

Table 17

Paired Sample t-Test: Results for Professional Practices Scores for the Sample

Paired Sample	Mean	Standard Deviation	95% Confidence Interval		t	df	Sig. (2-tailed)
			Lower	Upper			
2011-2012 to 2012-2013	-.42	1.31	-.62	-.23	-4.273	173	.000
2012-2013 to 2013-2014	0.06	1.19	-.12	.24	.666	173	.506
2011-2012 to 2013-2014	.36	1.36	-.57	-.16	-3.544	173	.001

Summary

The size of the sample and population of this study and distribution of the data supported the use of an independent samples t test and a paired samples t test. These tests were used to determine the significance of the difference, if any, between teacher performance as measured by value-added and professional practices observation scores before, during, and after implementation of Brevard's Peer Coaching Model for professional learning (BPCM). A significant difference was found between professional practices scores, resulting in the rejection of the null hypothesis for Research Questions 2 and 4. However, although value-added scores improved over the course of the study, the increase in performance was not statistically significant, resulting in the acceptance of the null hypothesis for Research Questions 1 and 3. Chapter 5 includes a summary of the study, discussion of the findings, implications for practice, recommendations for further research, and conclusions.

CHAPTER 5
SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Introduction

This chapter contains a summary of this study, a discussion of the findings, implications for practice, and recommendations for further research. The purpose of this chapter is to provide further examination of this study's findings. It has been written to support a deeper understanding of the underlying concepts and their connection to improving instructional practice and student learning. It contains suggestions for further research to connect professional learning, implementation in the classroom, and student achievement results.

Summary of the Study

The purpose of this study was to analyze whether one public school district's investment in peer coaching training over a two-year period resulted in improved instructional practice and increased levels of student achievement as measured by Florida's value-added model (VAM). Based on the participatory leadership model of Demings (Burke et al., 2012; Gruska,2000; Kezar, 2001; Rook and Torbert, 2005), growth in performance should occur when professional learning is embedded systematically in an organization through teamwork, collaboration, collegial support, and a focus on continuous improvement.

The professional practices observation ratings for teachers from Brevard Public Schools teachers who participated in this study were used for this research in addition to

their value-added (VAM) scores. Professional practices scores were assigned by trained

school-based administrators across seven dimensions using a 4-point Likert-type scale

across four levels of ratings with rubric definitions: Distinguished, Proficient,

Professional Support Needed, or Unsatisfactory. Scores for this study's sample ranged

from 13.00 to 21.00, with 21.00 representing the maximum, and 19.73 representing the

mean in 2013-2014, the final year of the study. Value-added scores for the sample

ranged from -3.72 to 1.89, with .11 representing the mean in 2013-2014, the final year of

the study.

The professional practices ratings were taken from each teacher's Summative Part

One Evaluation as submitted to the Human Resources Division of the school district and

provided to the researcher in this study. The three-year aggregated value-added scores

were provided by the Florida Department of Education and were provided to the

researcher by the Testing and Accountability Department of the school district. The

independent samples *t*-test was the statistical method used to determine whether a

significant difference existed between the scores of this study's sample and population as

Brevard's Peer Coaching Model (BPCM) was implemented. The paired samples *t*-test

was the statistical method used to determine whether a statistically significant difference

developed in the sample's scores after participation in BPCM.

This study's sample included 174 Brevard Public School teachers with three years

of professional practices observation scores and value-added results who completed a

minimum of six of nine days of BPCM professional development during the course of

this study (2011-2014). This study's population, used for comparison with the paired

samples *t*-test, included 3,435 teachers from Brevard Public Schools who did not

participate in BPCM professional development and who earned three years of

professional practices observation scores and value-added student achievement results.

The sample and the population excluded instructional personnel from charter schools, the

adult education centers, alternative learning centers, and virtual schools associated with

Brevard Public Schools. A complete demographic analysis was presented for gender,

ethnicity, and years of experience. This study was based on four research questions:

1. To what extent do the VAM scores of those teachers who participated in peer

    coaching training (BPCM) differ from the VAM scores of the population of

    Brevard Public School teachers who did not participate in BPCM?

    H_{01} There is no statistically significant difference in the VAM scores of those

    teachers who participated in peer coaching training (BPCM) from the VAM

    scores of the population of Brevard Public School teachers who did not

    participate in BPCM.

2. To what extent do the professional practices scores of those teachers who

    participated in peer coaching training (BPCM) differ from the professional

    practices scores of the population of Brevard Public School teachers who did

    not participate in BPCM?

    H_{02}: There is no statistically significant difference in the professional

    practices scores of those teachers who participated in peer coaching training

    (BPCM) from the professional practices scores of the population of Brevard

    Public School teachers who did not participate in BPCM.

3. To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

   H_{03}: There is no statistically significant difference in the VAM scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

4. To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

   H_{04}: There is no statistically significant difference in the professional practices scores of those teachers who participated in peer coaching training (BPCM) from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training.

Summary and Discussion of Findings

Previous researchers on the impact of a peer coaching professional learning model on systemic improvement in teaching and learning have reported positive findings but have also primarily relied on qualitative measures (Hill & Rapp, 2012; Johnson & Fiarman, 2012; Kohler et al., 1997; Sparks & Bruder, 1987). The purpose of this study

was to use quantitative measures to determine if significant improvements in observation ratings and value-added scores occurred after one large school district's significant investment in a multi-year peer coaching training program. The following summary and discussion of the findings has been organized around each of the study's four research questions.

Research Question 1

To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from the VAM scores of the population of Brevard Public School teachers who did not participate in BPCM?

The results of the data analysis for Research Question 1 indicated no significant difference existed between student achievement results of teachers in the sample and teachers in the population when measured by Florida's value-added model formula (VAM). Mean VAM scores of those teachers who participated in peer coaching training (BPCM), though higher than the mean VAM scores of those teachers who did not participate in BPCM throughout this study (non-BPCM), did not show a significant improvement over the scores of their peers, replicating the previous research findings of Murray et al. (2008). Over the course of this study, the mean VAM scores for teachers in both groups, BPCM and non-BPCM, increased as predicted, and the range in scores narrowed for BPCM and non-BPCM teachers. However, though BPCM teachers earned higher VAM scores than non-BPCM teachers in all three years of this study, both before and after BPCM implementation, and mean scores in both the sample and the population

increased over the course of the study, the difference in student achievement between the groups was not statistically significant despite the large study sample and duration of the professional development experience. Although the school district's use of three-year aggregated results in teacher evaluation minimized the impact of measurement error and variability for teacher performance ratings, it limited the researcher's efforts to capture statistically significant changes in student achievement after each year of implementation of the Brevard Peer Coaching Model (BPCM) project.

The data were also analyzed for differences in the value-added results by VAM type in order to determine if differences existed among teachers earning a teacher-aggregated VAM (TAV), a non-FCAT teacher-aggregated VAM (NFTAV), or a school-aggregated VAM (SAV). The results indicated that for BPCM and non-BPCM teachers earning a TAV or NFTAV, mean scores were higher for teachers who participated in BPCM than the mean scores of teachers who did not participate in BPCM. The differences were not, however, statistically significant. Teachers in BPCM who earned a SAV also had higher mean scores than teachers who did not participate in BPCM, but the difference between the mean SAV scores in this group was statistically significant. BPCM training had a stronger impact on all teachers in schools with higher levels of student growth than projected. An important distinguishing characteristic between the sample and the population related to student assignment. More BPCM teachers earned a TAV (n = 46.6%) than non-BPCM teachers (n = 28.5%), indicating that more of the BPCM participants were classroom-based reading, language arts, or mathematics teachers in Grades 4-10. Future studies would benefit from implementing new

professional learning experiences with a true random sample, representative of the general comparison population. They would also be enhanced by defining a measure of student performance more predictive of a single year's growth in performance for teachers across grade levels and content areas. Florida's value-added formula uses three-year aggregated scores from a single measure, the state assessment test in reading and mathematics.

The intent of the peer coaching training model was to improve teacher and student performance, both for those teachers who participated in the training and their colleagues, through implementation of the observation and feedback cycle in schools and based on Demings' guiding theory of participatory leadership (Burke et al., 2012; Gruska,2000; Kezar, 2001; Rook and Torbert, 2005). The overall effectiveness of a peer coaching model is measured in terms of impact on those trained and their colleagues, as the professional learning builds on and extends professional capital, thus keeping professional learning in the formal and informal conversations among teachers and between teachers and administrators. Over the course of this study, value-added scores improved, as predicted, even though the statistical analysis of the data demonstrated a weak effect. The analysis of the data indicated that BPCM had a positive impact on student performance, as value-added results improved consistently across the sample and the population, although clearly not at the desired level of influence. However, because the school district used a three-year aggregated value-added score in its teacher evaluation program to minimize the impact of standard error, the researcher's ability to evaluate significant changes in performance from year to year was limited. This

indicated that though increases in scores may appear promising, they require further and deeper analysis when conducting program evaluation for both effectiveness and efficiency. In addition, program evaluation that is planned strategically in the preliminary stages of a professional learning model's development and implementation is important. Early planning can be used to identify assessment tools and measurement outcomes that will provide results necessary for monitoring and analyzing the program's effectiveness over the course of the program.

Research Question 2

To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from the professional practices scores of the population of Brevard Public School teachers who did not participate in BPCM?

The results of the data analysis for Research Question 2 indicated there was a significant difference between the mean professional practices observation score of the teachers who participated in BPCM compared to the mean professional practices observation score of the teachers who did not participate in BPCM. Mean professional practices scores improved for teachers in both groups of teachers, BPCM and non-BPCM, from the baseline year of this study to 2013-2014, year two of BPCM implementation. There was a slight decrease, however, in mean professional practices scores between 2012-2013 and 2013-2014, years one and two of BPCM implementation, for BPCM and non-BPCM participants, as the school district increased training for

105

administrators on the observation rubrics, recertified all evaluators on use of the evaluation instruments, and emphasized the importance of inter-rater reliability (J. Respess, personal communication, March 24, 2014).

Overall, teacher performance in Brevard Public Schools, as measured by administrator-assigned observation scores, improved over the course of this study. Teachers who participated in BPCM improved significantly more than their colleagues. The data indicated that the professional learning experienced by BPCM participants had a positive impact on the performance of their peers, aligning with previous research on the positive impacts of peer learning and peer coaching (Joyce & Showers, 2002; Louis & Marks, 1998; Marzano & Simms, 2013). However, the data further indicated that throughout the three-year length of the study, the professional practices results were positively skewed. Administrators rated 92.5% of teachers in both the sample and the population as highly effective or effective in the area of professional practices. As found in the Widget Effect (Weisberg et al., 2009), principals struggled to distinguish between levels of performance among their teachers, even with clearly defined descriptors of teacher performance in the IPPAS rubrics.

In this study, school-based instructional leaders have been shown to continue to need additional training and support in distinguishing between levels of teacher performance and providing unbiased feedback that will improve instructional practice and increase student performance. It has also been demonstrated in the present study that the greatest impact of professional learning occurs in the initial year of training and for teachers with direct participation in the professional learning experience. Although the

investment in training yielded dividends for the school district beyond the participants themselves, additional research is needed to assess the level of impact and to identify strategies that will support improvements beyond year one.

Research Question 3

To what extent do the VAM scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

The results of the data analysis for Research Question 3 indicated no significant difference in the mean value-added scores (VAM) for those teachers participating in BPCM over the course of the model's implementation. The mean VAM score for BPCM teachers increased each year of BPCM implementation, from a baseline score of 0.77 to .11 in 2013-2014. Based on these results, BPCM teachers were having a greater impact on student learning over the course of this study, but the difference was not statistically significant. The range in value-added scores was small, and the school district's use of three-year aggregated value-added scores to reduce measurement error limited the researcher's ability to analyze changes in student learning in a single year. This indicated that accurate program evaluation requires statistical analysis beyond a comparison of raw mean scores. Establishing a direct link between professional development and student achievement remains a challenge (Borko, 2004; Gusky, 2002; Joyce & Showers, 2002; Yoon et al., 2007).

Research Question 4

To what extent do the professional practices scores of those teachers who participated in peer coaching training (BPCM) differ from prior to training, at the end of the first year after initial training, and at the end of the second year after additional follow-up training?

The results of the data analysis for Research Question 4 indicated a significant difference between the baseline professional practices scores of those teachers who participated in peer coaching training (BPCM) and their scores after years one and two of BPCM implementation. Teacher performance as measured by administrator-assigned observation scores improved most significantly after the first year of BPCM, indicating the greatest gains occurred after year one of the training. Teacher performance continued to improve in the second year, but at a less significant rate (p = .506). Overall, mean professional practices scores improved significantly from 19.37 in 2011-2012, prior to BPCM, to 19.73 in 2013-2014, after BPCM. In this study, the greatest improvements in teacher performance occurred after the first year of the peer coaching professional learning model, replicating the results of Lee and Maerten-Rivera (2012), and indicating a need for further examination of strategies to support BPCM participants in implementation and continuous improvement in year two and beyond.

Implications for Practice

Participatory leadership, the theoretical framework underlying the development of the Brevard Peer Coaching Model (BPCM) in Brevard Public Schools, describes the

importance of putting structures in place to transform schools and school districts into high quality learning organizations focused on continuous improvement. Through peer observation and meaningful feedback, peer learning promotes changes in professional practice that produce increases in student learning. In a large school district, building systemic capacity in schools is essential for success. BPCM was designed to identify high performing teacher teams at each school and to provide training in analyzing the teacher evaluation rubrics and using the rubrics to observe and provide specific feedback to peers. BPCM was intended to promote better understanding among teachers of the high quality instructional practices identified in the observation rubrics and to improve performance of both BPCM and non-BPCM participants over a two year period.

Results of this local study indicated increased value-added and professional practices scores for Brevard Public School teachers in the sample, BPCM, and in the population, non-BPCM. Furthermore, BPCM teachers consistently earned higher value-added and professional practices scores than their non-BPCM peers, before, during, and after the implementation of the professional learning model. However, improvements in value-added results both within the sample during the study and comparing the sample and the population mean scores were not at a statistically significant level within a 95% confidence interval. Florida's value-added model was designed to measure the impact of school and teacher performance on student learning in a more fair, equitable manner than previous tools which relied on proficiency ratings or learning gains and failed to consider student characteristics known to impact student growth (Florida Department of Education, 2011). This study has demonstrated that it is not an effective tool to capture

the impact of professional learning models in an academic year. The measurements designated for use in this school district's teacher evaluation system were not effective in evaluating the impact of a multi-year professional learning model on student learning. School districts planning to implement a professional learning model should plan for program evaluation in the initial stages of development, in order to identify more valid and reliable measurements to student performance to evaluate a model's effectiveness during and after implementation.

Professional practices scores used for evaluation in this study, though demonstrating significant improvements in instructional practices for both BPCM and non-BPCM participants, also reflected the continued failure of school administrators to use evaluation ratings and observation results to accurately distinguish between levels of teacher performance, a systemic problem reflected in research from other local school districts and beyond (Fritz, 2014; Weisberg et al., 2009). BPS school administrators rated 92.5% of all teachers as highly effective or effective in the professional practices portion of the teacher evaluation instrument during the final year of this study (2013-2014). This was consistent with administrators' ratings of previous years. Although the school district implemented additional training and certification requirements for all administrators in the same year, 2013-2014, objective and unbiased use of the observation framework to evaluate teacher performance clearly remains a significant concern. Continued training for administrators and emphasis on inter-rater reliability are needed. Furthermore, multiple observers have been demonstrated to minimize bias and increase validity and reliability of evaluation ratings in other school districts (Ensuring

Fair and Reliable Measures of Effective Teaching, 2013). This school district should consider a more multi-metric approach by building on the current level of BPCM training and taking steps necessary to further develop the capacity of the peer coaches to become true partners in the teacher evaluation process as peer evaluators.

One limitation of this study was related to the quasi-experimental design which lacked equally sized, randomly chosen sample and control groups for evaluation of the professional learning experience. This element of the study design limited interpretation of the results to impact in the local school district and should be given more consideration in planning for future program evaluation projects. In addition, the study design lacked a component to evaluate frequency or quality of feedback provided to peers after observations by the BPCM peer coaches. The school district has added an electronic system to capture and report observation and evaluation data to improve capacity to monitor results of formal and informal observations at the school and district level. The system includes a mentoring module, where peer coaches can record results from peer observations and provide specific feedback and suggestions for next steps via email to their peers in timely manner. This system should be fully implemented and monitored by school and district administrators to evaluate quality of the peer coaching encounters. It can be useful in the identification of needs for follow-up training and support to maximize efforts to improve teaching and learning using the peer observation process.

Finally, though not a part of this study, a correlation was run to determine the strength of the relationship between professional practices scores and VAM results for the BPCM participants in the final year of the study, 2013-2014. The results, $r(172) =$

+.005, p = .949, indicated a positive but very weak relationship between the two key elements of this program evaluation study and the BPS teacher evaluation system. Although professional practices scores increased as student achievement results improved, and vice versa, there was not a significant relationship between the two scores at the 95th confidence interval. Similarly, weak correlation results were found in the first year of implementation of the teacher evaluation system and BPCM (Mela, 2013). This indicated a continued need for the school district to review current research and analyze the teacher evaluation components. Continued attention to observation frameworks and student achievement outcomes will assist in ensuring that the system accurately identifies critical teacher competencies. Doing so will permit meaningful feedback to be shared with teachers and administrators that will result in improved instructional practice and increased student learning.

<div align="center">Recommendations for Further Research</div>

The publication of the updated Standards for Professional Learning by Learning Forward in 2011 launched a renewed focus on the development of high quality learning experiences for teachers and the need to monitor and measure the impact of professional learning on improving practice and increasing student learning. When school districts allocate resources to professional development, "PD to practice" should be the goal, and program evaluation is essential to ensure that investments of financial and human resources are producing a quality return. Another development in 2011, the implementation of new multi-metric teacher evaluation systems incorporating student

achievement outcomes, has also changed the landscape in public education. Both

developments informed the direction of this local study and provide opportunities for

future research.

Although the researcher in this study chose to focus on the quantitative measures

of professional practices observation scores and value-added results, previous studies

analyzing the effectiveness of peer coaching have often relied on qualitative measures

including survey results. A follow-up study for this local school district and other school

districts implementing a comprehensive professional learning model should include both

elements: (a) an analysis of the evaluation ratings from administrators and student

outcomes before, during, and after the professional learning; and (b) a perception survey

of study participants, in order to determine which, if any, elements of the professional

learning and follow-up were most helpful in improving instructional practice. The 21-

statement paper/pencil survey, *Teacher Perceptions of Professional Learning (TPLP)*

(Yates, 2007), gathers perception data around the experiential and reflective elements of

professional learning, as well as collaborative sharing and the relatedness of the

professional learning to a teacher's work with students. This would allow school districts

to study components of the training model to determine if different elements had greater

effects on professional practices or student achievement results. Survey results, as well

as interim quantitative measures, should be used to adapt training needs and resources as

needed during the study to ensure that significant impact continues beyond year one,

when impact has been historically greatest (Lee & Maerten-Rivera, 2012).

School districts developing and implementing a comprehensive professional learning model should also analyze attrition of a study's participants over time, in order to ensure that resources are being used effectively and efficiently to maximize professional growth. Over 100 teachers who participated in the first three-day summer training for BPCM dropped out of the program over the course of the study. Utilizing results from the TPLP survey to guide changes in the professional learning model's training sessions and follow-up would strengthen future professional development initiatives and potentially improve teacher retention, changes in practice, and increases in student learning.

A complicating factor in this study included the state-level differences in application of the value-added formula results for teachers assigned to grade levels and/or content areas not assessed by the state's standardized assessments for reading and mathematics. The local school district chose to apply either the NFTAV, the non-FCAT teacher aggregated VAM, or the SAV, the school-aggregated VAM, in teacher evaluation and for the analysis in this local study. Study results were similar for all three VAM types; however, issues of fairness and equity should guide future research to extend application of the principles and limiting characteristics of Florida's value-added model to all grade levels and content areas.

Finally, it is essential to ensure that elements guiding teacher evaluation and professional learning accurately identify instructional strategies, techniques, and skills that produce significant increases in student learning. The weak correlation between professional practices scores and value-added results demands further investigation and

analysis in this school district, and in school districts implementing similar evaluation

systems and frameworks. Effective management of human capital requires a clear vision

of high quality teaching that is research-based, outcome-driven, and the focus of all high

quality professional learning experiences.

Summary

Brevard Public Schools developed and implemented a comprehensive

professional learning model, Brevard's Peer Coaching Model (BPCM), during the

summer following the 2011-2012 school year, to support teachers and administrators in

improving instructional practice and student learning through a peer coaching and

feedback process. Designed using Learning Forward's Standards for Professional

Learning (2011), BPCM included nine days of training and follow-up over the course of

two school years and represented a significant investment of human and financial

resources by the school district. Building on the participatory leadership theoretical

framework, the training included multiple opportunities over the course of the study for

teachers to learn from teachers in a collegial, collaborative structure through peer

observation, feedback, and coaching, with the intent of improving practice for both

BPCM and non-BPCM participants.

To improve upon previous research and evaluate the effectiveness of the

professional learning experience, this local study was conducted to analyze the

quantitative measures of professional practices scores and value-added results before,

during, and after the training. Although both measures showed improvement over the

course of the study for both BPCM and non-BPCM participants, only professional

practices scores demonstrated a significant increase. VAM scores improved but not at a

statistically significant level. This confirmed the findings of previous research that

though peer coaching had a positive impact on teacher practice, a strong link between the

BPCM professional learning and student outcomes could not be determined from the data

available. The need for further analysis in the school district of the relationship between

professional practices scores and value-added measures was also demonstrated. Further

analysis could better ensure the use of observation frameworks and measurement

instruments that accurately identify instructional practices, techniques and skills that

produce significant increases in student learning, and that administrators are using the

evaluation tools with fidelity. The findings in the present study also were aligned with

those of previous researchers related to administrator bias, inter-rater reliability, and the

value of multi-metric evaluation systems in distinguishing between levels of teacher

performance.

Investments in professional learning represent a significant allocation of a school

district's limited resources, and high quality program evaluation is critical to ensure that

allocation decisions are research-based and outcome driven. The importance of planning

for program evaluation during the initial stages of development of a professional learning

model has been demonstrated in the present study. Planning is essential to ensure that

measurements are identified that will provide meaningful results before, during, and after

implementation. Also critical is aligning professional learning with valid and reliable

evaluation tools. Successful organizations recognize that teacher performance is crucial

in determining student learning and plan for professional learning to support continuous

improvement in both.

APPENDIX A
LITERATURE REVIEW: KEY TERMS AND AUTHORS

KEY TERMS AND AUTHORS ASSOCIATED WITH THREE MAJOR SEARCH CATEGORIES

| | | Evaluation |
Professional Learning	Peer Coaching	of Teacher Effectiveness
Key Terms		
Adult Learning	Peer Coaching	Teacher Effectiveness
Peer Learning	Instructional Coach	Teacher Quality
Teacher Training	Teacher Collaboration	Teacher Performance
Learning Centered Schools	Peer Learning	Teacher Evaluation
Professional Learning Communities		Teacher Observation
Professional Development		Accountability
Teacher Effectiveness		Value-added Measure
		Student Learning
Authors		
Barth	Bartalo	Baker
Borko	Bartholomew	Barton
Bruegmann	Summer	Briggs
Caffarella	Bearwald	Chiang
Curtis	Bowman	Coulter
Darling-Hammond	Brown	Danielson
Day	Bruce	Darling-Hammond
Demings	Tomlinson	Domingue
Dubois	Bruder	Dufour
Duff Gusky	Tschannen-Moran	Fritz
Hirsch	Bryan	Frontier
Horn	Christen	Gallagher-Fishbone
Ingvarson	Clifton	Gill
Islas	Collet	Glazerman
Jackson	Crilley	Goldhaber
Joyce	Curry	Green
Kezar	Cushman	Griffin
Knowles	Demonte	Haertel
Lee	Eisenbach	Hall
Little	Fiarman	Harris
Louis	Fullan	Hershberg
Maerten-Rivera	Gallimore	Hill
Marks	Garrett	Ingle
Merriam	Good	Kapitula
Sachs	Hall	Keeling
Showers	Harrison	Kelley
Valli	Hasbrouch	Kimball
Yates	Hill	Kupermintz
Yoon	Johnson	Ladd
	Joyce	Linn
	Killion	Lipscomb
	Knight	Loeb
	Koch	Martineau
	Kohler	Marzano
	Kretlow	Maslow
	Leana	McCarthy
	Ma	Mielke
	Marietta	Milanowok

Professional Learning	Peer Coaching	Evaluation of Teacher Effectiveness
	Marzano	Mulhorn
	Mazar	Oluwole
	McCormick	Papay
	Mousched	Pease
	Murray	Raedenbach
	Rapp	Ravitch
	Ross	Robertson-Kraft
	Shearer	Rothman
	Showers	Rothstein
	Simeral	Rutledge
	Simms	Sartain
	Sinkinson	Scannella
	Sparks	Sexton
		Shavelson
		Shepard
		Sparks
		Staiger
		Stecher
		Umland
		Weisberg
		Whitehurst
		Wise

APPENDIX B
BREVARD PUBLIC SCHOOLS:
INSTRUCTIONAL PERSONNEL PERFORMANCE APPRAISAL SYSTEM (IPPAS)

BPS IPPAS Survey Question 3 year scope

Metrix are based on subject matter identified as instructional only

1. The new Instructional Personnel Performance Appraisal System (IPPAS) has resulted in a fair evaluation of my professional practices as evidences in the 21 point scale score I received.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	186	8.4%	110	3.8%	122	3.4%
Agree	605	27.3%	395	13.6%	369	10.3%
Neither Agree Nor Disagree	359	16.2%	265	9.1%	258	7.2%
Disagree	443	20.0%	276	9.5%	284	8.0%
Strongly Disagree	562	25.4%	430	14.8%	344	9.6%
Don't Know	59	2.7%	29	1.0%	30	0.8%
Total including Skipped	2,214		2,900		3,568	

2. The elements of the rubric for professional practices help me understand the elements of teaching that would be evaluated.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	212	9.6%	191	6.6%	158	4.4%
Agree	882	39.8%	660	22.9%	606	17.0%
Neither Agree Nor Disagree	376	17.0%	231	8.0%	230	6.4%
Disagree	337	15.2%	183	6.3%	204	5.7%
Strongly Disagree	384	16.4%	218	7.6%	186	5.2%
Don't Know	21	0.9%	15	0.5%	18	0.5%
Total including Skipped	2,214		2,886		3,568	

3. The elements of the rubric for the Professional Growth Plan (PGP) helped guide me to develop a good PGP.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	212	10%	193	6.7%	161	4.5%
Agree	708	31.9%	528	18.3%	489	13.7%
Neither Agree Nor Disagree	406	18.3%	280	9.7%	275	7.7%
Disagree	426	19.2%	233	8.1%	229	6.4%
Strongly Disagree	364	16.4%	248	8.6%	231	6.5%
Don't Know	20	0.9%	12	0.6%	18	0.5%
Total Including Skipped	2,222		2,886		3,568	

4. The IPPAS supported more reflection upon my teaching and improved my practice.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	183	8.3%	133	4.6%	105	2.9%
Agree	646	29.2%	446	15.5%	383	10.7%
Neither Agree Nor Disagree	439	19.8%	325	11.3%	336	9.4%
Disagree	451	20.4%	288	10.0%	290	8.1%
Strongly Disagree	364	16.4%	287	10.0%	271	7.6%
Don't Know	25	1.1%	17	0.6%	17	0.5%
Total Including Skipped	2,215		2,877		3,568	

5. I was able to get answers I needed prior to PGP development

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	218	9.83	213	7.4%	218	6.1%
Agree	708	31.9%	628	21.7%	664	18.6
Neither Agree Nor Disagree	385	17.4%	292	10.1%	274	7.7%
Disagree	477	21.5%	189	6.5	125	3.5%
Strongly Disagree	364	16.4%	165	5.7%	108	3.0%
Don't Know	15	9.8%	15	0.5%	17	6.1%
Total including Skipped	2,218		2,893		3,568	

6. I had the opportunity to discuss my scores with my administration and explore through the rubric what behaviors justified my scores

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	430	22.1%	323	11.2%	338	9.5%
Agree	1,082	48.6%	765	26.0%	721	20.2%
Neither Agree Nor Disagree	270	12.2%	176	6.1%	184	5.2%
Disagree	189	8.5%	112	3.9	78	2.2%
Strongly Disagree	364	16.4%	102	3.5%	74	2.1%
Don't Know	19	0.7%	19	0.7%	16	0.4%
Total including Skipped	2,215		2,884		3,568	

July 30, 2014

125

7. I participated in the peer observation process through colleagues observing me/ me observing colleagues.

Label	2011-2012		2012-2013		2013-2014	
0-3 Times	1,181	53.5%	845	29.5%	814	22.8%
4-5 Times	684	31.0%	416	14.5%	390	10.9%
More then 5 Times	343	15.5%	212	7.4%	200	5.6%
Total Including Skipped	2,208		2,868		3,568	

8. I believe this process of sharing feedback with and giving feedback to my colleagues contributed to my professional growth.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	266	12.0%	226	7.6%	208	5.8%
Agree	819	46.6%	644	22.2%	605	17.0%
Neither Agree Nor Disagree	491	22.1%	300	10.3%	311	8.7%
Disagree	330	46.6%	178	6.1%	149	4.2%
Strongly Disagree	364	16.4%	125	4.3%	114	3.2%
Don't Know	21	0.9%	19	0.7%	17	0.5%
Total including Skipped	2,225		2,900		3,568	

5

July 30, 2014

9. Our collaborative team fostered sharing of strategies and professional practices.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	406	18.4%	334	11.6%	338	9.5%
Agree	1,041	47.1%	777	11.6%	712	21.8%
Neither Agree Nor Disagree	343	18.4%	197	6.8%	196	5.5%
Disagree	211	9.5%	106	6.8%	89	2.5%
Strongly Disagree	384	16.5%	63	2.2%	56	1.6%
Don't Know	13	0.6%	16	0.6%	13	0.4%
Total including Skipped	2,210		2,683		3,568	

10. Our collaborative team worked together on behalf of identified students.

Label	2011-2012		2012-2013		2013-2014	
Strongly Agree	475	21.5%	394	13.7%	372	10.4%
Agree	1,084	49.0%	778	27.0%	712	20.0%
Neither Agree Nor Disagree	300	13.6%	168	5.8%	186	5.2%
Disagree	180	8.1%	80	2.8%	76	2.1%
Strongly Disagree	143	6.5%	54	1.9%	39	1.1%
Don't Know	17	0.8%	17	0.6%	17	0.5%
Total including Skipped	2,214		2,879		3,568	

July 30, 2014

APPENDIX C
UNIVERSITY OF CENTRAL FLORIDA
INSTITUTIONAL REVIEW BOARD APPROVAL

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

University of
Central
Florida

NOT HUMAN RESEARCH DETERMINATION

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

To   :   **Debra Pace**

Date :   **August 05, 2014**

Dear Researcher:

On 8/5/2014 the IRB determined that the following proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56:

Type of Review:	Not Human Research Determination
Project Title:	A Study of the Impact of Brevard Public School's Peer Coaching Model on Student Achievement Outcomes and Teacher Evaluation Results
Investigator:	Debra Pace
IRB ID:	SBE-14-10479
Funding Agency:	
Grant Title:	
Research ID:	N/A

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

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

Signature applied by Patria Davis on 08/05/2014 02:48:30 PM EDT

IRB Coordinator

APPENDIX D
SCHOOL DISTRICT APPROVAL OF STUDY

School Board of Brevard County
2700 Judge Fran Jamieson Way Viera, FL 32940-6699
Dr. Brian Binggeli, Superintendent

School Board

July 29, 2013

Dear Mrs. Pace,

Thank you for your application to conduct research in the Brevard Public Schools. This letter is official verification that your application has been accepted and approved through the Office of Accountability, Testing, & Evaluation. Upon the completion of your research, please submit your findings to our office. If we can be of further assistance, do not hesitate to contact our office.

Sincerely,

Vickie B. Hickey

Vickie B. Hickey, Resource Teacher
Office of Accountability, Testing, and Evaluation

Office of Accountability, Testing & Evaluation
Phone: (321) 633-1000 FAX: (321) 633-3465

131

Office of Accountability, Testing, and Evaluations
Research Application

Assurances Form

I understand that I am requesting permission to engage in a research Project, and I am not requesting information pursuant to Open Records Legislation. If my research project requires participation with students, I understand that I may be subject to the appropriate School Board policy regarding background investigations, as well as any applicable costs associated. Additionally, if my request is granted, I agree to abide by all policies, rules and regulations of the District, INCLUDING **THE** SECURING OF WRITTEN PARENT PERMISSION PRIOR TO IMPLEMENTATION OF MY PROJECT.

X _____

Researcher

_____7/a4/13_____
Date

I have read the procedures for Research Projects in the Brevard County Public School System and understand that supervision of this project and responsibility for an outcome report rests with me. I also understand that the privileges of conducting future studies in the Brevard County Public School System is conditioned upon the fulfillment of such obligations.

Barbara A. Murray

Sponsor/Advisor of Research Project
(signature required for student research)

_____7/24/13_____
Date

Approval of Office of Accountability, Testing and Evaluation*.

_____
Signature

_____7/24/13_____
Date

*Approval of the study at the district level does not obligate principals to participate in the proposed research.

Approval of Principal*:

_____
Signature

_____
Date

*The principal's signature suggests that the research project has been reviewed and that the school will participate, subject to the researcher's compliance with District policies.

6

APPENDIX E
BREVARD PUBLIC SCHOOLS
INSTRUCTIONAL PERSONNEL EVALUATION INSTRUMENT

BPS Instructional Personnel Evaluation Instrument

Name _____ School Year _____
　　　　　　Last　　　　　　　First　　　　　MI

School/Dept Name _____ School/Dept # _____ Contract Status: INTERIM ☐ ANNUAL ☐

Principal/Dept Head _____ Assignment _____ New Hire ☐ AC ☐ CC/PSC ☐

A. Quality of Instruction *(Must be completed for all certificated employees)*:

Instructional Design and Lesson Planning	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN
1. Sets instructional outcomes and aligns instruction with state-adopted standards.								
2. Designs instruction using student prior knowledge and diagnostic student data to build coherent lessons.								
3. Designs ways to monitor learning, designs use of appropriate instructional resources, and designs formative assessments.								
4. Requires students to understand and demonstrate skills and competencies.								

SELF [_____]　　　ADMIN [_____]

Teacher Comments:

Administrator Comments:

Learning Environment	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN
1. Creates and maintains a safe and organized learning environment.								
2. Promotes a flexible, inclusive, collaborative, and student-centered learning environment.								
3. Allocates and manages time, space, resources.								
4. Manages student conduct.								
5. Conveys high expectations, respects students' cultural backgrounds and accommodates diverse needs of students.								
6. Models and teaches clear, acceptable communication skills.								
7. Maintains a climate of inquiry.								
8. Integrates appropriate technologies.								

SELF [_____]　　　ADMIN [_____]

Teacher Comments:

Administrator Comments:

Our Mission is to Serve Every Student with Excellence as the Standard!

Instructional Delivery and Facilitation

	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN

1. Utilizes knowledge of the subject to enrich students' understanding, identifies gaps in students' learning, and modifies instruction to respond to student misconceptions.
2. Employs higher order questions.
3. Applies varied instructional strategies and resources.
4. Delivers engaging, challenging and relevant lessons.
5. Differentiates instruction.
6. Provides immediate and specific feedback to students.

SELF [   ]   ADMIN [   ]

Teacher Comments:

Administrator Comments:

Assessment

	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN

1. Analyzes and applies data from multiple measures to diagnose students' learning needs, inform instruction, and monitor progress.
2. Designs and uses formative and summative assessments that lead to mastery.
3. Modifies teacher made assessments to accommodate diversity.
4. Communicates assessment data to students and parents.

SELF [   ]   ADMIN [   ]

Teacher Comments:

Administrator Comments:

B. Quality of Instruction *(Must be completed for all certificated employees)*:

Professional Responsibility and Ethical Conduct

	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN

1. Demonstrates punctuality, attendance, and timely completion of records and reports.
2. Performs assigned duties and complies with policies, procedures, programs, and the Code of Ethics.
3. Demonstrates professionalism.
4. Initiates professional communication with appropriate stakeholders.
5. Applies technology to organize and communicate assessment information.

SELF [   ]   ADMIN [   ]

Teacher Comments:

Administrator Comments:

IPPAS Appendix: 36

Our Mission is to Serve Every Student with Excellence as the Standard!

135

Relationships with Students

	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN
1. Displays knowledge and understanding of how students learn and applies knowledge to building positive relationships with students.								
2. Teacher brings student interest into content.								
3. Teacher interactions with students reflect respect and caring for the individual as well as groups of students.								

SELF [　　] ADMIN [　　]

Teacher Comments:

Administrator Comments:

Relationships with Parents and Community

	Distinguished		Proficient		Professional Support Needed		Unsatisfactory	
	3		2		1		0	
	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN	SELF	ADMIN
1. Establishes early contact with parents.								
2. Establishes a relationship of mutual trust and respect with parents.								
3. Maintains a family-friendly learning environment.								
4. Maintains two-way communication with parents.								
5. Responsive to parent needs.								
6. Promotes parent understanding of academic standards and expectations.								
7. Reports to parents about student performance and progress.								
8. Provides opportunities to assist student learning at home.								
9. Provides information about school and community resources to parents.								
10. Provides meaningful connections with learning and the community.								

SELF [　　] ADMIN [　　]

Teacher Comments:

Administrator Comments:

A. FORMATIVE SCORING BY ADMINISTRATOR:

Component	Possible Points	Points Earned
Total Observations Points	21	

Our Mission is to Serve Every Student with Excellence as the Standard!
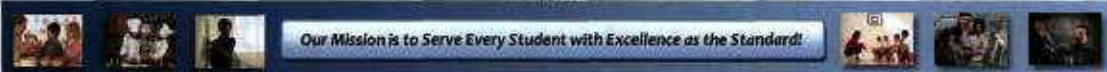
Principal Comments:

_____
_____
_____
_____

Teacher Comments *(Optional)*:

_____
_____
_____
_____

My signature indicates that this evaluation has been discussed with me:

_____ / _____
Signature of Employee   *(Blue Ink Only)*   Date

_____ / _____
Signature of Principal/Administrative Supervisor   (Required)   *(Blue Ink Only)*   Date

_____ / _____
Signature of Assistant Principal   (Required)   *(Blue Ink Only)*   Date

IPPAS Appendix: 38

Our Mission is to Serve Every Student with Excellence as the Standard!

137

APPENDIX F
SEVEN DIMENSIONS: FLORIDA EDUCATOR ACCOMPLISHED PRACTICES

Instructional Design & Lesson Planning

Learning Environment

Instructional Delivery & Facilitation

Assessment

Professional Responsibilities & Ethical Conduct

Relationship with Students

Relationships with Parents & Community

APPENDIX G
RUBRICS: BPS INSTRUCTIONAL PERSONNEL EVALUATION INSTRUMENT

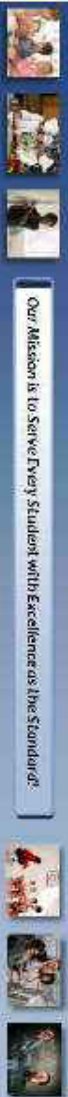| Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Sets instructional outcomes and aligns instruction with state-adopted standards	1. Instructional goals and objectives are consistently clear; focus on student learning and achievement of standards and the end result is complete and coherent; directions, procedures, and content are appropriate for and clear to all students. 2. Plans routinely to provide for instruction to meet the needs of students with varied levels.	1. Lesson plans are aligned to the district adopted curriculum maps and district/state assessments.	1. Lesson plans are not consistently aligned to the district adopted curriculum maps and district/state assessments.	1. Lesson plans, pacing guides, when available, do not align with the district adopted curriculum maps and/or district/state assessments.
Designs instruction using student prior knowledge and diagnostic student data to build coherent lessons	1. Regularly accesses and systematically uses data at the item level to find strengths and challenges both for disaggregated groups and for individual students. 2. Uses these data to accurately refine and modify instruction for whole group, small groups and for specific individuals.	1. Makes accurate use of student achievement data and uses prior student knowledge when making instructional decisions. 2. Examines data at the item level to find strengths and challenges for disaggregated groups of students to appropriately plan for instruction for whole and small groups.	1. Accesses data to view class achievement and examines data but shows little application to make instructional decisions linked to analysis of data. 2. Inferences about the data may not be complete. Examines data at the group level and uses these data when planning instruction.	1. Relies on someone else to access student achievement data. 2. When data is accessed, it may not be used to make decisions. 3. Does not consider data to make changes in instruction.
Designs ways to monitor learning, designs use of appropriate instructional resources, and designs formative assessments	1. High quality classroom formative and summative assessments are designed that accurately measure student learning of the lesson or unit objectives.	1. Lesson or unit is planned in detail around clearly defined lesson objectives. 2. Progression and pacing of the planned learning time (instructional strategies, student activities, use of resources, assessment tools) are constant, with reasonable time allocations.	1. Lesson plans or units are based on activities or resources rather than focused on objectives. 2. Progression and pacing of learning activities are sporadic, thus, time allocations are not always reasonable. 3. Lesson plans show little evidence of formative assessment.	1. There is little evidence of monitored learning or use of appropriate instructional resources. 2. No evidence of planning and design of formative assessment or progress monitoring.
Requires students to understand and demonstrate skills and competencies	1. Assessment tasks are embedded within the lesson and/or require a performance component for students to demonstrate what they know and are able to do.	1. Demonstration of skills and competencies is developed in advance of instruction and is aligned with the lesson objective. 2. Lesson objectives are written in student-friendly language and students understand what they are expected to know and be able to do by the end of each lesson.	1. Although students are aware of the posted lesson objective, they rely on teacher direction to focus them on what they are expected to know and be able to do. 2. The demonstration of skills and competencies is minimally developed and/or may be loosely connected to the lesson objective.	1. Lesson objectives are not known to students and students do not know what they are expected to know and be able to do.

BPS Instructional Personnel Performance Appraisal System

2011 - 2014

Our Mission is to Serve Every Student with Excellence as the Standard!

Dimension 2: *Learning Environment*

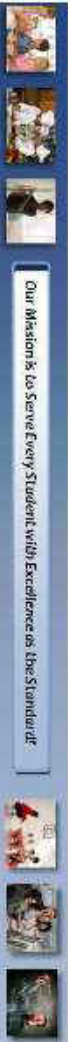| Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Creates and maintains a safe and organized learning environment	1. Implements and routinely reinforces, and students assume responsibility for school behavioral rules and regulations and established procedures to ensure student safety needs are consistently met. 2. Establishes and maintains a comfortable, safe, and inviting learning environment that is organized to facilitate a focus on learning.	1. Implements and routinely reinforces school behavioral rules and regulations and established procedures to ensure student safety needs are consistently met. 2. Implements classroom routines.	1. Does not consistently reinforce school behavioral rules and regulations and/or is not clear about established procedures in emergency situations. 2. Lacks an appropriate level of attention to the supervision of students inside and outside the classroom.	1. Allows for conflict in the classroom environment. 2. Is non-compliant and/or negligent with regard to supervision of students inside and outside the classroom setting.
Promotes a flexible, inclusive, collaborative, and student-centered learning environment	1. Fosters and creates, through collaboration among students, a stimulating and inclusive learning environment; maximizes potential for equal learning opportunities for every student. 2. Teachers create an environment where students assume responsibility and hold themselves accountable for their learning. 3. Instructional time is protected by students' adherence to classroom procedures.	1. Fosters a stimulating and inclusive learning environment. 2. Teachers assume responsibility for student learning. 3. Classroom procedures have been established and are followed by students.	1. Allows for too many students to "fall through the cracks". 2. Does not individualize student attainment of standards. 3. In the absence of known procedures, students spend time waiting for teacher re-direction in order to engage in independent work or cooperative group work.	1. Fails to use procedures for independent and/or cooperative group work. 2. Engages in learning solely on textbook or curriculum sequence with little or no individualized instruction for students to meet specific learning needs.
Allocates and manages time, space, resources	1. Arranges physical resources proactively to assure accessibility to all students, taking into account individual needs of students. 2. Learning experiences, activities and physical spaces are highly organized and efficiently facilitated by both the teacher and students who each assume responsibility for maximizing time for learning.	1. Use of physical resources and space (furniture, technology, and learning stations) contributes to all students' ability to access learning. 2. Learning experiences, activities and physical spaces are organized and facilitated by the teacher in such a way that students have adequate time for learning.	1. Use of physical resources and space limits student engagement and does not promote learning activities' accessibility to all students resulting in loss of learning time for some students. 2. Failure to utilize and maintain facilities/equipment results in loss of learning opportunities.	1. Poor use of physical resources and space is evident and/or learning is not accessible to some students resulting in considerable "down-time." 2. Teaching/ learning time wasted.

Dimension 2: Learning Environment

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Manages student conduct	1. Students contribute to designing the classroom rules and standards of conduct upheld by all members of the classroom. 2. Consequences for inappropriate behavior are reasonable, clear, and consistently applied. 3. Monitoring by the teacher of events in the classroom is subtle and proactive. 4. Students monitor their own and their peers' behavior, correcting one another respectfully. 5. The focus of discipline is to maximize student learning time and students show a respect for the rights of other students to learn.	1. Standards for student conduct are evident, clear, and the physical environment supports a nurturing learning environment. 2. Teacher is alert to student behavior at all times and manages student conduct. 3. Consequences for inappropriate behavior are reasonable, clear, and consistently applied.	1. Most students seem to understand rules of conduct. 2. Teacher is generally aware but misses the activities of some students. 3. Rules and consequences for inappropriate behavior are inconsistently applied. 4. Management techniques do not consistently assist students in developing individual responsibility for learning and behavior.	1. Classroom rules and consequences have not been established, communicated and/or enforced. 2. Teacher is unaware of what students are doing. 3. Teacher spends too much time dealing with student behavior and disciplinary consequences. 4. Response to student behavior is inconsistent or negative and counterproductive.
Conveys high expectations, respects students' cultural backgrounds and accommodates diverse needs of students	1. Establishes a culture of excellence that stretches achievement for all students. 2. Student to student interactions are highly respectful and are supportive of one another's individuality and developmental levels. 3. Every student receives verbal and nonverbal indications that he/she is valued and respected.	1. Encourages a culture in the classroom that challenges all students to continuously improve. 2. Teacher seeks to create an environment where student interactions are respectful. 3. Students receive verbal and nonverbal indications that he/she is valued and respected.	1. Supports student improvement efforts suitable for most students. 2. Teacher reinforces student to student interactions that are appropriate, but there may be an occasional display of insensitivity to cultural or developmental differences among students.	1. Reinforces low level learning expectations. 2. Teacher allows displays of insensitivity to cultural or developmental differences among students.
Models and teaches clear, acceptable communication skills	1. Directions, procedures, and feedback are clear to students and anticipate possible misunderstanding. 2. Teacher's spoken and written language conforms to standard English and contains well chosen vocabulary that enriches the lesson. 3. Teacher finds opportunities to extend student vocabulary. 4. Teacher/student interactions serve as a model of respectful rapport.	1. Directions, procedures, and feedback are clear to students. 2. Teacher's spoken and written language conforms to standard English. 3. Teacher/student interactions serve as a model of respectful rapport.	1. Directions, procedures, and feedback are clarified to students after initial student confusion. 2. Teacher's spoken language is audible and when written, is legible. 3. Teacher/student interactions are cordial, but may not transfer to positive student/teacher relations.	1. Directions, procedures, and feedback are confusing to students. 2. Teacher's spoken language is inaudible to some students. 3. Writing is illegible or contains errors in grammar or syntax. 4. Teacher/student interactions reflect a lack of rapport.

Dimension 2: Learning Environment

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Maintains a climate of inquiry	1. Facilitates a collaborative classroom culture that promotes student leadership in learning and in applying a variety of problem solving, inquiry-based strategies. 2. Engages students in problem solving inquiry-based activities through the use of high level questioning techniques, discovery learning, shared inquiry/Socratic discussions. 3. Problem solving culminates in student generated real-world applications.	1. Engages students in problem solving inquiry-based activities through the use of high level questioning. 2. Student participation and responses indicate individual understanding of content and/or concepts.	1. Includes some problem solving and inquiry-based activities. 2. Some students are not involved or adequately prepared, do not understand or are not engaged in the processes. 3. A limited number of effective inquiry-based strategies are used.	1. Instruction does not regularly include problem solving or inquiry-based activities.
Integrates appropriate technologies	1. Creates a rich learning environment in which students regularly engage in activities that would have been difficult to achieve without technology. 2. Students are actively engaged in using available technology as a tool rather than passively receiving information from the technology. 3. Students engage in ongoing meta-cognitive activities and collaborate with peers and/or experts on educational projects.	1. Uses technology to construct teaching materials and learning activities. 2. Students are actively engaged in using technology as a tool rather than passively receiving information from the technology. 3. Teaches students to use available computers and other forms of technology to create products.	1. Demonstrates awareness of and models acceptable use of technology; collaborates with other educational professionals, parents and students, as appropriate. 2. Teacher uses technology to deliver curriculum content to students. 3. Students use technology passively for drill and practice and computer-based training.	1. Technology to promote organizational efficiency and support instruction is not evident.

Dimension 3: Instructional Delivery & Facilitation

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Utilizes knowledge of the subject to enrich students' understanding, identifies gaps in students' learning, modifies instruction to respond to student misconceptions	1. Instruction is based on rich content knowledge that is accurate, current and consistent with sound practices of the discipline. 2. All students are engaged in learning experiences, discussions, questioning, and demonstrations of learning. 3. Uses a variety of checks for understanding during strategic points in the lesson to identify learning gaps and adjust instruction. 4. Anticipates problems and uses multiple intervention strategies to assist student understanding and performance.	1. Instruction is based on content knowledge that is accurate and current. 2. Recognizes problems and makes adjustments during the lesson in response to student understanding and performance. 3. The teacher engages students, models instruction, guides practice and provides feedback, and promotes student demonstration of learning.	1. Instruction is based on a general understanding of content knowledge that is accurate but fails to reflect the most current knowledge of the discipline. 2. Teacher does not check for understanding or recognize gaps in learning during the lesson. 3. Teacher adjusts future lessons based on student response data at the conclusion of the lesson.	1. Instruction references inaccurate, outdated content knowledge. 2. Makes ineffective use of instructional materials. 3. Teacher does not use checks for understanding during the lesson. 4. No adjustments are made to the lesson sequence or pacing when students are confused.
Employs higher order questions	1. Strategic teacher questioning requires students to think critically, problem solve, and defend or justify their answers. 2. Much of the questioning involves analysis and synthesis. 3. Students contribute to the discussion by commenting on or adding to answers given by their classmates. 4. Ideally teachers can take a back seat & discussions can be student-led.	1. Most teacher questions encourage thoughtful and extended responses while some are of lower level and can be answered with few words. 2. Teacher encourages all students to participate.	1. Few teacher questioning strategies invite quality responses. 2. Teacher is quick to provide the correct answers and students are not given the opportunity to think things through. 3. Only a few students respond.	1. Questioning is knowledge level and is solely teacher-driven. 2. Students appear apathetic and the teacher provides the answers without any effort on their part.
Applies varied instructional strategies and resources	1. Uses multiple resources and an extensive repertoire of instructional strategies that engage and challenge all students and support instructional outcomes. 2. Teacher stays abreast of current research through professional organizations, universities, local resources, and through the use of technology. 3. Teacher incorporates innovative ideas and materials.	1. Shows awareness of resources for students available through the school and district and some familiarity with resources external to the school. 2. Teacher uses multiple instructional strategies to maintain focus, engage students and support instructional outcomes for all students.	1. Makes use of a limited number of instructional strategies to support student understanding. 2. Teacher is aware of and uses resources available through the school and district but does not use resources available more broadly.	1. Develops lessons that lack instructional focus and do not engage students and support instructional outcomes. 2. Teacher does not use resources available through the school and district to enhance content. 3. Classroom time is often filled with activities that merely consume time and do not contribute to instructional outcomes.

Dimension 3: Instructional Delivery & Facilitation

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Delivers engaging, challenging and relevant lessons	1. Lessons (and units) are designed to maximize productive time. 2. High levels of rigor and relevance challenge students to be intellectually engaged throughout. 3. Teacher clearly demonstrates and articulates how content relates and applies to instructional activities, life, work and community. 4. Students can explain how the lesson relates to other disciplines and to themselves.	1. Lessons (and units) are designed to provide students with rigorous curriculum that is intellectually engaging. 2. Lessons often demonstrate the subject's importance and relation to life situations, and students can state how the subject being studied is relevant to their own learning.	1. Lessons (and units) spend the majority of the class time on knowledge level tasks that may require only some intellectual engagement. 2. Lessons do not consistently demonstrate an understanding of the subject's relationship to life situations. 3. Whether or not they find it interesting, the students follow the teacher's directions and do what is asked of them.	1. Lessons (and units) are too easy for the majority of students and class time is spent on busy work with low intellectual engagement. 2. Lessons do not help the students make the connection between the instructional goals and what they are learning. 3. Presentation of the lesson is not interesting and the students fail to apply the lesson to themselves.
Differentiates instruction	1. Differentiates instruction efficiently and appropriately by content, process and product to address the unique learning differences of students. 2. Both daily instruction and unit design clearly address the wide range of learning styles and abilities present in the classroom. 3. Teacher provides opportunities for student choice in demonstrating mastery of learning goals.	1. Provides evidence of incorporating various differentiated instructional strategies (e.g. ability grouping or compacting of lessons) to meet the needs of students with varying learning styles and abilities.	1. Provides little evidence of incorporating various differentiated instruction. 2. Recognizes the need for differentiation but requires support to differentiate instruction for students with varying learning styles and abilities.	1. Does not differentiate instruction to address the needs of students with varying learning styles and abilities.
Provides immediate and specific feedback to students	1. Provides timely, deliberate, and consistent feedback during guided practice, discussion and major activities. 2. Students use feedback in their learning to improve their performance.	1. Provides timely and consistent feedback. • Students are clear on what areas need improvement for the next assigned task.	1. Provides feedback at the end of the unit. 2. Feedback does not provide specific information for students to make error corrections or is not timely enough to improve performance.	1. Provides inadequate or no feedback. Feedback provided is not provided in a timely manner.

BPS Instructional Personnel Performance Appraisal System

2011 - 2014

Our Mission is to Serve Every Student with Excellence as the Standard?

Dimension 4: Assessment

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Analyzes and applies data from multiple measures to diagnose students' learning needs, inform instruction, and monitor progress.	1. Consciously plans and uses pre-assessments, formative and summative assessments in a continuous improvement cycle based on student achievement. 2. All assessment criteria and standards are clear to students and contribute to their learning. 3. Students and the teacher analyze data results from multiple assessments to make decisions about their progress and develop appropriate interventions relative to their needs. 4. Progress monitoring data is regularly used to provide specific feedback to individual students and analyze teaching results. 5. Students make changes in their learning paths based on the analysis of assessment data.	1. Uses pre-assessment, formative assessment, and summative assessment data strategically to set and adjust learning goals based on student needs. 2. All assessment criteria and standards are clear. 3. Gathers data during instruction and uses data to inform instruction. 4. Analyzes multiple assessment data regularly and adjusts instruction based on analysis. 5. Assessment methodologies may be differentiated for groups of students.	1. Regularly uses summative assessment data to determine that students have achieved instructional outcomes. 2. Assessment criteria and standards have been developed, but may not be clear. 3. Use of formative assessment data is rudimentary, assessing some instructional outcomes. 4. Little intervention or differentiation is evident.	1. Assessment procedures are not congruent with instructional outcomes. 2. No criteria or standards for assessment are evident. 3. Teacher does not incorporate assessment data to monitor and adjust instruction or to differentiate assessment for individuals. 4. Assessment measures are not used to diagnose student learning or to develop interventions to improve instruction.
Designs and uses formative and summative assessments that lead to mastery	1. Formative and summative assessments are developed as a part of unit design prior to instruction. 2. Performance results are used to build a shared understanding with students of what they should know and be able to do as a result of instruction. 3. Formative assessment tasks are embedded within the lesson and/or require a performance component for students to demonstrate what they know and are able to do. 4. The teacher evaluates the results from the formative assessment tasks and uses that information to re-teach or improve future lessons.	1. Some formative assessments are used to measure achievement in both content and higher-level thinking. 2. Summative assessments are given at the end of lessons or units and provide varied options for students to demonstrate what they know and are able to do. 3. Results of the assessments are the basis for instructional planning and used to evaluate overall student progress toward meeting the standards.	1. Formative and summative assessments are developed but may not measure beyond understanding of content or may not use the tools to help students understand individual achievement. 2. Assessments after instruction are not used for instructional planning but evaluate student progress toward meeting standards. 3. Some summative assessment content material is considered prior to instructional lesson and unit design.	1. No formative assessment is given. 2. A single type of classroom assessment is used that may or may not be aligned with the curriculum and is used only to justify student grades. 3. Lessons and units are not designed prior to assessments being administered.

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Modifies teacher made assessments to accommodate diversity	1. Efficiently and appropriately differentiates assessments by content, process and/or product to address the unique learning differences of students that have a wide range of learning styles and abilities. 2. Students are aware of their learning styles and given choices to demonstrate their learning.	1. Appropriately differentiates assessments to address the unique learning differences of students that have a wide range of learning styles and abilities. 2. Provides a variety of assessments to meet the needs of students. 3. Students may have choices in their assessments.	1. Little evidence is provided that assessments are differentiated to meet the needs of students' learning styles or abilities.	1. Assessments are "one size fits all."
Communicates assessment data to students and parents	1. Assessment criteria, due dates, and grading methodology are clearly communicated in Edline in such a way that promotes student learning. 2. Teacher is proactive in creating and maintaining frequent and effective two-way communication with students and parents. 3. Grades and progress reports are completed and updated regularly to allow students to monitor their own proficiency.	1. Grades and progress reports are completed and updated regularly in Edline to allow students to check their own progress. 2. Teacher has designed a system for documenting and reporting student learning. This is shared with students and parents in advance. 3. Teachers lead conferences with parents and students to share assessment data with parents.	1. Grading and reporting of student learning is completed at the end of the unit. 2. Teacher does not consistently communicate progress with students and parents in a timely fashion. 3. Grades and progress reports are completed per the school schedule.	1. Grading is inconsistent and may be subjective. 2. Students and parents do not know what the cumulative grade will be until the official report comes out. 3. Grades and progress reports are not completed per the school schedule.

Dimension 5: Professional Responsibilities & Ethical Conduct

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Demonstrates punctuality, attendance, and timely completion of records and reports	1. Complies fully with all school rules, policies and procedures regarding punctuality and attendance. 2. Has developed a system for maintaining information on students, lesson plans, reports and other data which is efficient, current and useful.	1. Complies with school rules, policies and procedures. 2. The system for maintaining information on students, lesson plans, reports and other data is effective.	1. Inconsistently complies with school rules, policies and procedures. 2. The system for maintaining student information, lesson plans, reports and other data is rudimentary and only partially effective.	1. Does not comply with school rules, policies and procedures. 2. There is no system of record keeping, lesson plans, records or other data or the system is in disarray.
Performs assigned duties and complies with policies, procedures, programs, and the Code of Ethics	1. Fully understands, abides by and supports the legal and professional responsibilities pertaining to education. 2. The teacher complies with district and school policies, rules, procedures and the Code of Ethics. 3. The teacher acts proactively in fostering compliance amongst colleagues with standards of excellence and demonstrates leadership in exhibiting a high ethical standard of decorum throughout the school setting.	1. Understands and abides by the legal and professional responsibilities pertaining to education. 2. The teacher complies fully with district and school policies, rules, procedures and the Code of Ethics. 3. The teacher models professional and ethical standards of decorum throughout the school setting.	1. Demonstrates a limited understanding through inconsistent behaviors. 2. Inconsistently complies with district and school policies, rules, procedures and/or the Code of Ethics. 3. The teacher exhibits minimal deference to professional and ethical standards of decorum throughout the school setting.	1. Disregards or has no awareness of legal and professional responsibilities pertaining to education. 2. The teacher does not comply with district and school policies, rules, procedures and/or the Code of Ethics. 3. The teacher's behavior is below acceptable ethical standards of decorum throughout the school setting.
Demonstrates professionalism	1. Displays conduct based on the highest professional standards. 2. Acts honestly and with integrity when dealing with students, parents, colleagues and/or the community. 3. The teacher demonstrates a pattern of participation in district and school initiatives, contributes to decision-making processes, serves on subcommittees and/or disseminates information when appropriate.	1. Deals with students, parents and/or colleagues with honesty and integrity. 2. The teacher participates in and implements school and district initiatives. 3. The teacher clearly implements the evidence.	1. Uses poor judgment when dealing with students, parents and/or colleagues. 2. The teacher implements most decisions made at the school and district levels.	1. Displays unethical or dishonest conduct when dealing with students, parents and/or colleagues. 2. The teacher does not implement decisions made at school or district levels.

149

Dimension 5: Professional Responsibilities & Ethical Conduct

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Initiates professional communication with appropriate stakeholders	1. Promotes a two-way partnership between school and home. 2. Maintains regular communication between classroom and student's family regarding the instructional program and their child's progress. 3. Efforts to connect families and communities to the instructional program are frequent and successful. 4. These may include, but are not limited to, newsletters, family nights, websites, electronic communication, and phone calls. 5. Students and their families understand what they are expected to know and be able to do.	1. Provides regular information about the instructional program. 2. Is available as needed to respond to parental concerns. 3. Efforts to connect families, school, and communities are successful.	1. Participates in school's activities for parent communication but offers limited information regarding the instructional program. 2. Parents' concerns are only partially addressed with little or no follow up. 3. Makes minimal attempts to connect families and communities to the instructional program.	1. Provides little or no information about the instructional program. 2. Does not respond or responds inappropriately to parental concerns. 3. Makes no attempt to connect families and communities to the instructional program.
Applies technology to organize and communicate assessment information	1. Uses technology to organize, monitor, and communicate student learning and assessment information to appropriate stakeholders. 2. Integrates authentic tasks and provides increased opportunities for independent learning for all students through the use of technology tools.	1. Uses technology to communicate student learning and assessment information to appropriate stakeholders in a timely manner.	1. Struggles to use technology to communicate student learning and assessment information to appropriate stakeholders in a timely manner.	1. Does not use technology to communicate student learning and assessment information to appropriate stakeholders.

Dimension 6: Relationship with Students

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Demonstrates Knowledge of Students	1. Teacher recognizes the value of understanding students' interests and cultural heritage and displays this knowledge by using interest surveys, invitational relationship-building activities, and equitable procedures. 2. Teacher displays understanding of individual students' skill, knowledge, and language proficiency and has a strategy for maintaining such information. 3. Teacher possesses information about each student's learning and medical needs, collecting such information from a variety of sources. 4. Teacher displays extensive understanding of how students learn and applies this knowledge to develop a positive relationship with individual students.	1. Teacher recognizes the value of understanding students' interests and cultural heritage and displays this knowledge for groups of students. 2. Teacher recognizes the value of understanding students' skills, knowledge, and language proficiency and displays this knowledge for groups of students. 3. Teacher is aware of students' special learning and medical needs. 4. Teacher's knowledge of how students learn is accurate and current. 5. Teacher applies this knowledge in a positive relationship to the class as a whole and to groups of students.	1. Teacher recognizes the value of understanding students' interests and cultural heritage but displays this knowledge only to the class as a whole. 2. Teacher recognizes the value of understanding students' skills, knowledge, and language proficiency but displays this knowledge only for the class as a whole. 3. Teacher displays awareness of the importance of knowing students' special learning or medical needs, but such knowledge may be incomplete or inaccurate. 4. Teacher recognizes the value of knowing how students learn, but this knowledge is limited or outdated. 5. Students appear to lack a positive relationship with the teacher.	1. Teacher displays little or no knowledge of students' interests or cultural heritage and does not indicate that such knowledge is valuable. 2. Teacher displays little or no knowledge of students' skills, knowledge, and language proficiency and does not indicate that such knowledge is valuable. 3. Teacher displays little or no understanding of students' special learning or medical needs or why such knowledge is important to a positive relationship. 4. Teacher does not seek information about how students learn and appears to lack positive relationships with them.
Builds Relationships Through Instructional Interactions	1. Teacher allows and encourages all students to be part of class discussions and interactions. 2. Teacher brings student interests into the content.	1. Teacher allows and encourages most students to be part of class discussions and interactions. 2. Teacher brings some student interests into the content.	1. Teacher allows some students to be involved in class discussions and interactions. 2. There is little evidence of the teacher's bringing student interests into the content.	1. Teacher does not engage students in class discussions and interactions. 2. Awareness of student interests is not evident in content delivery.

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Creates a Positive Environment of Respect and Rapport	1. Teacher interactions with students reflect genuine respect and caring for individuals as well as groups of students. 2. Students appear to trust the teacher with sensitive information. 3. Students demonstrate genuine caring for one another and monitor one another's treatment of peers, correcting classmates respectfully when needed.	1. Teacher interactions with students reflect the same amount of respect given to and expected from every student. 2. Teacher-student interactions are friendly and demonstrate general caring and respect. Such interactions are appropriate to the age and cultures of the students.	1. Teacher –student interactions are generally appropriate but may reflect occasional inconsistencies, favoritism, or disregard for students' cultures. 2. Students exhibit only minimal respect for the teacher. 3. Students are disrespectful to one another.	1. Teacher interaction with at least some students is negative, demeaning, sarcastic, or inappropriate to the age or culture of the students. 2. Students exhibit disrespect for the teacher. 3. Student interactions are characterized by conflict, sarcasm, or put-downs.

Dimension 7: Relationship with Parents and Community: Building Relationships with Parents

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Establish contact with parents early in the school year	1. Use of multiple methods to reach out and welcome all parents within the first two weeks of the school year and elicits feedback from parents about their child.	1. Welcome parents with a personal contact early in the school year (phone calls, welcome letter, home visits).	1. Attempts to contact parents early in the school year, but may not have a consistent plan and/or needs guidance to do so effectively.	1. Makes no effort to contact parents early in the school year.
Establish a relationship of mutual trust and respect with parents	1. Consistently considers family background and cultural diversity when working with parents. 2. Demonstrates sensitivity, respect, and understanding of families to build positive relationships with all families.	1. Demonstrates awareness of family background and cultural diversity. 2. Demonstrates sensitivity, respect, and understanding of families to build a positive relationship.	1. Lacks awareness of cultural diversity of families. 2. Needs guidance to learn effective strategies for building positive relationships with parents. 3. Does not establish a relationship of trust and mutual respect with parents.	1. Does not maintain a family friendly learning environment. 2. Does not consistently reach out to invite parents to collaborate and participate in learning in innovative ways.
Maintain a family friendly learning environment	1. Consistently reaches out to invite parents to collaborate and participate in learning in innovative ways.	1. Maintains a family friendly environment that encourages engagement (inviting climate, opportunities to volunteer in the classroom, and encourage collaboration with parents).	1. Attempts to maintain a family friendly environment that encourages engagement, but needs guidance to do so effectively.	1. Does not maintain a family friendly environment that encourages engagement.
Maintains two-way communication with parents	1. Multiple methods of two-way communication are used regularly to maintain communication with all parents (See examples in proficient). 2. Consistently implements strategies to maximize participation of all parents in parent teacher conferences (See examples in proficient). 3. Student-led conferences are evident when appropriate.	1. Two-way communication is used to keep parents informed (agendas, Edline, email, phone calls). 2. Implements strategies to maximize participation in parent-teacher conferences (offering multiple locations, convenient times, opportunities for student-led conferences, follow up with parents who do not initially reply).	1. Inconsistent in communication with parents. 2. May need strategies to implement two-way communication. 3. Minimal effort to encourage parent participation in parent-teacher conferences. 4. Does not implement multiple efforts to reach parents and may only conferences with parents when there is a problem.	1. Does not maintain communication with parents. 2. Makes no effort to encourage parent participation in parent-teacher conferences.
Responsive to parent needs	1. Consistently accessible to parents, actively listens to concerns, immediately addresses issues, elicits feedback regularly, and engages in open dialogue.	1. Accessible to parents, listens to concerns, addresses issues in a timely manner and engages in open dialogue.	1. Limited accessibility to parents and does not consistently respond to parents in a timely manner.	1. Not accessible to parents. 2. Not responsive to parent concerns.

Dimension 7: Relationship with Parents and Community: Engaging Parents in Learning

Elements	Distinguished (3 pts)	Proficient (2 pts)	Professional Support Needed (1 pt)	Unsatisfactory (0 pts)
Promotes parent understanding of academic standards and expectations	1. Provides all parents with information, specific to their child, about standards and expectations in a format that parents can understand.	1. Provides parents with information about academic standards and expectations in a format that parents can understand (Open House, handouts, post on website, parent conferences).	1. Limited information provided to parents about academic standards and expectations. 2. May need assistance in providing it in a format that parents can understand.	1. Does not provide parents with information about academic standards and expectations.
Reports to parents about student performance and academic progress	1. Consistently interprets and explains individual students' academic progress (assessment results, accomplishments, and needs for improvement) in a manner that provides parents with an ongoing picture of children's performance.	1. Interprets and explains students' academic progress (assessment results, accomplishments, and needs for improvement) in a manner that provides parents with an accurate picture of children's performance.	1. Limited information shared with parents on student progress. 2. Does not provide an accurate picture of children's performance in a timely manner.	1. Does not provide parents with information about students' performance and academic progress.
Provides opportunities to assist with student learning at home	1. Consistently provides all parents will multiple strategies and opportunities to assist with student learning (See examples for proficient). 2. Offers individualized support	1. Provides strategies and opportunities for parents to assist with student learning at home (parent workshops, meetings, suggestions for home activities, access to resources).	1. Provides limited strategies and opportunities for parents to assist with student learning.	1. Does not provide or promote opportunities for parents to assist with student learning at home.
Provides information about school and community resources to parents	1. Consistently provides school and community event information in newsletters, email, or Edline about available resources. 2. Consistently provides information to parents about community events and resources.	1. Provides school and community event information in newsletters, email, or Edline about available resources. 2. Provides timely information to parents about community events and resources.	1. Limited information is provided to parents in newsletters, email, or Edline about available resources. 2. Limited information is provided to parents about community events and resources.	1. Does not provide information to parents in newsletters, email, or Edline about available resources. 2. Does not provide information to parents about community events and resources.
Provides meaningful connections with learning and community	1. Consistently integrates everyday activities to real-life situations 2. Consistently integrates community service opportunities with enhanced learning	1. Integrates everyday activities to real-life situations. 2. Integrates community service opportunities with learning	1. Limited integration of everyday activities to real-life situations 2. Limited integration of community service opportunities with learning	• Does not integrate everyday activities to real-life situations. • Does not integrate community service opportunities with learning.

LIST OF REFERENCES

Abrams, J. (2009). *Having hard conversations*. Thousand Oaks, CA: Corwin Press.

American Recovery and Reinvestment Act of 2009, Public Law111-5 (H.R. 1). (2009). 123 Stat. 115 as of May 21, 2012. Retrieved from http://www2.ed.gov/policy/gen/leg/recovery/statutory/stabilization-fund.pdf

Baker, B. D., Oluwole, J. O., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the race-to-the-top-era. *Educational Policy Archives, 21*(5), 1-71.

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. *Economic policy institute briefing paper 278*. Retrieved from http://epi.3.cdn.net/b9667271ee6c154195_t9m6iij8k.pdf

Barth, R. S. (2006) Improving relationships within the schoolhouse. *Educational Leadership, 63*(6), 8-13.

Bartolo, D. B. (2012). *Closing the teaching gap: Coaching for instructional leaders*. Thousand Oaks, CA: Corwin Press.

Bearwald, R. R. (2011). It's about the questions. *Educational Leadership, 69*(2), 74-77.

Bergquist, C. C. (2012). Florida Department of Education. *Professional development system evaluation protocol: Third cycle/second year technical report (school year 2011-2012)*. Tallahassee, FL: Evaluation Systems Design.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15. Retrieved from set.stanford.edu/sites/default/files/files/documents/publications/Borko-Professional%20Development%20and%20Teacher%20Learning.p

Bowman, C. L., & McCormick, S. (2000). Comparison of peer coaching versus traditional supervision effects. *The Journal of Educational Research, 93*, 256-261.

Brevard Public Schools. (2011). *Instructional personnel performance appraisal system*. Retrieved from http://benefits.brevard.k12.fl.us/HR/comp/pas/pinfo/IPPAS.pdf

Brevard's Effective Strategies for Teaching [B.E.S.T.] (2009). Retrieved from http://best.brevardschools.org/best/default.aspx

Briggs, D. & Domingue, B. (2011). *Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District teachers by* the Los Angeles Times. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/due-diligence.

Brown, C. S. (2010). *Implementing preschool curriculum: Mentoring and coaching as key components to teacher professional development.* (Doctoral dissertation). Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/7535078 88?accountid=10003

Bruce, C. D., & Ross, J. A. (2008). A model for increasing reform implementation and teacher efficacy: Teacher peer coaching in grades 3 and 6 in mathematics. *Canadian Journal of Education 31*(2), 346-370.

Burke, W., Marx, G. E., & Lowenstein, E. (2012). Leading, leadership, and learning: Exploring new contexts for leadership development in emerging school environments. *Planning and Changing, 33*, 113-126.

Chetty, R., Friedman, J. N., & Rockoff, J. (2012). Great teaching: Measuring its effects on students' future earnings. *Great Teaching: Education Next, 12*(3). Retrieved from http://educationnext.org/great-teaching/

Collet, V. S. (2012). The gradual increase of responsibility model: Coaching for teacher change. *Literacy Research and Instruction, 51*(1), 27-47.

Coulter, M. P. (2013). A qualitative study of teacher and principal perceptions of Washington State teacher evaluation instruments: Danielson, Marzano, and CEL 5D+. Retrieved from ProQuest Dissertations and Theses (1426646878).

Cross, J. (2012). Lasting legacy for achievement: Colorado coaches boost teacher effectiveness with a rubric for professional learning. *JSD: The Learning Forward Journal, 33*(5), 36-42.

Curtis, R. (2013). *Finding a new way: Leveraging teacher leadership to meet unprecedented demands.* Washington, DC: Aspen Institute. Retrieved from http://www.aspendrl.org/portal/browse/DocumentDetail?documentId=1574&dow nload

Cushman, C. S. (2013). *The effects of instructional coaching on collective teacher efficacy and school-level effect in Tennessee public middle schools.* (Doctoral dissertation). Retrieved from

http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/1499258 713?accountid=10003

Danielson, C. (2006). *Teacher leadership that strengthens professional practice.* Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2007*). Enhancing professional practice: The framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future.* New York: Teachers College Press.

Darling-Hammond, L. (2013). When teachers support and evaluate their peers. *Educational Leadership, 71*(2), 24 – 29.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Oxford, OH: NSDC.

Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*(3), 285-328. Retrieved from http: www.jstor.org/stable/1170367

Day, C., & Sachs, J. (2004). *International handbook on the continuing professional development of teachers.* Maidenhead: Open University Press.

Demonte, J. (2013, July). High quality professional development for teachers: Supporting teacher training to improve student learning. *Center for American Progress.* Retrieved from http: www.americanprogress.org/issues/education/report/2013/07/15/69592/high-quality-professional-development-for-teachers.pdf

DuBois, L. (2012). Principals' leadership and leadership principles: Creating professional development to help today's principals excel at leading teachers and schools. *Peabody Reflector*. Retrieved from http://www.vanderbilt.edu/magazines/peabodyreflector/2012/07/principals-leadership-an

Duff, V., & Islas, M. R. (2013). Partners in learning: Teacher leaders drive instructional excellence. *JSD: The Learning Forward Journal, 34(*6), 10-14.

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

Eisenbach, R., & Curry, R. R. (1999). Ideas, observations, and inquiry: The emotional reaction to classroom visitation and peer coaching. *Journal of Management Education, 23*(4), 416-427. Retrieved from http://jme.sagepub.com.ezproxy.net.ucf.edu/content/23/4/416.full.pdf

Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study. (2013). Seattle, WA: The Bill and Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf

Florida Department of Education. (2010a). *Educator recruitment, development, & retention: Professional development standards*. Retrieved from http://www.fldoe.org/profdev/pdstandards.asp

Florida Department of Education. (2010b). *Florida Educator Accomplished Practices FEAPS)*. State Board of Education Rule 6A-5.065. Retrieved from http://www.fldoe.org.

Florida Department of Education (2011a). *Educator Quality Update on Teacher Evaluations*. Retrieved from http://www.fldoe.org/profdev/pdf/OverviewFloridaTeacherEvaluationSystem.pdf

Florida Department of Education (2011b). *Florida's value-added model*. [Powerpoint].Retrieved from www.fldoe.org/committees/pdf/Presentation Value-addedModel.pdf.

Florida Department of Education (2012). *Professional development system evaluation protocol, third cycle/second year technical report*. Retrieved from http://www.fldoe.org/profdev/pdf/ytcdrs.pdf

Florida Government (2011). *Governor Scott signs student success act*. Retrieved from http://www.flgov.com/2011/03/24/governor-scott-signs-student-success-act/

Florida State Statute 1012.34. (2012).

Florida State Statute 1012.98, (2003).

Florida's value-added model: Overview of the model to measure student learning growth on FCAT as developed by the student growth implementation committee (2013). Retrieved from http://www.fldoe.org/committees/pdf/PresentationValue-addedModel.pdf

Fritz, R. S. (2014). *An analysis of the difference among assessed instructional practices, value-added measures, and learning gains of secondary reading teachers.* (Unpublished doctoral dissertation). University of Central Florida Orlando, FL.

Fullan, M., & Knight, J. (2011). Coaches as system leaders. *Educational Leadership, 69*(2), 50-53.

Gallagher-Fishbaugh, S. (2011, January 29). Effective teacher evaluation programs can improve schools. *The Salt Lake Tribune.* Retrieved from archive.sltrib.com/printfriendly.php?id=510892239type=cmsid

Gallimore, R., Ermeling, B. A., Saunders, W. M., & Goldenberg, C. (2009). Moving the learning of teaching close to practice: Teacher education implications of school-based inquiry teams. *The Elementary School Journal, 109*(5), 537-553.

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W. . . Silverberg, M. (2008, September). *The impact of two professional development interventions on early reading instruction and achievement.* Retrieved from http://ies.ed.gov/ncee/pdf/20084030.pdf

Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. (2012, January). Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D., & Whitehurst, G. J. (2010, November). Evaluating teachers: the important role of value-added. *Brookings Report.* Retrieved from http://www.brookings.edu/research/reports.2010/11/17-evaluating-teachers

Griffin, L. (2013). Charlotte Danielson on teacher evaluation and quality. *School Administrator. 70(1)*, 27-31.

Growth models: Issues and advice from the states (2012, July). Institute of Education Sciences (IES) Statewide Longitudinal Data Systems (SLDS) Grant Program. Retrieved from nces.ed.gov/programs/slds/pdf.guide_growth-model.pdf

Gruska, G. F. (2000). W. E. Deming in the new millennium. *Total Quality Management, 11*(4-6), 9. Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/219800405?accountid=10003

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership, 59*, 45 – 51.

Hall, C. (2013). Building a culture of growth and evaluation in schools. *Independent School, 73*(1), 88-93. Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct =true&db=eft&AN=90400179&site=ehost-live

Hall, P., & Simeral, A. (2008). *Building teachers' capacity for success: A collaborative approach for coaches and school leaders.* Alexandria, VA: ASCD.

Harris, D. N. (2010, October). Value-added measures of education performance: Clearing away the smoke and mirrors. *Pace Policy Brief.* Retrieved from http://www.edpolicyinca.org/publications/value-added-measures-education-performance-clearing-away-smoke-and-mirrors

Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014, February). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*(1), 77-112.

Hasbrouck, J. E., & Christen, M. H. (1997, January). Providing peer coaching in inclusive classrooms: A tool for consulting teachers. *Intervention in School & Clinic, 32*(3), 172-178. Retrieved from ds.a.ebscohost.com.ezproxy.net.ucf.edu/eds/delivery?sid=9f4fcf1b-f9a8-4ce8-86c7-cc6b3970c6f8%40sessionmgr4004&vid=15&hid=4211

Haslam, M. B. (2010). Teacher professional development evaluation guide. *National Staff Development Council.* Retrieved from http://learningforward.org/docs/pdf/evaluationguide.pdf?sfvrsn=0

Hawley, W., & Valli, L. (1999). The essentials of professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 127-150). San Francisco, CA: Jossey-Bass.

Hershberg, T., & Roberston-Kraft, C. (2010). Rewards and supports. *School Administrator, 67*(3), 28-31.

Hill, H. C., Kapitula, L., & Umland, K. (2010). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal, 48*(3), 794-831. DOI: 10.3102/0002831210387916

Hill, R., & Rapp, L. (2012). School-based coaches plant seeds of learning. *JSD: The Learning Forward Journal, 33*, 36-39.

Hirsch, S. (2007). NSCD standards and tools help strengthen professional development. *SEDL Letter, XIX*(1). Retrieved from www.sedl.org/pubs/sedl-letter/v19n01/nsdc-standards-tools.html

Holland, S. (2011, November 10). Building inter-rater reliability and accuracy into a system of teacher observation. *Learning Sciences International Blog Updates*. Retrieved from www.learningsciences.com/news/building_inter_rater_reliability_and_accuracy_into_a_system_of_teacher_obse/

Horn, I. S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. *American Education Research Journal, 47*(1), 181-217. doi:10.3.3102/0002831209345158

Ingvarson, L. (2005). *Getting professional development right*. Retrieved from http://research.acer.edu.au/professional_dev/4

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *Teacher Education and Professional Development Commons.* Retrieved from http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1079&context=workingpapers

Johnson, S. M., & Fiarman, S. E. (2012). The potential of peer review. *Educational Leadership, 3*, 20–25.

Joyce, B., & Showers, B. (2002). *Student Achievement through staff development* (3rd ed.), Alexandria, VA: Association for Supervision and Curriculum Development.

Kezar, A. (2001). Investigating organizational fit in a participatory leadership environment. *Journal of Higher Education Policy and Management, 23*, 85-101.

Killion, J., Harrison, C., Bryan, C., & Clifton, H. (2014). Teacher-coach relationships. *Tools for Learning Schools, 17*(4), 1-3.

Killion, J. (2014, August). The quality of professional learning will influence its results. *JSD: The Learning Forward Journal, 35*(4), 60-62.

Kimball, S. M. & Milanowski, A. (2009, February). Examining teacher evaluation validity and leadership decision-making within a standards-based evaluation system. *Education Administration Quarterly, 45*(1), 34-70.

Kline, T. J. B., & Sulsky, L. (2009). Measurement and assessment issues in performance appraisal. *Canadian Psychology, 50*(3), 161-171.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction.* Thousand Oaks, CA: Corwin Press.

Knight, J. (2011). What good coaches do. *Educational Leadership, 69*(2), 18 – 22.

Knowles, M. S. (1980). *The modern practice of adult education: From pedagogy to andagogy 2nd ed.* New York, NY: Association Press.

Koch, M. (2014). *The relationship between peer coaching, collaboration and collegiality, teacher effectiveness and leadership.* (Doctoral dissertation). Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/1525980958?accountid=100003

Kohler, F. W., Crilley, K. M., Shearer, D. D. & Good, G. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90*, 240-250.

Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education, 33*(4), 279-299.

Kupermintz, H. (2003, Fall). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value-added assessment system. *Educational Evaluation and Policy Analysis, 25*(3), 287-298. Retrieved from http://epa.sagepub.com/content/25/3/287.refs.html

Leana, C. R. (2011). The missing link in school reform. *Stanford Social Innovation Review.* Retrieved from http://www.ssireview.org/images/articles/Missing_Link_Cover.pdf.

Lee, O., & Maerten-Rivera, J. (2012). Teacher change in elementary science instruction with English language learners: Results of a multiyear professional development intervention across multiple grades. *Teachers College Record, 114*(8), 1-44.

Lipscomb, S., Chiang, H., & Gill, B. (2012, April). Value-added estimates for phase 1 of the Pennsylvania teacher and principal evaluation pilot (Executive Summary). *Mathematica Policy Research.* Retrieved from http://www.mathematica-mpr.com/~/media/publications/PDFs/education/value-added_PAteachprin_techrpt.pdf

Little, J. W. (2006). *Professional community and professional development in the learning-centered school.* Washington, DC: National Education Association. Retrieved from http://www.nea.org/assets/docs/HE/mf_pdreport.pdf

Lomax, R. G. (2007). *An introduction to statistical concepts for education and the behavioral sciences* (2nd ed.). New York, NY: Routledge.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education 106(*4), 532-575. Retrieved from http://www.jstor.org/stable/1085627

Marietta, G. (2011). *Multiple measures of teacher effectiveness in Hillsborough County Public Schools: The roles of peers and mentors.* Retrieved from http://www.fadss.org/_docs/_content/eet/caseStudyHCPS/MultipleMeasuresHCPS.pdf

Martineau, J. A. (2006). Distorting value-added: The use of longitudinal, vertically scaled student achievement data for growth-based, value-added accountability. *Journal of Educational Behavioral Statistics, 31*(1), 35-62. Retrieved from http://jeb.sagepub.com/content/31/1/35.refs.html

Marzano Art and Science of Teaching Framework: Learning Map. (2011). Retrieved from education.ucf.edu/rtp3/docs/RTP_Marzano_Art%20_Science_of_Teaching_Framework.pdf

Marzano, R. J. (2007). *The art and science of teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19. Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=83173912&site=ehost-live

Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching.* Alexandria, VA: Association for Supervision and Curriculum.

Marzano, R. J., Marzano, J. S., & Pickering, D. J. (2003). *Classroom management that works: Research- based strategies for every teacher.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., & Simms, J. (2013). *Coaching classroom instruction.* Bloomington, IN: Marzano Research Laboratory.

Maslow, V., & Kelley, C. (2012). Does evaluation advance teaching practice? The effects of performance evaluation on teaching quality and system change in large diverse high schools. *Journal of School Leadership, 22*(3), 600-632.

Mela, C. (2013). *An analysis of the school board of Brevard County instruction personnel performance appraisal instrument and the validity of its components.* (Unpublished doctoral dissertation). University of Central Florida, Orlando, FL.

Merriam, S., & Caffarella, R. (1999). *Learning in adulthood: A comprehensive guide, 2*[nd] *ed.* San Francisco, CA: Jossey-Bass.

Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership, 70(*3*),*10-13. Retrieved from http://ezproxy.net.ucf.edu/longin?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=83173911&site=ehost-live

Mourshed, M., Chijioke, C., & Barber, M. (2010, November). *How the world's most improved school systems keep getting better.* Retrieved from http://mckinseyonsociety.com/how-the-worlds-most-improved-school-systems-keep-getting-better/

Murray, S., Ma, X., & Mazur, J. (2008). Effects of peer coaching on teachers' collaborative interactions and students' mathematics achievements. *The Journal of Educational Research*, *102*(3), 203-212.

NSDC (2001). Professional development (definition of). Retrieved from http://www.aypf.org/documents/62609NSDCDefinitionofProfessionalDevelopment908.pdf

Papay, J. P. (2011). Different tests, different measures: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal, 48*(1), 163-193. Retrieved from http://aer.sagepub.com.ezproxy.net.ucf.edu/content/48/1/163.ful

Petersen, P. B. (1999). Total quality management and the Deming approach to quality management. *Journal of Management History, 5*(8), 468. Retrieved from http://ezproxy.net.ucf/login?url=http://search.proquest.com/docview/210946964?accountid=10003

Professional Career Development (2014). Retrieved from www.professionaldevelpmentpath.com

Rook, D., & Torbert, W. R. (2005). Seven transformations of leadership. *Harvard Business Review*. Retrieved from http://hbr.org/2005/04/seven-transformations-of-leadership/ar/1

Rothman, R. (2010). Beyond test scores: Adding value to assessment. *School Administrator, 67*(2), 20-24.

Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluations in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation.* (Research Report). Retrieved from http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED527619

Scannella, A., & McCarthy, S. (2014). Adversity or opportunity? *Principal Leadership, 14(5),* 52-55. Retrieved from http://ezproxy.net.ucf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=93685692&site=ehost-live

Sinkinson, C. (2011). An assessment of peer coaching to drive professional development and reflective teaching. *Communications in information literacy, 5(1*), 9-20.

Sparks, G. M., & Bruder, S. (1987). Before and after peer coaching. *Educational Leadership, 42*(3), 54-57.

Sparks, S. D. (2012, October 31). Caution urged on 'value-added' reviews. *Education Week, 32*(10), p. 6.

Standards for professional learning. (2011). Retrieved from learningforward.org/standards.

Stecher, B., Garet, M., Holtzman, D., & Hamilton, L. (2012). Implementing measures of teacher effectiveness. *Phi Delta Kappan, 94(3),* 39-43.

Steinberg, W. J. (2011*). Statistics alive!* (2nd ed.) Los Angeles, CA: Sage.

Student Success Act. (2011). CS/CS/SB736: Education Personnel. Retrieved from
www.flsenate.gov/session/bill/2011/0736

The System of profound knowledge. (2014). Palos Verdes Estates, CA: The W. Edwards
Deming Institute. Retrieved from
https://www.deming.org/theman/theories/profoundknowledge

Thornton, S. (2012). *The importance of reliability in performance appraisals.* Retrieved
from http://www.ehow.com/about_5445066_importance-reliabilty-performance
appraisals.html

Tomlinson, C. A. (2011). Every teacher a coach. *Educational Leadership, 69*(2), 92-93.

Tschannen-Moran, B., & Tschannen-Moran, M. (2011). The coach and the evaluator.
Educational Leadership, 69(2), 10-16.

U. S. Department of Education. (2009). *Race to the top program executive summary.*
Retrieved from http://www2ed.gov/programs/racetothetop/executive-
summary.pdf.

U. S. Department of Education (2010). Nine states and the District of Columbia win
second round of race to the top grants. [Press release]. Retrieved from
http://www.ed.gov/news/press-releases/nine-states-and-district-columbia-win-
second-round-race-top-grants

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our
national failure to acknowledge and act on teacher effectiveness.* New York:
NY: The New Teacher Project.

Yates, S. M. (2007). Teachers' perceptions of their professional learning activities.
International Education Journal, 8(2), 213-221. Retrieved from
files.eric.ed.gov/fulltext/Ef834237.pdf

Yoon, K. W., Duncan, T., Lee, S. W., Scarloss, B., Shapley, & Shapley, K. L. (2007,
October*). Reviewing the evidence on how teacher professional development
affects student achievement*. (Issues & Answers Report, REL 2007-No. 033.)
Retrieved from
http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033.pdf