2015

Policy Implications of a Teacher Evaluation System: The Relationship of Classroom Observations, Levels of Feedback, and Students Achievement Outcomes

Shana H. Rafalski
University of Central Florida

POLICY IMPLICATIONS OF A TEACHER EVALUATION SYSTEM: THE RELATIONSHIP OF CLASSROOM OBSERVATIONS, LEVELS OF FEEDBACK, AND STUDENT ACHIEVEMENT OUTCOMES

by

SHANA H. RAFALSKI
B.S. University of South Florida, 1993
M.S. University of South Florida, 2001

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 purpose of this study was two-fold: (a) to determine the relationship between the number of classroom observations and teacher VAM scores and (b) to identify the relationship between the types of feedback provided to teachers and student achievement outcomes as measured by VAM scores. De-identified data for the sample set of teachers in a large urban school district was gathered for the 2013-2014 year from iObservation by administrators observing teachers using the domains of the Marzano instructional model. The number of observations were compared to VAM scores to determine if teachers with a greater number of observations received higher VAM ratings. The comments recorded and submitted as feedback were also reviewed.

Data were analyzed to identify relationships between the types of feedback provided to teachers and student achievement outcomes as measured by VAM scores. No significant relationship existed between VAM scores and number of observations or percentage of comments for teachers at any grade level. In addition, no significant relationship existed between predominant feedback for teachers and VAM scores.

The information in this study was valuable for understanding the relationships that exist among instructional practice scores, value-added measures, and learning gains to drive conversations with teachers regarding rigorous instruction. Observations and feedback should be a tool for improvement of instruction, but the data confirmed this process continues to be compliance based with inflated scores that do not match the level of performance of students. Changing this is strongly linked to the provision of feedback associated with improving instruction and holding teachers accountable in meeting the

standards outlined in the feedback. Observers are in need of professional development on how to provide effective feedback in the areas of instruction that will make the biggest impact on student achievement. Continuing to put time and effort into implementing and monitoring evaluation systems without further training and emphasis on feedback will result in the same lack of impact on student achievement outcomes and may even undermine the role of observers in providing support to teachers.

To my grandmother and hero, Emma Howard

ACKNOWLEDGMENTS

Benjamin Disraeli has been quoted as saying, “One secret of success in life is for a man to be ready for his opportunity when it comes.” Being ready has come with sacrifice. My unending love and gratitude is offered to my sweetheart Andrew who walks softly and steps deliberately with me on the stepping stones of life. Shelby, Nicholas, Emily and Connor, you are the constant source of joy, inspiration and hope! Thank you to my parents for showing me anything is possible and who never doubted I could do anything I set my mind to; and for my sister for providing me with a healthy dose of sibling rivalry neutralized by the kind of love only sisters know.

From the beginning to the end, my chair Dr. Barbara Murray provided guidance, encouragement and pushed me to aim higher. I would also like to thank the members of my committee, Dr. Lee Baldwin, Dr. Walter Doherty, and Dr. Cynthia Hutchinson. I appreciate the time and expertise each has selflessly given toward getting me ready to be a better educational leader and researcher.

My career has been filled with a series of exceptional educational leaders who have been responsible for pushing, pulling and preparing me for each next opportunity. Faychone Durant (“Mama Faye”) taught me how to lead by example, and her words still roll off my tongue. I will be eternally grateful to Dr. Beverly Carbaugh for being my benefactor. To Dr. Okhee Lee, I am still in awe of how you set me on this journey over coffee and then never let me slow down or stumble. I don’t know where I would be without Dr. Brandon McKelvey who kept me from “squirrelling” and helped me to see that the impossible is just within reach! And finally, for reminding me this is only the

beginning and exceeding expectations has value, a hearty thank you is extended to Dr. Scott Fritz who always stayed one step ahead!

Completing this dissertation has been as much about the journey as the destination. It allowed me to reach just beyond my comfort zone and taught me as much about myself as the topic. I commit to paying it forward, being a lifelong learner, and being ready for my time when it comes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
PROBLEM OF PRACTICE

Introduction

Across the country and around the world, measuring teacher effectiveness has become an issue policy makers are writing into law as they demand more accountability for student success. It has also given teachers a reason to speak out as the demands on them have increased. With this accountability, many questions resonate as to the content that is most important for students to learn, the strategies that improve student learning, and the most effective and reliable ways to measure learning. Much of the recent focus can be attributed to the introduction of the federal government's competitive Race to the Top grant and the involvement in the educational arena of philanthropies such as the Gates Foundation. The resolution of this issue also has implications for policy as the way the improvement process is managed will have an influence on costs to the public, on the earnings of individuals and on the future of the economy as a whole. As Hanushek (2011) observed in noting that lower achievement means slower growth in the economy, "Thus the achievement gap between the US and the world's top-performing countries can be said to be causing the equivalent of a permanent recession," (p. 40).

From a historical perspective, the purpose of the compulsory public education system from its inception was to develop an informed and responsible citizenry. Although many additional demands have been placed on the system, the fundamental purpose has not changed. What has not been fully agreed upon, however, is how to measure its success. *A Nation at Risk* (The National Commission on Excellence in

Education, 1983) "sparked increased dissatisfaction with public education, resulting in public demand for improved levels of student achievement" (Doran & Izumi, 2004. p. 13); and sent the message for the need for teacher salaries to be "professionally competitive, market-sensitive and performance-based" (Toch & Rothman, year, p. 1). Though No Child Left Behind (NCLB), requiring each state to make Adequate Yearly Progress (AYP) towards state proficiency goals, was signed into law by President George W. Bush on January 8, 2002, there was little hope of reaching the goal of proficiency for 100% of all students by 2013-14. NCLB did, however, expand state oversight and regulation of evaluation practices. Still, almost all teachers continued to receive the highest performance ratings, and a great number of students were not attaining proficiency. According to Fusarelli (2002), "Despite wave after wave of school reform, student achievement, particularly in urban schools, remains abysmally low" (p. 561).

Based on a large body of research there are sets of generally accepted strategies, which have yielded positive results related to student outcomes. These strategies have emerged from meta-analysis of many researchers' work. What actually will spur student progress toward wide-spread student achievement has continued to be the topic of a great deal of conversation in the field of education today, as many of the strategies and initiatives tested over the years have not brought the change needed for global competitiveness. Edmonds (1979) has been quoted as saying

> We can, whenever and wherever we choose, successfully teach all children whose schooling is of interest to us. We already know more than we need to do that.

> Whether or not we do it must finally depend on how we feel about the fact that we haven't so far (p. 23).

Researchers have repeatedly indicated that nothing measures up to the quality of a teacher in determining student achievement. Initiatives, such as "class-size reduction, curriculum revamping, reorganization of school schedules, investment in technology, all fall far short of the impact that good teachers can have in the classroom" (Hanushek, 2011, p. 40). Individual teachers have been acknowledged to have a greater influence on student achievement than any other factor.

With the push to better assess effectiveness, there has been a call for a more systematic approach to classroom observation through the use of multiple measures, with student outcomes as evidence of teacher effectiveness. Emerging as a better method for improving instruction is providing timely feedback following a large number of unannounced classroom visits, allowing observers to get a clear picture of a teacher's performance across a school year rather than using the "drive-by" observations of the past (Marshall, 2011). Contemporary studies have confirmed the value of targeted feedback as a highly effective strategy (Hattie, 2009; Taylor & Tyler, 2011).

As school districts continue to implement standards-based initiatives and the Common Core State Standards, teacher and administrator evaluation systems, computer-based testing, new data systems and measuring student learning growth through a Value Added Model (VAM), it is important to be aware of the impact on curriculum (what is taught), instruction (how it is taught) and assessment (how results are measured). There is also a need to build capacity within the system to carry out reform efforts. The main underlying assumption is

that instruction will improve if leaders are willing and able to provide detailed feedback to teachers, including suggestions for change (Louis, Leithwood, Wahlstrom, & Anderson, 2010). At the beginning of this new policy era, educators would be remiss if they ignored the potential for feedback to be at the center of improvements in classroom instructional practice. The connection between research on teacher evaluation and feedback, coupled with the high stakes being placed on VAM results in numerous school districts, led to the problem and purpose of this study.

Problem Statement

There is widespread understanding that the teacher evaluation process should be conducted for the ultimate purpose of improving student achievement. To date there has been little research into the influence on student achievement of the frequency of classroom observations or their relationship to forms of feedback provided by observers following classroom observations.

Purpose Statement

Building on current research on observation frequency, one purpose of this study was to determine the relationship between the number of classroom observations and teacher VAM scores. In addition, building on the research on effective feedback, another purpose of this study was to identify the relationship between the types of feedback provided to teachers and student achievement outcomes as measured by VAM scores.

The findings of this study were intended to inform policy on current teacher observation practices and feedback.

Definition of Terms

The following definitions, presented in four categories, are offered to ensure clarity of understanding in this document. Included are: (a) definitions related to evaluation and effectiveness, (b) definitions related to value-added measures, (c) definitions related to feedback, and (d) operational definitions.

Definitions Related to Evaluation and Effectiveness

Accountability systems**--** Accountability systems provide useful data on instructional delivery to educational practitioners and provide accurate and reliable information reflecting the quality of the educational program (Doran & Izumi, 2004). School accountability systems

> have 3 distinguishing characteristics: 1. A shift from input (process) to output (performance) standards; 2. Greater emphasis on what students should know and be able to do; and 3. A push to link often fragmented state policies into a coherent framework (systematic accountability reform). (Fusarelli, 2002, p. 570).

Cognitive Complexity--A structure for identifying the alignment of the cognitive demands placed on learners (CPALMS, 2014)

Deliberate Practice--Deliberate practice involves specifically identifying the elements that will have the biggest impact on student achievement, practicing them, and

getting feedback on progress. A baseline for performance is established in a focus area; and teachers engage in cycles of focused practice, feedback, and monitoring of progress within a time-bound goal for improvement.

Design Question (DQ)--In *The Art and Science of Teaching* (Marzano, 2007), the characteristics of effective teaching have been organized into broad categories framed as "design questions." These are questions that teachers ask themselves when planning a lesson or unit of instruction. The Marzano Instructional Model Learning Map is included in Appendix A.

Element--Research-based strategies are interpreted in the Marzano model as elements. These elements are described through desired effects, and evidence is gathered through teacher and student observed behavior. The Marzano Instructional Model Learning Map is included in Appendix A.

Formal observation--Generally, the formal observation is used as the observation for summative evaluation, lasts for an entire class period and provides a rich source of feedback to teachers regarding their instructional practice and professional growth. It includes a pre-conference and a post-conference for reflection with the teacher (RTTT glossary). For the sample district, one formal observation is required annually for a teacher with three or more years of experience in the district. Three years of experience is the point at which under state statute a teacher is no longer considered to be in their developmental period.

Informal observation-- Informal observations can be announced or unannounced and typically last from 10 minutes to a full class period. They are used to provide

feedback, track deliberate practice growth, and to collect evidence to inform the annual evaluation process. For the sample district, two informal observations are required annually for a teacher with three or more years of experience.

Instructional practices-- These are comprised of a set of observed teaching strategies by which principals assign levels of competence to teachers in delivering instruction (Marzano, 2007).

Marzano Protocol--This protocol consists of 41 key strategies revealed by research for effective teaching presented in a robust, easy-to-understand model of instruction based on *The Art and Science of Teaching* (Marzano, 2007).

Multiple measures-- In an evaluation system, three factors are examined. They include, "teacher effectiveness-classroom observations, student achievement gains and feedback from students--meant to compensate for the imperfections of each individual measure" (Marshall, 2012, p. 50).

Observer --Anyone trained and authorized to do informal or formal teacher observations, including rating elements and giving feedback to teachers. This could include principals, assistant principals, instructional coaches, directors, senior administrators or coordinators, as well as other administrators.

Observation Rating Scale--The Marzano observation system includes the following rating scale: (Appendix B: Teacher Domain 1 Observational Protocol).

Innovating (4)--Adapts and creates new strategies for unique student needs and situation.

Applying (3)--Teacher used the strategy and monitors the extent to which students understand their level of performance.

Developing (2)--Engages students in the use of a strategy but does not monitor the use.

Beginning (1)--Uses strategy incorrectly or with parts missing.

Not Using (0)--Strategy was called for but not exhibited. (Appendix B: Teacher Domain 1 Observational Protocol)

Teacher effectiveness-- Effective teachers are those whose students experience high academic growth, while the students of less effective teachers experience less academic growth, (Stronge, Ward & Grant, 2011, p. 339). It is the combination of teacher inputs (qualifications), the teaching process (instructional practices), and the product of teaching (effects on student learning), (Stronge et al., p. 341).

Teacher evaluation system**--**The Marzano Instructional Model (carried out through iObservation) served as the basis of the teacher evaluation system in this study.

Walkthrough observation--As cited in *Effective Supervision: Supporting the Art and Science of Teaching* (Marzano, Frontier & Livingston, 2011, p, 57) and *Teacher Evaluation That Makes a Difference* (Marzano & Toth, 2013, p. 64), the Downey walkthrough approach is outlined as being a short, focused, yet informal observation used to identify a possible area for teacher reflection, rarely requiring follow-up and would not include a checklist for introspection. Lasting 3-10 minutes, the observer gathers evidence regarding classroom instructional practices and behaviors.

Definitions Related to Value-added Measures

Outcome-based performance measures--For the purpose of this study, the measures are "the various models produce a positive or negative number that describe either a teacher's performance in relation to that of typical teachers or the average growth of students in typical teachers' classes" (Gagne, 2011, p. 4).

Value added model--"In general, value-added models are a class of statistical procedures that use longitudinal test score data, i.e., data collected over a period of time, to measure the change in a student's performance during a specific period of time" (Doran & Izumi, 2004, p. 3). A value-added measure is the metric assigned to specific teachers based on growth in the learning of the students they taught during a specified period of time (Ravitch, 2010) or the difference between the predicted performance and the actual performance represents the value-added by the teacher's instruction (Florida Department of Education, 2014).

Definitions Related to Feedback

Feedback-- For the purpose of this study, feedback is defined as "information about how we are doing in our efforts to reach a goal" (Wiggins, 2012, p. 11), and "Feedback needs to provide information specifically relating to the task or process of learning that fills the gap between what is understood and what is aimed to be understood" (Hattie & Timperley, 2007, p. 82).

Targeted feedback--This refers to feedback that is informative, constructive, objective, actionable, and focused on specific classroom strategies and behaviors during a set time interval, (Florida RTTT glossary).

Negative reinforcement--This refers to a response or behavior strengthened by stopping, removing or avoiding a negative outcome or aversive stimulus (Skinner).

Positive reinforcement-- This refers to the addition of a reinforcing stimulus following a behavior that makes it more likely that the behavior will occur again in the future (Skinner).

Operational Definitions

The following operational definitions outline the way terms were interpreted for this study:

Feedback alignment-- For the purpose of this study, feedback alignment is appropriate and matched commentary given in observations based on the content teachers are teaching and the method they are using teaching it in relation to effectiveness.

Reinforcement theory-- Behavior is a function of consequences creating a cause and effect relationship. Behaviors followed by positive consequences (i.e. reinforced) will occur more frequently, and behaviors either followed by negative consequences or not followed by positive consequences will occur less frequently, (Williams, 2014).

Rubric-- A rubric is a guide for communicating expectations of quality for a task by setting clear criteria and listing specific measures for scoring. On the rubric, the following categories are organized by level:

Level 1-No feedback-- The observer provides no opinion in the comment section of the protocol.

Level 2-Unrelated feedback or General Statement-- The observer gives some information in the comment section but it is not relevant to the element or meaning cannot be interpreted.

Level 3-Recount of Observation Events-- This could include a narrative of what the teacher and students were doing during the observation, general statements of events, or notes the observer took to justify the rating given. In some instances the observer included statements to support the effectiveness of a strategy.

Level 4 -General Affirmation or Praise Statement—The observer either leaves a single word or phrase to indicate approval or adds a complement to the end of a recount of observation events.

Level 5-Reflective feedback or Reflective Question--The observer asks the teacher to think about their practice or a specific element in either a general or specific way.

Level 6-Standardized feedback-- The observer uses the cut and paste option in the protocol to leave systematized feedback.

Level 7-Specific targeted feedback-- The observer leaves differentiated and meaningful statements intended to improve the impact of an instructional strategy.

Standards-based instruction--This type of education is based on standardized measures. It is the connection between curriculum and assessment.

Significance of the Study

Evaluating teacher effectiveness has become a dominant theme in 21st century American education reform efforts, in no small part due to the extreme variation in teacher expertise as measured by capacity to stimulate growth in student achievement (Taylor & Tyler, 2011). "The current emphasis on, and the pervasiveness of, performance-based accountability distinguishes this era from previous eras," (Fusarelli, 2002, p. 562), but the concept is not entirely new. As early as 1949, Tyler emphasized prioritizing outcomes of learning as opposed to content to be taught (Conklin, 2005). There has been a significant amount of time, energy, and money devoted to building systems and skills to improve teacher effectiveness that can be documented by student achievement outcomes.

Clarity in defining teacher effectiveness is important because what gets measured is a reflection of what is valued (Goe, Bell, & Little, 2008, p. 4). Measuring the worth of teachers seems to be of great value based on the development of various evaluation systems and the numbers of teachers and administrators being trained in their implementation. Despite teachers learning about effective strategy use and observers attending inter-rater reliability training to calibrate and ensure the validity of the instruments, uncertainty remains whether all of the emphasis is impacting instruction and student learning.

For many things in life, more is better. "It is often suggested that the more frequently feedback is provided the more effective the resultant performance will be" (Fedor & Buckley, 1987, p. 172). The question exists as to whether this holds true for

teacher observation as it relates to improving student learning. If it is widely believed that "improvement of teacher quality is essential to improve educational achievement," (Yeh, 2007, p. 221) then knowing how to improve this quality is paramount.

Researcher and author, Kim Marshall (2011), suggested a better alternative to the past practice of teacher evaluation, which she equated to nothing more than infrequent scripted events likened to a "dog and pony show." She has recommended engaging in 10 brief, unannounced classroom visits of 10-15 minutes each in a year by the same administrator. These visits would vary, occurring at the beginning, middle, and end of lessons and for different subject areas or classes or different times of the day and days of the week. Each observation, according to Marshall (2011), should be followed promptly by a face-to-face coaching conversation and then brief written feedback. Further supporting this idea, Marzano and Toth's (2013) recommendation was to collect multiple samples of data regarding classroom practices by "increasing the number of observations required within the evaluation systems" (p. 13).

Stronge and Hindman (2003), like other researchers, reinforced the importance of teachers, stating that "the common denominator in school improvement and student success is the teacher" (p. 48). However, the reports of past practice and the importance of teachers have not contributed to documented, significant changes in student achievement.

Many researchers have investigated ways to improve teacher effectiveness and identified strategies most likely to lead to gains in student achievement. In discussing *The Widget Effect*, a report released by The New Teacher Project (TNTP), Jerald (2012)

found teacher evaluations to be infrequent, subjective, producing inflated performance ratings which "provided almost no useful feedback to help teachers improve" (p. 6). Another project, Measures of Effective Teaching (MET), funded by the Bill and Melinda Gates Foundation, was initiated to connect teacher observations and student performance. Jerald observed that "accurate feedback based on observation instruments can be a powerful resource for improving teaching and learning," (p. 3). He also referred to the findings of other studies, suggesting that inaccurate feedback or non-specific feedback will fail to provide opportunities for growth or improve teaching and learning.

Feedback can be quantitative or qualitative. Quantitative feedback is measured by capturing a teacher's overall skillset over an extended period of time. In contrast, qualitative feedback takes the form of coaching following an observation (Jerald, 2012). By increasing the number of feedback opportunities, the feedback gap can be closed and improvement compounded. This speaks to quantity and quality.

Building on the current body of knowledge, this study aimed to focus specifically on the feedback teachers receive during and after classroom observations. This feedback was classified using a rubric to determine what type of feedback was given and if the feedback was specific, relevant and targeted. Outcome data, including effectiveness ratings and VAM scores were analyzed to determine if correlations existed between (a) number of feedback sessions, (b) if feedback was given, and (c) the type of feedback given. The results of this study were intended to inform practice by providing information on new practices, including whether increasing the number of classroom observations and providing teachers with specific, targeted feedback has a relationship to

student achievement outcomes as measured by VAM. The results of the study have the potential to impact professional development recommendations. If the analysis of data revealed that more observations did not yield higher VAM scores, the current training schedule and recommended number of observations could change.

Conceptual Framework

To understand the relationship between improved performance and feedback, the researcher reviewed the work of numerous theorists to determine the conceptual origins. From the myriad of behavior theories, the work of several theorists emerged. These include: Lewin's idea that behavior is the result of interactions between the individual and environmental factors; Watson's thought that behaviorism is stimulus and response; Pavlov's conditioning; Skinner's distinction between operant and respondent behavior focused on conditioning aspects; Weiner's reinforcement theories; and deCharms' intrinsic and extrinsic motivation construct. For the purpose of this study, a combination of theories were considered, and reinforcement theory emerged as predominant after control theory and cognitive theory were reviewed.

Control theory proposes that feedback could be considered an intrinsic motivator based on how an individual actively seeks and interprets feedback (Fedor & Buckley, 1987). Cognitive theory or cognitive process is described as organismic and is linked directly to stimulus combined with a behavioral response. Bandura (1977) explained that human behavior is developed through modeling. This is unlearned behavior linked to the respondent, controlled or elicited by prior stimulation (Pate, 1977).

Comparatively, reinforcement theory or reinforcement process is mechanistic and performance-based. This operant or learned behavior is influenced by events and linked to successful performance and self-correction from feedback (Pate, 1977).

> Reinforcement theory says that behavior is a function of its consequences, that behaviors followed by positive consequences (i.e., reinforced) will occur more frequently, and that behaviors either followed by negative consequences or not followed by positive consequences will occur less frequently. (Williams, 2014)

Positive and negative consequences are also known as positive and negative reinforcement.

Brauer and Tittle (2012) asserted that differential reinforcement is "the balance of anticipated or actual rewards and punishments that follow or are consequences of behavior" (p. 165) and that learning is the result of both direct and vicarious behavioral reinforcement.

Research Questions

For the purpose of this study, "teachers" were those with a matched Value Added Model (VAM) score, more than three years of teaching experience, and no National Board Certification. Observations and feedback were provided through the Marzano Instructional Model and student achievement was measured by the VAM. Following are the research questions which were used to guide the study:

1. What is the frequency of classroom observations and comments for teachers, including formal, informal, and walkthrough observations?

2. What relationship, if any, exists between the frequency of observations by observers as measured by the number of classroom observations and elements scored during a school year and student achievement outcomes as measured by teacher VAM scores?

   H_{01}. There is no significant relationship between the number of classroom observations during a school year and student achievement outcomes as measured by VAM.

3. What is the frequency by level of feedback defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?
4. What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

   H_{02}. There is no significant difference between classroom observations by feedback category and student achievement outcomes as measured by VAM scores.

Delimitations

This study was delimited by the following:

A large Florida urban school district employing approximately 14,000 teachers was chosen. To account for the variable of professional growth, novice teachers with

fewer than three years of teaching and expert teachers with National Board Certification were excluded from the study. The researcher reviewed one year of observation data and VAM scores from 2,718 teachers, based on the criteria for teacher selection being used in the study. The sample was drawn from a population of elementary, middle and high school teachers.

The number of walkthrough, informal and formal observations teachers in the sample received was compared to their VAM score, and descriptive statistics were used to interpret the results. Implementation included the use of the scales on the protocols within the Marzano model. Fidelity to this model was increased through initial training, inter-rater reliability follow-up, and master observer training.

Limitations

The researcher identified the following limitations for this study:

The policy context for the selected evaluation system and the scales included within the model were recognized as delimitations. The Marzano Instructional model, on which the study was based, was the Florida state model for teacher evaluation and the model selected for use by the school district. Although accounted for through inter-rater reliability training, it was assumed that there would be variability in the interpretation of the Marzano instructional model by administrators. Additionally, the years of service and level of training of administrators could have an impact on the results of the study. Finally, the extent to which schools were following mandates as they pertain to fidelity of program implementation could have affected outcomes. The limitations of the classroom

observation model include the use of 41 instructional strategies identified in Domain 1 of the Marzano evaluation model. The model limits administrators to rating teachers during observations using a four-point scale, and only trained administrators may conduct classroom observations. The maximum number of observations completed by an administrator was beyond the control of the researcher. The observation protocol considers only the minimum number of observations required for teachers.

All feedback may not have been recorded within the iObservation tool as prescribed by the program. Many administrators may have given "off the record" feedback in the form of written or verbal communication which is not part of a data source for the study. Electronic feedback was the only data source for the study.

Assumptions

As part of the implementation of the Marzano Instructional Model there are a minimum number of observations that administrators must complete per teacher per year. This number is greater for teachers with fewer than three years of experience; however, they were excluded from the study.

District administrative personnel, principals and assistant principals, and other trained personnel referred to in the study as "observers" were required to be trained on the use of the system and participated in follow-up inter-rater reliability training, as well as master observer training.

Administrators were expected to follow the fidelity of implementation, which included performing a minimum of two observations during the school year using the

rating scales within the protocols, to ensure the appropriate number of monitoring and feedback opportunities occurs.

All teachers in the sample were working on deliberate practice strategies as part of their individual professional development plans. All teachers are required to submit their choices for deliberate practice at the beginning of each school year. It was assumed that this focused practice on specific strategy use would be impacted by classroom observational feedback.

All students in Grades 3-10 take state high-stakes assessments in the areas of reading and mathematics, the results of which were used to determine student achievement outcomes and VAM scores. Florida state law as it relates to VAM is dictated and thus was outside of the control of this study. At the time of the study, as part of Race to the Top participation, the Florida Legislature required that 40% of teachers' evaluations be based on value-added measures (Florida State Statute 1012.34, 2012).

Organization of the Study

This report of the research has been organized in five chapters. Chapter 1 has included a statement of the problem, the purpose and significance of the study, definition of terms, conceptual framework, research questions and the limitations, delimitations, and organization of the study. Chapter 2 contains a review of the related literature. Chapter 3 describes the methodology used for the research study. Chapter 4 presents the findings of the study. Chapter 5 provides a summary of the study, discussion and implications of

the findings for educational policy and practice on teacher evaluation, and recommendations for further research.

CHAPTER 2
REVIEW OF LITERATURE

Introduction

Researchers have frequently observed that student learning is a result of the quality of teaching that occurs in a classroom. To this end, the components for what makes up effective teaching has become a controversial topic of great interest. Changes in national and state policy requiring the measurement of teacher quality has raised questions as to the impact of the many changes on results. These questions have fueled the accountability issue, requiring valid and reliable systems to be built around measuring the effects of specific teaching strategies and behaviors.

Initially the researcher used the UCF Library One Search online reference tool to find resources to support the topics incorporated in the research questions which guided this study. She initially narrowed the search based on the big ideas within the research questions. As a strategy to ensure that all sources were located, however, she scheduled a research consultation with a research librarian. During the session, assistance was provided in using the ERIC library, Web of Science, WorldCatDissertations, Dissertations & Theses Full Text, and PsycInfo databases.

This process narrowed the focus of topics for deeper investigation and formed a basis for understanding the components of this study. The researcher subsequently engaged in a review of the literature surrounding the following five topics associated with the research questions: (a) reinforcement theory, (b) teacher evaluation, including the Marzano instructional model, (c) teacher effectiveness, (d) Value-added measures of

student achievement, and (e) feedback. The literature review presented in this chapter has been organized around these five areas of interest.

Section 1 of the review focuses on the behavioral aspects of evaluation and feedback, educational feedback, and behavior and reinforcement theory. As noted by Lunenburg and Irby (2008), "The theoretical framework for the first section of the review of literature builds a base for the researcher's dissertation topic as is relates to the behavioral aspects of evaluation and feedback, change theory, and behavior and reinforcement theory" (p. 122).

Sections 2 and 3 concentrate on the process of teacher evaluation, and models related to evaluation and teacher effectiveness. An abundance of articles and books on teacher evaluation systems and teacher effectiveness emerged in the review, exploring the categories and components of a variety of evaluation systems, thereby providing information on program effectiveness. Further investigation enabled comparisons to alternative evaluation systems presently in use around the country.

Much of the contemporary information related to teacher evaluation included references to value-added models (VAM) as a component. This led the researcher to a variety of articles related to the topic of VAM, their definition, and studies by schools or districts using VAM to measure student growth. Section 4 provides a summary of the VAM information that was reviewed.

Section 5 contains a review of literature and research related to feedback. Articles on feedback initially appeared to be abundant; however, information on teacher performance feedback for improvement was veiled behind other terms such as cognitive

feedback, self-efficacy, evaluative feedback, corrective feedback, and information processing. Regardless of context, the definition of feedback appeared to be consistent throughout the literature reviewed.

Throughout the process of reviewing the literature, the researcher continued to acquire resources by reviewing reference lists of related books, articles, and documents. This exponentially increased the references at the researcher's disposal and provided a basis for the conclusion that there were adequate resources available to support the present research.

Research Related to Reinforcement Theory

The research related to reinforcement theory includes definitions and suggestions for how to shape behavior in others.

> Reinforcement theory says that behavior is a function of its consequences, that behaviors followed by positive consequences (i.e. reinforced) will occur more frequently, and that behaviors either followed by negative consequences or not followed by positive consequences will occur less frequently (Williams, 2014, p. 95).

Understanding what motivates others is important, especially for those who manage the work of subordinates. According to Brauer and Tittle (2012), "Learning occurs through both direct and vicarious behavioral reinforcement," (p. 159). Increasing productivity and meeting or exceeding desired outcomes is reliant upon being able to motivate and encourage others to maximize their potential.

This theory includes reinforcement contingencies as one of the components, which are the cause and effect relationships between performance and consequences (Williams, 2014). They are related in part to what Brauer and Tittle (2012) described as differential reinforcement, or the relationship between anticipated or actual rewards and the punishments that follow as consequences of behavior. Other components and terminology include positive reinforcement, negative reinforcement, punishment, extinction, and reinforcement schedules (Williams, 2014). Negative feedback or sensitivity to punishment is associated with negative outcomes, and positive feedback or sensitivity to reward is associated with positive outcomes (Hundt et al., 2012). "Studies showing the highest effect sizes involved students receiving information feedback about a task and how to do it more effectively. Lower effect sizes were related to praise, rewards and punishment" (Hattie & Timperley, 2007, p. 84).

Reinforcement theory is directly related to teacher evaluation because it makes the connection between what teachers do in their classrooms and the results, i.e., the evaluation judgments that are made. Observation, feedback and evaluation have a natural connection to behavior theory. "Feedback enables individuals to understand and improve their judgments, improve their expertise in the judgment task, and reduce commitment to incorrect judgment strategies" (Balzer, Doherty, & O'Connor, 1989, p. 412). Depending on the type, specificity, and follow-up from this feedback, the feedback is considered reinforcement. There are steps in the process of motivating employees with reinforcement theory that make it effective. They include identifying observable behaviors, measuring baseline frequencies, analyzing causes and consequences for

behaviors, intervening by using reinforcement and evaluating how reinforcement changed the behavior (Williams, 2014). There are also applications for reinforcement theory within personal relationships.

Researchers have shown that motivation, whether positive or negative, can have an impact on performance. Rowan, Chiang, & Miller (1997) reinforced this idea by stating that differing "perspectives suggest that higher motivation increases the performance of tasks, but each proposes a different source of motivation" (p. 260).

Teacher Evaluation

There are six specific areas of interest related to teacher evaluation: (a) data surrounding the persistence of teachers nationwide receiving inflated evaluation ratings while the performance of students remains low, (b) ratings in high cognitive complexity strategies, (c) the need for change, (d) types of evaluation systems and evaluation processes, and (e) the linkage between teacher evaluation and student learning gains.

"There is widespread agreement among researchers and policymakers that teachers matter significantly in improving student learning," (Little, 2009, p. xi) and "may be the most important school-based factor in increasing student achievement" (Little, 2009, p. xi). Teacher evaluation has come under a great deal of scrutiny in recent years as results from around the country have shown that more than 95% of teachers are rated at the highest level and yet student achievement is not reflective of these same levels.

> In recent years evaluating teacher effectiveness has become a dominant theme in American education reform efforts, an emphasis motivated in no small part by the apparently large variation in teacher productivity as measured by ability to promote student achievement growth. (Taylor & Tyler, 2011, p. 1).

For over 20 years, personnel evaluations for educators have been mandated in every state; however, the results in the form of student achievement continue to be lacking (Frase & Streshly, 1994). In the state of Florida, even in the wake of major systemic change, teacher ratings have remained stable, i.e., inflated; and student learning has remained flat in both reading and mathematics with fewer than 60% of students performing at or above proficiency. This has fed the belief that evaluation is a "perfunctory bureaucratic requirement that yields little help for teachers and little information on which a school district can base decisions" (Darling-Hammond, 1986, p. 530).

In addition, it has been found, in decades of research, that administrators pervasively inflate evaluation ratings. Studies have repeatedly revealed that even teachers assigned the highest observational and evaluation ratings use some of the most ineffective instructional practices, e.g., drill and practice, copying, lack of lesson planning, low-quality worksheets, and lack of student engagement (Frase & Streshly,1994). In his keynote speech at the Building Expertise Marzano International Conference on June 18, 2014, Toth noted that administrators continue to inflate observation scores by a full scale rating (applying level) above student performance even as teachers self-evaluate themselves as a level below (beginning level). Furthermore, a

review of more than two million data points analyzed by the Learning Sciences Marzano Center revealed that teachers continue to "rely heavily on traditional teacher-centered strategies to deliver content," rather than move to student-centered activities to help move pupils toward being able to solve complex problems and apply their knowledge (Marzano & Toth, 2014). Unfortunately, despite the awareness of the consequences of such a trend, Little reported in 2009 that the problem persisted and had been reported for two decades, even in schools with dismal student achievement scores.

The remedy is said to be implementing instructional strategies associated with high cognitive complexity. According to Marzano, speaking at the same international conference, "We should see evidence of students wrestling with new content as they build the stamina required to reach higher levels of thinking." The way this can be accomplished is through a transition from overusing strategies associated with lecture, practice and review to the strategies most critical for developing cognitive complexity. At the same conference, Marzano discussed what he termed 13 essential strategies (including the Super 7*) associated with cognitive complexity. They include elements from the Marzano Instructional Model as follows:

Element 6--Identifying Critical Information*

Element 8-- Previewing New Content

Element 7-- Organizing Students to Interact with New Knowledge

Element 10 --Processing of New Information

Element 11--Elaborating of New Information*

Element 12--Recording and Representing Knowledge*

Element 26--Managing Response Rates (with tiered Questioning Techniques)

Element 14--Reviewing Content

Element 19--Practicing Skills, Strategies, and Processes

Element 17--Examining Similarities and Differences*

Element 18-- Examining Errors in Reasoning*

Element 20--Revising Knowledge*

Element 22--Engaging Students in Cognitively Complex Tasks Involving Hypothesis Generation and Testing* (Marzano & Toth, 2014)

There have been other variations of this theme such as the Excellent 11, and Access to Common Core Strategies. Regardless of terminology, there has been a call for the use of such strategies in moving from teacher-centered to student-centered classrooms. In addition, the need for teachers to have models, training, and feedback in becoming facilitators of learning has been identified.

If educators draw upon research to guide action based on the current status of education and what is known about effective teaching, the need for change becomes obvious. Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012) suggested that effective teachers have the ability to do the following:

- Understand subject matter deeply and flexibly;
- Connect what is to be learned to students' prior knowledge and experience;
- Create effective scaffolds and supports for learning;
- Use instructional strategies that help students draw connections, apply what they're learning, practice new skills, and monitor their own learning;

• Assess student learning continuously and adapt teaching to student needs;

• Provide clear standards, constant feedback, and opportunities for revising work; and

• Develop and effectively manage a collaborative classroom in which all students have membership (p. 13).

If these abilities reflect effectiveness, one would wonder why more teachers have not been recognized as achieving the goal. What is known is, "A focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their request to improve" (Black & William, 2010, p. 81). According to Darling-Hammond (1986), "personnel evaluation in an organization reveals what is valued in the organization," (p. 530). This belief has led to the development of several teacher evaluation models with a strong correlation to high-yield instructional strategies that have become extremely popular and even touted as the silver-bullet for our American education system. Unfortunately, in some instances evaluation systems have been part of the problem for being off target in measuring the correct elements. These systems, lacking in on-target, actionable feedback, have resembled checklists rather than tools to guide outcomes. Other problems that have emerged include the increased burdens of principals who have had to shift from managerial roles to roles as instructional leaders with no relief from the demands on their time. "A typical principal has from 20-100 teachers to supervise, as compared with the supervisory ratio of no more than one to 10 in most other types of organizations" (Darling-Hammond, 1986, p. 533).

It is an understatement to say that a metamorphosis has occurred in teacher evaluation systems over the past century (Ellet & Teddlie, 2003). Most recently, teacher evaluation has been "used for three major purposes internationally--accountability, promotion and staff development, but rarely for teacher or school improvement" (Ellet & Teddlie, 2003, p. 102). Teacher evaluation systems should mirror what is valued in classrooms today, but in order to come to a consensus about what is valued, multiple stakeholders should be involved in the process and multiple measures should be embedded in the system to make them credible (Little, 2009). Additionally, Little (2009) warned, "An evaluation system should be established before the link to pay is made" (p. viii). It seems that in a rush to hold teachers accountable for student outcomes, this step has more often than not, been an afterthought. There is hope that direction can be provided by further examining the work of past researchers. "Successful systems share several common components about effectively measuring teaching and reforming evaluation and compensation systems" (Little, 2009, p. 11). Successful systems include not only the components mentioned by Little but also are connected to standards, have embedded professional development, and provide for targeted support and feedback.

The Bill and Melinda Gates Foundation has funded research to establish the connection between teacher behaviors and student learning outcomes. As a result of this support,

> The Measures of Effective Teaching (MET) Project has developed a number of tools, including observations or videotapes of teachers, supplemented with other artifacts of practice (lesson plans, assignments, etc.), that can be scored according

to standards that reflect practices associated with effective teaching. (Darling-Hammond, L, Amrein-Beardsley, Haertel, & Rothstein, 2012, p. 13).

In their research, Darling-Hammond et al. found that these tools work best when accompanied by a system that has evaluators who are well-trained, provide frequent and targeted feedback and provide follow up coaching and support.

The goal in any system is to measure the cause and effect between teaching and learning. The National Board for Professional Teaching Standards (NBPTS) defined accomplished teaching and established criteria to demonstrate the resulting standards through performance-based assessments. The National Board certification process is an example of a standards-based process that reflects the connection between what teachers do and the results in student learning (Darling-Hammond et al., 2012, p. 13). The students of teachers who have completed the process, gaining certification, have been shown to have increased learning gains.

Researchers have also found that evaluation differs based on teachers' experience. New teachers require different types and levels of support than end-of-career teachers. Data have shown, however, that "high-quality, classroom observation-based evaluation improves mid-career teacher performance both during the period of evaluation and in subsequent years" (Taylor & Tyler, 2011, p. 3).

Overall, the trend in evaluation shows a strong linkage between teacher evaluation scores and student learning. Top rated teachers produced more learning gains than teachers with lower evaluation scores, (Odden, 2004). Odden also suggested that the correlation was sufficient enough to assume that pay increases could be based on the

results of certain performance systems. However, Darling-Hammond et al. (2012) contended that this "assumes that student learning is measured well by a given test, is influenced by the teacher alone, and is independent from the growth of classmates and other aspects of the classroom context" (p. 8). They proposed that school factors, home and community supports, individual and peer attributes, learning styles, prior educational experiences, and even specific tests played a role in influencing learning gains.

Effectiveness

Darling-Hammond (1986) referenced teacher effectiveness as being more than "a monolithic construct" (p. 535). Six key elements of teacher effectiveness are addressed in this review: (a) teacher effectiveness as it relates to student achievement, (b) the multifaceted nature of teacher effectiveness, (c) the link between effectiveness and teacher quality, (d) the impact of National Board Certification, (e) the special needs of new teachers, and (f) the economic impact of teachers.

"The question of whether teachers differ dramatically in their effectiveness in promoting their students' academic achievement is fundamental to educational research" (Nye, Konstantopoulos & Hedges, 2004. p. 237). Numerous researchers have supported the notion that the impact of a teacher, regardless of characteristics such as educational preparation, experience or salary is related to student achievement (Nye et al., 2004). "Researchers have worked hard to isolate the impact of teachers from other influences. Rigorous studies consistently show that the impact of a more-effective teacher is substantial" (Hanushek, 2011, p. 41).

Researchers have found that effective teachers make a difference for students when it comes to their success in terms of learning outcomes. However, the focus in the nation's public schools over the last 40 years has focused on the development of curriculum standards and assessments to measure student achievement. "Unfortunately, much of the foregoing policy discussion has overlooked the most fundamental unit of change--the classroom--and the primary catalyst for improvement in our schools--the teacher" (Stronge, Ward, Tucker & Hindman, 2008, p. 167). This was further supported by Kane and Staiger (2008) who reported that researchers have shown, over a 30-year period, "considerable heterogeneity in teacher impacts on student achievement" (p. 1). Stronge et al. (2008) stressed the importance of having qualified teachers in every classroom in the following statement:

> Given the clear and undeniable link that exists between teacher effectiveness and student learning, the use of student achievement information, when it is curriculum based, can provide an invaluable tool to explore the classroom practices of teachers who enhance student learning beyond predicted levels of accomplishment" (p. 181).

Goe et al. (2008) provided the following five-point description of the characteristics of effective teachers:

- Have high expectations for all students and help them learn;
- Contribute to positive academic, attitudinal and social outcomes for students;
- Use diverse resources to plan and structure engaging learning opportunities, monitor student progress formatively, and adapt instruction as needed;

- Value diversity and civic-mindedness;
- Engage in collaboration. (p. 8)

Thus, if the expectation is for an increase in student achievement, there are steps that need to be taken to improve teacher expertise "through programs of professional development that build on existing good practice" (Black & William, 2010, p. 89). The movement towards the use of value- added models is an attempt to measure the impact of teachers on students, and there has been some indication, according to Kane and Staiger (2008), that "standard teacher value-added models are able to generate unbiased and reasonably accurate predictions of the causal short-term impact of a teacher on student test scores" (p. 33).

The quality of a teacher can make a profound difference for a student. It can mean the difference between success or failure, catapulting ahead or falling behind, or learning and growing or wasting precious time. "Research literature provides a surprisingly precise estimates of the impact of students achievement levels on their lifetime earnings and by combining this with estimated impacts of more effective teachers on student achievement" (Hanushek, 2011, p. 41). Data have shown that there is a positive residual impact on students lasting several years when they are taught by a high quality teacher for even one year. Goldhaber and Anthony (2004) noted that "All else equal, a student with a very high quality teacher will achieve a learning gain of 1.5 grade level equivalents, while a student with a low-quality teacher achieves a gain of only .5 grade level equivalents" (p. 4). Unfortunately, the inverse is also true. It has been determined that it could take up to three years to remediate students who have been taught by an

ineffective teacher (Stronge et al., 2008).

The quality gaps that exist among teachers within schools can be measured. Rockoff (2004) found that a one-standard-deviation increase in teacher quality raises test scores by approximately .1 standard deviation in reading and mathematics. Empirical evidence has suggested that simply reducing the gap in teacher quality by raising instructional skill levels among teachers will result in improved student outcomes (Rockoff, 2004).

By observing the parallels between improving teaching and improving the overall workforce, there have been those who believe that research from outside the field of education can be used to improve the quality of the teaching force (Rowan et al., 1997). Others seeking to improve teacher quality have proposed raising the requirements for admission into teacher education programs and raising the qualifications for earning certification. Darling-Hammond and Youngs (2002) used the results from empirical studies and scholarly articles to refute the wisdom of the following four propositions, providing facts to be considered in future policy recommendations regarding teacher effectiveness.

Proposition 1: Teachers matter for student achievement, but teacher education and certification are not related to teacher effectiveness. This proposition was determined to be false. Darling-Hammond and Youngs (2002) found that researchers employing different units of analysis who had examined the influence of teacher education and certification on student achievement had often found significant relationships between measures of teacher expertise and student achievement.

Proposition 2: Verbal ability and subject matter knowledge are the most important components of teacher effectiveness. In challenging this proposition, Darling-Hammond and Youngs (2002) cited research calling into question the Education Secretary's assertion that subject area knowledge and verbal ability is more important than knowing how to teach.

Proposition 3: Teachers who have completed teacher education programs are academically weak and underprepared for their jobs. The researchers observed that the Secretary's report included several misleading assertions regarding the qualification of the teacher workforce.

Proposition 4: Alternative certification programs have academically stronger recruits, high rates of teacher retention, and produce more successful teachers. Darling-Hammond and Youngs (2002) reviewed research on alternative certification programs in terms of program design and determined that more carefully designed programs yielded stronger outcomes in terms of teacher effectiveness and retention than those that provided less training and support.

Vandevoort, Amrein-Beardsley and Berliner (2004) summarized the status of knowledge regarding teacher effectiveness, concluding that "Contemporary research on teaching indicates that teachers are powerful contributors to students' academic achievement, though the set and interrelationships of characteristics that make for high-quality and effective teaching have yet to be satisfactorily determined" (p. 1).

The National Board Certification process has offered a way to improve and measure teacher quality. The National Board for Professional Teaching Standards

(NBPTS) was founded in 1987 to establish "the definitive standards of accomplished teaching and the process by which the profession would certify whether or not a teacher had met those standards," (NBPTS, n.d.). Just as doctors and lawyers have peer governing bodies, the founders' goal was to elevate the status of teachers to that of other professions by self-regulating the standards of entry, practice, and advancement. Over the next 15 years, the numbers of National Board Certified (NBC) teachers grew to over 40,000 nationwide, each whose application fee cost $2,300 (Cavalluzzo, 2004; Sanders, Ashton & Wright, 2005). In 2013, just a year after celebrating its 25th anniversary, NBPTS reached a new milestone; 100,000 NBC teachers, and claimed to have a disproportionate positive impact on improving education in classrooms in all 50 states (NBPTS, 2014). Results from studies, however, are inconsistent. Cavalluzzo (2004) stated,

> When compared with students whose teachers had never been involved with NBC, we found that students with otherwise similar teachers made larger gains if their teacher had a NBC and smaller gains if their teacher failed or withdrew from the NBC accreditation process. . . NBC proved to be an effective signal of teacher quality. (p. 3).

Further supporting this claim, Vandevoort et al. (2004) found "students in the classes of National Board Certified Teachers surpassed students in the classrooms of non-Board certified teachers in almost three quarters of the comparisons" (p. 2) They further asserted that "Teachers identified through the assessments of the National Board for Professional Teaching Standards are, on average, more effective teachers in terms of

academic achievement, one of the many outcomes of education for which teachers are responsible" (Vandevoort et al., 2004, p. 2).

Sanders et al. (2005) contradicted the positive findings regarding the success of NBCT, writing "Students of NBCTs did not have significantly better rates of academic progress than students of other teachers and estimated effect sizes were relatively small" (p. 2). They further suggested that "a student randomly assigned to a NBCT is no more likely to get an "effective" (or an "ineffective") teacher than a student assigned to a non-NBCT" (p. 2). These findings were in contrast the findings of other researchers, leaving the statement of disproportionate impact by the NBPTS in question.

There are two sides to the issue of economic impact: the cost to society of ineffective teachers instructing the nation's youth and the cost to society to replace them. Economists have estimated that the cost is great for students who have poor teachers. The far reaching impact of quality instruction is a well-educated society that thrives economically. Hanushek (2010) wrote,

> Recent analysis has demonstrated a very close tie between cognitive skills of a country's population and the country's rate of economic growth. . . .The magnitude of the effects is truly large. For the United States, Hanushek and Woessmann (2010) calculate that the present value of increased Gross Domestic Product (GDP) from improving scores by 0.25 standard deviations would be $44 trillion" (p. 19).

This, alone, has implications for communities and businesses.

There has been a great deal of discussion about the need to supply classrooms with sufficient numbers of highly qualified new teachers to meet the potential demand (Hanushek, 2010, p. 1). At the same time, there has been an emphasis on improving the quality of the current teaching force. Hanushek stated, "US achievement could reach that in Canada and Finland if we replaced with average teachers the least effective 8-12 percent of teachers" (2011, p. 42). Some economic data has emerged, however, indicating that the system could not handle the costs associated with replacing all the low-performing teachers with high-performing replacements, even if they were available. A more reasonable solution seems to be better training. Yeh and Ritter (2009) posited, "It may be more effective to shift the entire distribution of teacher performance through a fundamental advance in technology and knowledge" (p. 426). Thus, putting systems in place to train and support an existing teaching force may be equally as important to address Goldhaber and Anthony's (2004) concern that "a growing body of research shows that the quality of the teacher in the classroom is the most important schooling factor predicting student outcomes," (p. 4).

Hanushek (2011) has written extensively about the impact of students' achievement levels on their future earnings, noting that projections provide "surprisingly precise estimates of the impact of students' achievement levels on their lifetime earnings" (p. 41). Yeh and Ritter (2009) observed that "Economic studies suggest the differences among teachers contribute to significant differences in student achievement" (p. 426). Rockoff's 2004 report equated the difference as a one-standard-deviation increase in teacher quality raises test scores by approximately .1 standard deviation in reading and

mathematics. Stronge et al. (2008) wrote that "More can be done to improve education by improving the effectiveness of teachers than by any other single factor" (p. 168). If the empirical evidence suggests that raising teacher quality is the key to improving student outcomes, there would appear to be ethical and personal responsibilities associated with providing feedback to teachers and helping them improve. By doing so, leaders commit to the growth of teachers, which subsequently translates to academic gains for students.

Overall, multiple sources are needed to determine teacher effectiveness, as no single indicator offers sufficiently strong evidence regarding performance. "Classroom observations, teacher examinations, and student performance measures considered independently offer a disjointed and narrow view of what constitutes effective teaching. The most accurate assessment of a teacher's performance requires input from multiple sources" (Flowers & Hancock, 2003, p. 162).

Kyriakides,, Demetriou, and Charalmbous (2006) observed that "It should be acknowledged that most teacher effectiveness studies have mainly elaborated on the classroom activities, failing to take into consideration other school factors" (p.18) which, although not at the magnitude of teachers, impacts classrooms. Table 1 provides a listing of the dimensions of teacher effectiveness and the research base for each provided by Strong et al. (2008). Though, this summary of teacher effectiveness dimensions and related research extends beyond the narrower focus considered in the present research, it is helpful in understanding the complexity of teacher effectiveness.

Table 1

Teacher Effectiveness Dimensions and Representative Research Base

Dimensions	Representative Research Base
Instruction	
Focus on instruction	Allington, 2002; Darling-Hammond, 2000; Johnson, 1997;Wenglinsky, 2000.
Expectations for achievement	Peart & Campbell, 1999; Wenglinsky, 2002
Planning for instruction	Good & Brophy, 1997; Jay, 2002; Shellard & Protheroe, 2000.
Range of strategies	Pressley et al., 2004; Walsh & Sattes, 2005; Weiss et al., 2003.
Questioning	Eisner, 2003/2004; Peart & Campbell, 1999; Sternberg, 2003; Zahorik et al., 2003.
Student engagement	Cawelti, 2004; Walsh & Sattes, 2005; Wenglinsky, 2002.
Student assessment	
Monitor student progress	Cotton, 2000; Foegen et al., 2007; Janisch & Johnson, 2003; Yesseldyke & Bolt, 2007.
Learning environment	
Classroom management	Johnson, 1997; Marzano et al., 2003; Pressley et al., 2004; Wang et al., 1993.
Organization	McLeod et al. 2003; Zahorik et al., 2003.
Personal qualities	
Caring	Boyle-Baise, 2005; Collinson et al., 1999.
Fairness and respect	McBer, 2000; Peart & Campbell, 1999.
Interactions with students	Corbett & Wilson, 2002; Cruickshank & Haefele, 2001;Darling-Hammond, 2001; Peart & Campbell, 1999.
Enthusiasm and motivation	Rowan et al., 1997; Quek, 2005.
Attitude toward teaching	Hamre & Pianta, 2005; Southeast Center for Teaching Quality, 2003.
Reflective practice	Cruickshank & Haefele, 2001; Good & Brophy, 1997.

Source. Stronge, J. H., Ward, T. J., Tucker, P.D., & Hindman, J. L. (2008). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education 20* (3-4), 165-184.

Value-Added Modeling (VAM)

There are six components related to the concept of value-added modeling (VAM). In this section, the history of VAM is presented along with a clear definition of VAM and a description of various VAM models. Outcome data from previous studies on the topic are reviewed and the advantages and disadvantages of using VAM are considered. Finally, summary of the review of the literature considering economic impact of implementing VAM is shared.

Although there has been recent controversy surrounding value-added models (VAM), the history is longer than one might expect. In the 1970s, educational economist Hanushek argued that it should be possible to judge the effectiveness of teachers by measuring the learning gains of their students (Asay & Schafer, 2013). However, the real philosophical underpinning of VAM started when Tyler proposed "evaluation should be a process of comparison between stated objectives and actual outcomes" (Sanders & Horn, 1994, p. 301). Louisiana was the first state in the nation to develop and implement a VAM model to assess teacher preparation programs in their state, but Sanders, an educational statistician, put Hanushek's original idea into practice when he introduced the first value-added assessment models in Tennessee in 1992 (Asay & Schafer, 2013). This made Tennessee "the first state to implement a value-added model as the basis of a school accountability program" (Doran & Izumi, 2004, p. 3). Louisiana subsequently adapted their original model to create a teacher evaluation model for practicing teachers.

Under No Child Left Behind (NCLB), signed into law by President George W. Bush on January 8, 2002, value-added assessments gained in popularity across the nation

as both researchers and educators recognized them as a new opportunity to ensure that all students benefited from an effective teacher (Asay & Schafer, 2013). Race to the Top (RTTP) funding soon brought other states into the ranks to meet the requirements of the grant.

Chetty, Friedman, & Rockoff (2011) expressed the belief that whether VAM is impactful or not can be resolved by answering two questions: (a) Is student achievement a result of the teacher or student grouping? and (b) Are highly effective teachers truly improving learning or just better at teaching to the test?

Goldhaber and Theobald (2009) defined value-added models in the following way: "Value-added models are statistical models that generally try to isolate the contribution to student test scores by individual teachers or schools from factors outside the school's or teacher's control" (p. 3). Doran (2003) and Goe et al. (2008) noted that VAM analysis seeks to quantify school factors that contribute to students' learning growth. Papay (2011) discussed underlying assumptions when he said, "All value-added models rely on the assumption that teacher effectiveness can be estimated reliably and validly through student achievement tests," (p. 168). Gagne (2011) expanded further, stating that such models were "based on complex statistics that attempt to measure a teacher's impact on students' academic growth over time," (p. 4). According to Doran (2003), "Value added analysis, combined with other valid indicators, can more reliably assess school quality without punishing or rewarding schools for preexisting differences related to student background and other non-school related factors" (p. 57).

An original version of an early value-added method based on student test scores, listed 44 variables (Sanders & Horn, 1994). Ballou (2010) viewed the central idea of value-added assessment as straight-forward: educators are to be evaluated based on the progress of their students, or the difference between incoming and outgoing levels of achievement.

In the State of Florida, a teacher's value-added score "reflects the average amount of learning growth of the teacher's students above or below the expected learning growth of similar students in the state, using the factors accounted for in the model" (Florida Department of Education, 2014). The Florida Value Added Model (VAM) was developed and recommended by the Student Growth Implementation Committee (SGIC) after reviewing eight different types of models used around the country.

> Goldhaber and Theobald (2009) used four categories in discussing VAM models: "models that do not control for student background; models that do control for student background; models that compare teachers within rather than across schools; and student growth percentile (SGP) models, which measure the achievement of individual students compared to other students with similar test score histories" (p. 3).

According to Goldhaber and Theobald (2009), these models vary in the way they account for student background and resources available to students. Table 2 displays the major vendors and models most frequently associated with value-added models at the time of the present study.

Table 2

Large Vendors That Estimate Teacher Effectiveness Using Student Test Scores

Vendor	Name of Model	Brief Description
American Institutes for Research (AIR)	Varied	In most situations, models control for student background.
Mathematica	Varied	In most situations, models control for student background.
National Center for the Improvement of Educational Assessment (NCIEA)	Student Growth Percentile (SGP) Models	Models a descriptive measure of student growth within a teacher's classroom.
SAS	EVAAS	Models control for prior test scores but not other student background variables.
Value Added Research Center (VARC)	Varied	In most situations, models control for student background.

Source. Goldhaber, D., & Theobald, R. (2009). *Do Different Value-Added Models Tell Us the Same Things*? Carnegie Knowledge Network, p. 9.

In the large urban Central Florida school district used for this study, the following student characteristics were used to adjust VAM scores and to control for the amount of expected growth in the study year:

Up to two prior years of achievement scores

Number of subject-related courses in which the student is enrolled

Students with disabilities status

English language learners status

Gifted status

Daily attendance

Mobility (number of transitions)

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

Classroom characteristics

Class size

Homogeneity of students' entering test scores in the class

Student characteristic such as gender, race, ethnicity, and socio-economic status are not included in the VAM, because under Florida law the Student Success Act specifically prohibits their inclusion into the model (Orange County Public Schools, 2013).

Districts and schools themselves impact student learning that may be attributed to teachers by inequitably distributing talent across schools (Goldhaber & Theobald, 2009). Chetty, Friedman, & Rockoff, (2013, p. 26) found that (a) the long-term impacts of teacher value-added measures were slightly larger for females than males; (b) improvements in the quality of English teachers had larger impacts than improvements in the quality of mathematics teachers; and (c) the impacts of value-added measures were roughly constant in percentage terms by parents' income.

Goldhaber & Theobald, (2009) found that teachers of advantaged students benefited from models that did not control for student background factors. In contrast, teachers of disadvantaged students benefited from models that did not control for student background factors even though a class showed less actual growth. This was attributed to the lack of control of most models for covariates such as race and poverty. Policy makers

have avoided including these covariates because the model "would expect low-income students to show lesser gains than high-income students" (Goldhaber & Theobald, 2009, p. 8). These researchers did, however, find a strong correlation between statistical models that did not account for student background factors and estimates from value-added models that controlled for student backgrounds when measures of prior student achievement was included (Goldhaber & Theobald, 2009).

There are several arguments for and against the value-added models, and it is important to understand both sides of the issue. Those who have advocated for value-added measures assert that VAM focuses on individual-level learning as opposed to group-level learning, (Asay & Schafer, 2013) making the system one that is more individualized and differentiated. It also allows for greater flexibility, because both academic achievement and academic growth is important (Ready, 2013). Sanders and Horn (1994) had earlier noted that "By focusing on outcomes rather than the processes by which they are achieved, teacher and schools are free to use whatever methods prove practical in achieving student academic progress" (p. 301).

In addition, VAM provides strong incentives for teachers to teach to all students regardless of individual abilities (Asay & Schafer, 2013). Teachers stand to gain as much from high performing students as low performing students based on their trajectory of growth. Unlike state assessments that traditionally have focused on measuring minimal proficiency levels, in a VAM system there is no particular disadvantage to being assigned subgroups of poor, or minority students (Ballou, 2010; Ready, 2013). A flaw of traditional school accountability systems has been that they disproportionately punish

socioeconomically disadvantaged schools (Ready, 2013). VAM statistically isolates the effects of teachers from various non-school factors out of the control of the teacher and school but could impact student learning. These could include student intelligence, family socio-economic level, parent involvement and neighborhood characteristics (Asay & Schafer, 2013). Ultimately, the model has been able to determine that "students in some teachers classrooms score higher than their previous scores would have predicted" (Goe et al., 2008, p. 5).

Those opposed to the use of VAM have cited several disadvantages to the models. An issue raised is that VAM does not take into account some of the confounding influences that impact resulting scores, e.g., the non-random assignment of students to teachers. Critics have argued that VAM does not account for selection biases, e.g., that teachers do not have an equal chance of being assigned any student in the district of the appropriate grade and subject (Ballou, 2010) and that some schools attract more high-ability students with more supportive families and more positive neighborhood contexts than other schools (Asay & Schafer, 2013). Scherrer (2011) provided an extended explanation of the impact of the lack of randomness in assigning students to schools:

> Most neighborhoods in the United States are strikingly homogenous, and the schools that house the children in each neighborhood are as well. Simply stated, children are not randomly assigned to schools. Advantaged students usually find themselves with other advantaged students, and disadvantaged students find themselves with other disadvantaged students. These nonrandom living arrangements introduce many factors (e.g., families with the most resources tend

> to move to school districts with the best teachers, better teachers gravitate toward schools with more resources, differences in social capital) that all complicate the measurement of the value a teacher adds to achievement. Most VAM attempts to "control" for these differences and allow for comparisons between "like" populations. But precision in matching is extremely difficult. For example, qualification for free/reduced lunch often labels a child as having low socioeconomic status. Beyond this single variable, it is difficult to obtain (on a large scale) information about a child's family that would tell more about her or his disadvantagedness (e.g., parents' education level). Disadvantaged is a relative term, and not all disadvantaged children are equally disadvantaged (p. 127).

There are practical realities to teaching in most public schools which include having to think creatively about staffing. The "one size fits all" design of VAM tends to assume students are taught by one teacher throughout the day, without taking into account the variety of groupings where students switch from teacher to teacher throughout the day (Asay & Schafer, 2013).

According to Goe et al. (2008), "The validity of using VAM for measuring teacher effectiveness is dependent in part on whether the statistical models are correctly specified and whether the inferences drawn are appropriate and defensible" (p. 47). Goldhaber & Theobald (2009) found that "even when correlations between models are high, different models will categorize many teachers differently" (p. 2). Ballou (2010) commented on the imprecision with which teacher effects are estimated, noting that it contributes to instability of teacher's estimated value-added across years. Goe et al.

(2008) observed, “Even if teachers could be cloned, the teaching context (students, curriculum, resources, parental support, school leadership, etc.) results would vary” (p. 47). VAM opponents have also cited the difficulty in establishing learning trajectories for transient students or those with varied school or teacher experiences, concluding the task is fraught with complexities. These issues have been likely to bring the issue of fairness into the debate about using VAM scores in teacher evaluation.

There is an economic impact of using VAM to be considered, as well. In their study, Chetty et al. (2011; 2013) investigated more than one million children and found that being assigned to high-value-added teachers had substantial impacts on a broad range of outcomes for students. Findings included: (a) increased probability of attending college and also increased quality of institutions students attend; (b) improved earning potential and trajectories throughout their 20s, (c) reduced chance of teenage pregnancy, (d) improved living conditions based on neighborhood residence throughout adulthood; and (e) increased ability to contribute to retirement savings plans. According to Chetty et al. (2013), “Replacing a teacher whose value-added is in the bottom 5% with an average teacher would increase the present value of students' lifetime income by approximately $250,000 per classroom” (p. 3).

The debate has continued in regard to VAM with some experts arguing that VAM is unreliable with a margin of error too big to justify the risk to teachers and others countering with their view that VAM provides valuable feedback and is a much better tool than the subjective tools that have been in place for decades. All sides agree that VAM should never be used as the single measure of teacher effectiveness (Scherrer,

2011). Papay (2011) cautioned high-stakes decision-makers to think carefully about the consequences of not recognizing all the variables within a school and a classroom that could affect effectiveness outcomes. Most authorities agree that the changes needed in measuring teacher quality will only come with alterations to the way teacher evaluation is conducted. The time for accountability for results has arrived and it is recognized that, "Any effort to create a quality teaching force should inarguably include a system that holds teachers accountable; teachers do indeed have a large effect on student outcomes" (Scherrer, 2011, p. 123).

Feedback

Feedback as a catalyst for change has become a topic of great interest, especially as it pertains to teacher performance outcomes. Six areas of feedback are discussed in this section to clarify its meaning within the context of this research: (a) an exploration of the many definitions and interpretations of feedback; (b) descriptions of types of feedback including positive feedback, negative feedback, formative feedback, facilitative feedback, directive feedback, descriptive feedback, prescriptive feedback, targeted feedback and feedback alignment; (c) feedback within evaluation, (d) the frequency of feedback; (e) the interpretation of feedback; and (f) feedback related to results.

According to Latham and Locke (1991), "Few concepts in psychology have been written about more uncritically and incorrectly than that of feedback. Actually, feedback is only information, that is, data, and as such has not necessary consequences at all" (p. 224). Ovando (1992), using Roget's Thesaurus, defined feedback as being associated with a "response, especially to one in authority about an activity" (p. 3). Ovando (1992)

also referenced Bloom's suggestion that "feedback can reveal errors in learning shortly after they occur. . . a self-correcting system so that errors made at one time can be corrected before they are compounded with later errors" (p. 3). Generally, feedback has been defined as providing information about the gap that exists between the actual level of performance and the desired level of performance, as well as the actions needed to close the gap. Cognitive psychologists who study expert performance have found that high-quality, targeted, immediate feedback is necessary to reach high levels of performance in any field (Jerald, 2012). It plays a decisive role in learning and development, within and beyond formal educational settings, and if done well, creates faster, more effective learning by guiding progress and giving precise information about what can be done in order to improve (Carless, 2006). More specifically, formative feedback is "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (Shute, 2008, p. 154). Shute also states that feedback reduces the uncertainty within the feedback gap, reduces the cognitive load of learners and helps to correct errors or misconceptions in thinking (2008, p. 157). Feedback is central to the development of effective learning, yet sadly it has been grossly underutilized even though it is recognized that most people are starved for effective feedback.

Feedback has many conditions, connotations and applications but "accurate feedback based on observation instruments can be a powerful resource for improving teaching and learning" (Jerald, 2012, p. 9). Hattie and Timperley (2007) asserted that, "Feedback needs to provide information specifically related to the task or process of

learning that fills the gap between what is understood and what is aimed to be understood" (p. 82). Ramaprasad (1983) defined feedback as "information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way," (p. 4) and Fernandez-Toro, Truman & Walker (2013) further clarified this as the "gap between a learner's present and desired level of knowledge, understanding and skill, together with information about the action necessary to close this gap" (p. 817). The action taken to close this gap is frequently called a feedback intervention (FI). Conflicting evidence exists about the effectiveness of providing such action. "FIs have highly variable effects on performance, such that in some conditions FIs improve performance. In other conditions FIs have no apparent effects on performance, and in yet others FIs debilitate performance" (Kluger & DeNisi, 1996, p. 254). This may be because there are limitations to this type of feedback including cues, personality, situational variables, and task characteristics like time constraints, complexity, novelty, and duration (Kluger & DeNisi, 1996).

Although athletes and musicians often receive regular doses of high-quality feedback, most professionals do not. Herold and Greller (1977) wrote that feedback was central to issues of training, performance, motivation and satisfaction. Fedor and Buckley (1987) observed that organizational members, like artists and athletes, have the right to be informed about the quality of their performance so they can self-correct if needed. Cognitive psychologists have suggested that feedback is the key to reaching high levels of performance in any field, including education (Jerald, 2012). According to Hattie and Timperley (2007), effective feedback must answer three major questions:

"Where am I going? (What are my goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?)" (p. 82). Feedback provides the learner with two types of information: verification (simple judgment of correctness) or elaboration (guiding cues toward a correct answer)" (Shute, 2008 p. 158).

In addition to knowing what feedback is and is not, the various types of feedback, how to give it, and how to receive it become crucial. Feedback can take several forms, the first of which researchers simply call positive feedback or praise. "It is understandable why praise would be considered positive feedback, but it is more interesting to note that receipt of information indicating a lack of adequate performance is viewed favorably," possibly because this type of feedback is perceived as useful and viewed as a supervisor showing concern (Geddes, & Linnehan, 1996). Negative feedback, in contrast, is often ambiguous and complex, with distorted messages (Geddes & Linnehan, 1996). Organizations should recognize that "negative feedback can have a disastrous effect on persons with low self-esteem, and managers should seek ways to minimize this effect" (DeNisi & Kluger, 2000, p. 135).

Other types of feedback include facilitative feedback, directive feedback and formative feedback. Facilitative feedback is feedback given to the learner with guidance and cues while directive feedback is providing corrective information. "Conventional wisdom suggests that facilitative feedback would enhance learning more than directive feedback, and yet this is not necessarily the case" (Shute, 2008, p. 163). The lack of clear outcomes may be due, in no small part, to the variation is learners. This is where

the concept of formative feedback is important. Formative feedback should "take into consideration instructional context as well as characteristics of the learner to provide effective feedback for complex learning tasks" (Shute, 2008, p. 172-173). Furthermore, Learning Sciences International has coined the phrases, descriptive feedback and prescriptive feedback. These terms are similar to Shute's (2008) terms and used as a way to determine how feedback is worded to describe behavior or provide guidance for improvement. Shute's feedback types are described in Table 3.

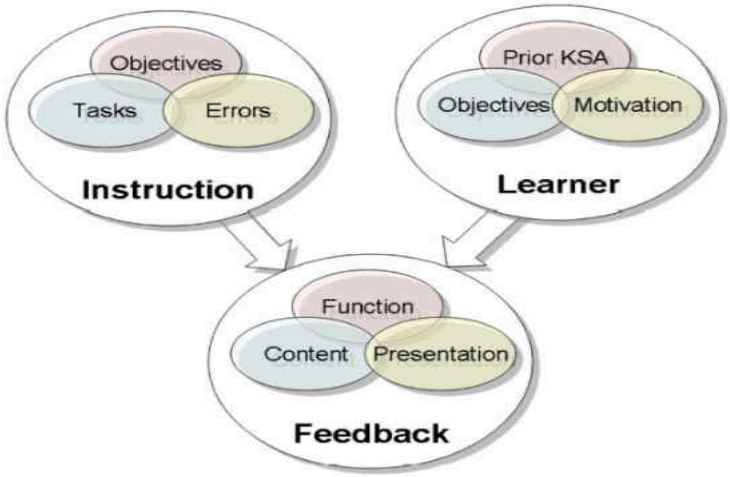
Table 3

Feedback Types Arrayed Loosely by Complexity

Feedback type	Description
No feedback	Refers to conditions where the learner is presented a question and is required to respond, but there is no indication as to the correctness of the learner's response.
Verification	Also called "knowledge of results" or "knowledge of outcome." It informs the learners about the correctness of their responses (e.g. right-wrong, or overall percentage correct).
Correct response	Also known as "knowledge of correct response." Informs the learner of the correct answer to a specific problem, with no additional information.
Try again	Also known as "repeat-until-correct" feedback. It informs the learner about an incorrect response and allows the learner one of more attempts to answer it.
Error flagging	Also known as "location of mistakes." Error flagging highlights errors in a solution, without giving correct answer.
Elaborated	General term relating to the provision of an explanation about why a specific response was correct or not and may allow the learner to review part of the instruction. It may or may not present the correct answer.
Attribute isolation	Elaborated feedback that presents information addressing central attributes of the target concept of skill being studied.
Topic contingent	Elaborated feedback providing the learner with information relating to the target topic currently being studied. May entail simply reteaching material.
Response contingent	Elaborated feedback that focuses on the learner's specific response. It may describe why the incorrect answer is wrong and why the correct answer is correct. This does not use formal error analysis.
Hints/cues/prompts	Elaborated feedback guiding the learner in the right direction, e.g., strategic hint on what to do next or a worked example or demonstration. Avoids explicitly presenting the correct answer.
Bugs/misconceptions	Elaborated feedback requiring error analysis and diagnosis. It provides information about the learner's specific errors or misconceptions (e.g., what is wrong and why)
Informative tutoring	The most elaborated feedback (from Narciss & Huth, 2004). This presents verification feedback, error flagging and strategic hints on how to proceed. The correct answer is not usually provided.

Source. Shute, V. (2008, March). Focus on Formative Feedback. *Review of Educational Research*. 78(1), p. 160.

Researchers have developed a prescription for providing effective feedback, outlined by the following characteristics: (a) be descriptive rather than evaluative; (b) be specific rather than general; (c) take into account the needs of both the receiver and the giver of feedback; (d) be directed toward behavior that the receiver can control; (e) be solicited rather than imposed; and (f) be well-timed (immediate) and checked to insure clear communication. Feedback should also be as positive, specific, and timely as possible (Fedor & Buckley, 1987, p. 171). Likewise, Shute (2008) proposed three main elements in feedback: (a) the content of the feedback; (b) the function of the feedback; and (c) the presentation of the feedback components (p. 173). This includes objectives related to curriculum and content, cognitive operations, metacognitive skills, background knowledge, skill level and motivation of the learner. Figure 1 illustrates this concept.



Source. Shute, V. (2008, March). Focus on Formative Feedback. *Review of Educational Research*. 78(1), p. 173.

Figure 1. Factors interacting with feedback to influence learning.

In addition, Ovando (1992) created the following list of distinguishing characteristics of constructive feedback in education:

- Relevant: Addresses student and teacher specific achievements, needs, and interests as well as specific learning and teaching behaviors;
- Immediate: Provided as soon as information about student and teacher performance is available;
- Factual: Based on actual student achievement (performance on a test, assignment, or project) and teacher's instructional behaviors;
- Helpful: Provides suggestions for improvement of teaching and learning;
- Confidential: Given directly to student or teacher without an intermediary;
- Respectful: of student's and teacher's integrity and needs;
- Tailored: Designed to meet individual student or teacher's specific needs and circumstances;
- Encouraging: Motivates student and teacher to continue and to increase teaching and learning efforts (p. 5)

Some researchers have indicated that feedback provided by managers is often lacking in specificity, harsh in tone, delivered in an untimely manner, or worst of all, simply not provided, (London, 1997). Cognitive psychologists who study expert performance have found that high quality, targeted, immediate feedback is necessary to reach high levels of performance in any field, (Jerald, 2012). "Feedback serves a purpose in organization; it may be stabilization, control, growth or change," (Ramaprasad, 1983, p. 5). This is because feedback can reinforce, incentivize, reduce role ambiguity, and

improve performance (Fedor & Buckley, 1987). Frase and Streshly (1994) expressed their opinion that whether it is positive or negative, accurate and straightforward feedback regarding performance is crucial for improvement. Although positive feedback is more readily accepted, negative feedback is sometimes necessary for corrective actions to occur in a timely fashion (Fedor & Buckley, 1987). It can be argued that frequency will affect, in yet to be determined ways, the interpretation of even these simple patterns of positive and negative feedback (Fedor & Buckley, 1987, p. 178). Overall however, "more specific feedback may be generally better than less specific feedback", but it should not be too wordy or complex (Shute, 2008, p. 159).

As it pertains to teacher evaluation, feedback can be quantitative or qualitative; however, it is more difficult to capture the data on the quality of feedback as much of it is subjective. Quantitative feedback is measured by capturing a teacher's overall skillset over an extended period of time, and qualitative feedback takes the form of coaching following an observation (Jerald, 2012). As part of the evaluation cycle, feedback provides information regarding performance and "is information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, met-cognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies," (Hattie & Timperley, 2007, p. 82). Researchers have reported for decades on the inequalities that exist in student learning gains from school to school and classroom to classroom, with little collective effort given to providing high-quality, targeted feedback and instructional quality matters (Kane & Staiger, 2012). Recent findings by Taylor and Tyler (2011) have indicated that providing clear feedback

to teachers leads to substantial increases in student outcomes. Ovando (2004) referred to a U.S. Department of Education Office of Innovation and Improvement report suggesting that effective principals who "analyze instruction and student learning through regular classroom observations can provide detailed feedback to teachers that supports instructional improvement" (p. 172). With these types of findings, schools have been urged to look for ways to facilitate development by utilizing feedback during classroom observations in a more targeted way (Kane & Staiger, 2012). As Ovando observed in 2004, however, training is required to deliver the type of constructive written feedback needed for growth. Shute (2008) supported this by warning that non-specific feedback may lead to uncertainty for how to respond or require greater information processing (p. 157), adding that feedback should provide learners with "information about their progress toward a desired goal" (p. 161).

Additionally, Shute (2008) discussed the importance of scaffolding feedback. This is important because it motivates interest, simplifies tasks, provides direction, indicates the difference between current work and the standard, reduces frustration and risk as well as modeling and clearly defining expectations, (p. 163). "Like training wheels, scaffolding enables learners to do more advanced activities and to engage in more advanced thinking and problem solving than they could without such help" (Shute, 2008, p. 162).

Teacher evaluation methods, i.e., classroom observations, principal evaluation, instructional artifacts, portfolios, teacher self-report measures, student surveys and the value-added model, vary. There can also be interpretation within these models both by

the giver and the receiver (Goe et al., 2008). Clear and accurate communication of progress shapes the influence of feedback and can impact the receiver's interpretation of that information. Simply making an employee aware of a shortfall in performance is not feedback. It becomes feedback when the awareness translates into action, (Ramaprasad, 1983). As with any system, variability exists. It is when this variability impacts results that reliability is compromised. Goe et al. (2008) expressed concerns about validity in teacher evaluation as well as other measurement concerns such as comprehensiveness, generality, utility, practicality, reliability, and credibility. Failure to convey specific negative feedback can lead to employees' incorrect beliefs that their performance is acceptable and can result in legal problems for an organization, (Sawyer, Hollis-Sawyer & Pokryfke, 2002). Relationship bias can also convolute accurate feedback. In the MET Study Policy Brief, *Ensuring Fair and Reliable Measures of Effective Teaching,* it was recommended that teachers receive additional observations from observers outside the teachers schools to combat this bias. Furthermore, the reluctance to provide accurate feedback or being lenient with appraisals can have costly consequences for students and the economy as a whole. Although awareness of this reluctance "to deliver bad news" is widespread, the issue continues to persist. According to Frase and Streshly (1994), few school districts have fulfilled their responsibility to provide accurate feedback to teachers even though researchers like Ovando (2004) have determined that "principals are in a key position to influence the teaching and learning process and that feedback is an important component of such influence" (p. 171).

What remains unclear is how frequently feedback should be provided to organizational members to achieve maximum results (Fedor & Buckley, 1987). Frase and Streshly (1994) found that successful principals spend 40-50% of their day in classrooms, allowing them to provide worthwhile and timely feedback to teachers.. Increasing feedback frequency is a relatively inexpensive method for better utilizing a sometimes scarce organizational resource (Fedor & Buckley, 1987). Over the short run, there may be a linear relationship between feedback frequency and positive organizational outcomes (Fedor & Buckley, 1987).

When DeNisi and Kluger (2000) reviewed the literature on various types of feedback interventions dating from the turn of the century, they discovered several inconsistencies concerning the effectiveness of feedback. "Some early experiments found that feedback improved performance for some performance indicators, but actually hurt performance for other indicators" (DeNisi & Kluger, 2000, p. 130).

Relationships also play an important part in how feedback is received or interpreted. This is because "feedback is a social process in which elements, such as discourse, power and emotion, impact on how messages can be interpreted" (Carless, 2006, p. 221). If no relationship building has occurred, the credibility of the persons giving the feedback may be compromised; and the message they are giving could be lost. The goal of feedback is for it to be meaningful. "Feedback is constructive when it offers concrete information that can be used. The intent is to help (i.e. maintain, correct, or improve behavior). It is provided in such a way that it is used by the recipient," (London, 1997, p. 513). Practitioners need to be aware that "more is not always better" (Fedor &

Buckley, 187, p. 179) in reference to feedback, and the perceptions of the recipients are important so that what is intended is perceived.

Although "feedback has emerged as a means to facilitate the learning process as well as teaching performance" (Ovando,1992, p. 2), the literature on the subject has been contradictory, recognizing that changing teaching behavior is a difficult undertaking (Ilgen, Fisher, & Taylor, 1979). A further complication has been related to determining if feedback results are different for learning a skill or managing employees (Kluger & DeNisi, 1998).

In her research, Ovando (1992) showed a positive relationship between feedback and student learning, indicating the importance of implementing a "systematic feedback process which may lead not only to effective teaching, but more importantly to successful learning" (p. 2.). Later, she wrote, "According to others, a systematic feedback process aims at enhancing both students' learning and teachers' delivery of instruction so that learning outcomes can be achieved" (Ovando, 2004, p. 173). Conversely, DeNisi and Kluger (2000) found in a meta-analysis "a modest, but positive effect of feedback on performance overall (fewer than one-half of one standard deviation improvement in performance), but 38 percent of the feedback effects were actually negative" (p. 130). In an earlier statement, Balcazar, Hopkins, & Suarez (1985) agreed, writing, "Feedback does not uniformly improve performance (p. 65). As a result of the conflicting results, DeNisi and Kluger (2000) concluded that the answer to whether feedback works should be, "Usually, but not always" (p. 131).

The findings have not hindered progress in making recommendations for the effective use of feedback. DeNisi and Kluger (2000, p. 134) suggested the following:

> Focus on the task and task performance only, not on the person or any part of the person's self-concept.
>
> Be presented in ways that do not threaten the ego of the recipient.
>
> Include information about how to improve performance.
>
> Include a formal goal-setting plan along with the feedback.
>
> Maximize information relating to performance improvements and minimize information concerning the relative performance of others.

Additionally, constructive feedback from instructional leaders should have specific process steps, be systematically collected, and analyzed and be aligned with goals, objectives and instructional strategies (Ovando, 2004, p. 5). Furthermore, according to Shute (2008), "Researchers report that feedback is more effective when learners are given specific and clear details for how to improve rather than just indicating if their work is correct or not" (p. 157).

Summary

This study was conducted to determine the relationship between the number of classroom observations and teacher VAM scores and to identify the relationship between the types of feedback provided to teachers and student achievement outcomes as measured by VAM scores. This review of the literature provides a basis for better understanding and clarifying the components of this study. The review addressed literature and research related to the five topics associated with the research questions

which guided the study: (a) reinforcement theory, (b) teacher evaluation, including the Marzano instructional model, (c) teacher effectiveness, (d) Value-added measures of student achievement, and (e) feedback. Chapter 3 describes the methods and procedures used to conduct the research.

CHAPTER 3
METHODOLOGY

Introduction

The findings in this study were intended to inform policy on current teacher observation practices and feedback by determining if there was a relationship between the number of classroom observations and teacher VAM scores. In addition, data were analyzed to identify relationships between the types of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores. This was interpreted through the following research questions.

1. What is the frequency of classroom observations and comments for teachers, including formal, informal, and walkthrough observations?
2. What relationship, if any, exists between the frequency of observations by observers as measured by the number of classroom observations and elements scored during a school year and student achievement outcomes as measured by teacher VAM scores?

   H_{01}. There is no significant relationship between the number of classroom observations during a school year and student achievement outcomes as measured by VAM.
3. What is the frequency by level of feedback defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?

4. What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

   H_{02}. There is no significant difference between classroom observations by feedback category and student achievement outcomes as measured by VAM scores.

For the purpose of this study "teachers" were instructional personnel in a large urban school district who received an individual VAM score based on the results of the learning of students they directly instruct, had more than three years of teaching experience, and did not hold National Board Certification. Observations and feedback were provided through the Marzano Instructional Model and the statewide Value Added Model (VAM) produced by the Florida Department of Education (FDOE) with technical assistance provided by the American Institutes for Research (AIR).

Instrumentation and Sources of Data

In 2011, supported by Florida's successful Race to the Top (RTTP) application which 62 of the 67 Florida school districts had agreed to implement, Governor Rick Scott signed into law Senate Bill 736, ushering in a new era of evaluation and accountability. The bill revised "the evaluation, compensation, and employment practices for classroom teachers, other instructional personnel, and school administrators to refocus the education system on what is best for students" (S.B. 736, 2011). Further clarification of the legislation was outlined in Fla. Stat. § 1012.34 (2011) Personnel Evaluation Procedures

and Criteria, and Florida State Board of Education Rule 6A-5.030 (2011), Instructional Personnel and School Administrator Evaluations. The new evaluation criteria reflected the complexity of teaching, facilitated a system-wide common language of instruction, and supported the expectation that all teachers can increase their expertise resulting in gains in student achievement (FDOE, 2014).

The Marzano Evaluation Model

The psychometric adequacy for the teacher evaluation framework was supported by a research-based selection process. The Marzano Evaluation Model was chosen as the state model framework and was approved by the Florida Department of Education (DOE) for districts to use or adapt as their teacher evaluation model. Of the 67 counties that use teacher evaluation systems in Florida, 25 opted to use the Marzano model, 18 chose the Danielson Model, 14 selected the Educational Management Consultant Services (EMCS) model, and 11 used other models that included indicators from the state model.

The Marzano Evaluation Model is based on the accumulated results of "thousands of studies that span multiple decades" (Marzano, 2011, p. 5) and were published in several books widely accepted as research-based, each of which was a result of synthesis of research and theory. Supporting the research behind the model, Marzano (2011) stated, "Experimental/control studies have been conducted that establish more direct causal linkages with enhanced student achievement that can be made with other types of data analysis" (p. 6). Marzano (2011) observed that "Correlation studies (the more typical approach to examining the viability of a model) have also been conducted

indicating positive correlations between the elements of the model and student mathematics and reading achievement" (p. 6). In advocating for the Marzano model, Learning Sciences International (2010) wrote that "a district can transform its teacher evaluation system from an exercise in compliance into an effective engine of incremental growth, one that reflects parallel gains between teacher assessment and student performance".

The model includes four domains with a total of 60 elements, distributed as follows: Domain 1, Classroom Strategies and Behaviors, 41 elements; Domain 2, Preparing and Planning, eight elements; Domain 3, Reflecting on Teaching, five elements; and Domain 4, Collegiality and Professionalism, six elements. As stated by Marzano (2013), "Given that forty-one of the sixty elements in the model are from Domain 1, the clear emphasis in the Marzano model is what occurs in the classroom" (p. 2). Administrators and teachers are trained in specifics of each of the elements, and scales have been developed to determine the level of proficiency in the implementation of each. Inter-rater reliability training ensures fidelity and calibration of observer interpretation within the model.

The Marzano Evaluation Model has a feedback component as does the state of Florida, i.e., State Board of Education Rule 6A-5.030-Instructional Personnel and School Administrator Evaluations (2011) which requires the following:

> Processes for providing feedback to the individual being evaluated, including a description of how the feedback will be timely and will promote the continuous

quality improvement of professional skills, and how results from the evaluation system will be used for individual professional development. (pp. 8-9)

Marzano's model met the state's requirement. In addition, the commitment made by the chosen school district to mandate inter-rater reliability training for any observer speaks to the reliability and validity of the process.

Student Performance Data

Student performance data is also an important aspect of multi-faceted teacher evaluation systems. The Student Success Act (2010), Senate Bill 736 (2011), and Fla. Stat. § Section 1012.34(3)(a)1. (2011) have required that school districts implement personnel evaluations based on several criteria. Instructional practice is one measure within this equation, and the law heavily values student learning growth for the other component, stating that

> at least 50 percent of a performance evaluation must be based upon data and indicators of student learning growth assessed annually by statewide assessments or, for subjects and grade levels not measured by statewide assessments, by school district assessments as provided in s. 1008.22(8) (Fla. Stat. § 1008.22).

For teachers with fewer than three years of experience, the provision allows for student learning growth to account for only 40% of their evaluations. This system of yearly evaluations went into effect in the 2011-2012 school year, and the large urban school district studied chose to use additional flexibility provided by the Florida Department of Education to count all teachers as "first-year teachers" for the purpose of

evaluation. All teachers in the target school district received final evaluation scores that consisted of 60% instructional practice and 40% student learning growth in both 2011-12 and 2012-13.

The amount of a teacher's contribution to student learning is interpreted through a value-added score. Value-added models have been used in business, health care, education and economics. In the State of Florida, a teacher's value-added score "reflects the average amount of learning growth of the teacher's students above or below the expected learning growth of similar students in the state, using the factors accounted for in the model" (FDOE, 2014, para. 3).

The Florida Value-added Model (VAM)

As mandated by the Florida Legislature, Florida school districts must consider multiple sources of data and include indicators of performance in instructional practice. The Florida Value Added Model (VAM) was developed and recommended by the Student Growth Implementation Committee (SGIC) after reviewing eight different types of models used around the country. The Florida Department of Education convened the SGIC of stakeholders to identify the type of model to be used in Florida to meet the requirements of the Student Success Act and examine the factors that should be accounted for in Florida's value-added model. To provide technical expertise, the Department contracted with the American Institutes for Research (AIR) to help the SGIC develop the recommended model from the class of covariate adjustment models that were approved by the Commissioner of Education and subsequently adopted (Florida

Department of Education, 2014). Implied psychometric adequacy was built into the model through the SGIC process. At the time of the present study, this model, selected at the state level for use for applicable teachers, was the official measure of student learning growth for relevant teachers and was not able to be altered by the school district or the researcher.

In Florida, the VAM model measures the difference in student performance on a statewide assessment from one year to the next, accounting for specific student, classroom and school characteristics shown to impact student learning (Florida Department of Education, 2014). Florida State Board of Education Rule 6A-5.0411 (2011), Calculations of Student Learning Growth Using Statewide Assessment Data for Use in School Personnel Evaluations, outlines the formula and the factors accounted for within the model as follows:

1. The formula for measuring student learning growth beginning in the 2011-12 school year using student FCAT Reading and Mathematics results is a value-added model from the class of covariate adjustment models. A value-added model is a statistical calculation employed for the purpose of determining an individual teacher or principal's contribution to student learning. Mathematically, the formula for this model is $y_i = \mu + \sum_{g=1}^{M} \delta_g x_g + \sum_{j=1}^{K} \beta_j x_j + \theta_{(s)i} + \omega_{(ms)i} + \varepsilon_i$, where y_i denotes the test score for student *i*, δ_g is the coefficient associated with g^{th} prior test score, β_j is the coefficient associated with variable *j*, θ is the common school

component of school *k* assumed $\theta \sim N(0, \sigma_\theta^2)$, ω is the effect of teacher *m* in school *k* assumed $\omega \sim N(0, \sigma_\omega^2)$, and ε is the random error term assumed $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

2. The value-added model estimates a student's performance based on variables, which represent student, classroom, and school characteristics. The variables included in the value-added model are:

a. The number of subject-relevant courses in which the student is enrolled. This variable counts, for each student, the number of courses he or she is enrolled in that are associated with FCAT Reading and Mathematics. The courses associated with the subjects of the state assessment will be published by the Department on its website at http://www.fldoe.org/committees/sg.asp.

b. Up to two (2) prior years of achievement scores for each student. This variable captures each student's most immediate prior scale score on FCAT, as well as the student's scale score from two (2) years prior, if available.

c. The student's primary disability. This is a series of variables, each which identifies a student's primary disability.

d. The student's English Language Learner (ELL) status. This variable indicates if the student has been identified as an ELL and is enrolled in a program or receiving services that are specifically designed to meet the instructional needs of ELL students for two (2) years or fewer.

e. Gifted status. This variable indicates if the student has been identified as Gifted or not.

f. Student attendance. This variable is an indicator of the days the student was present during the school year.

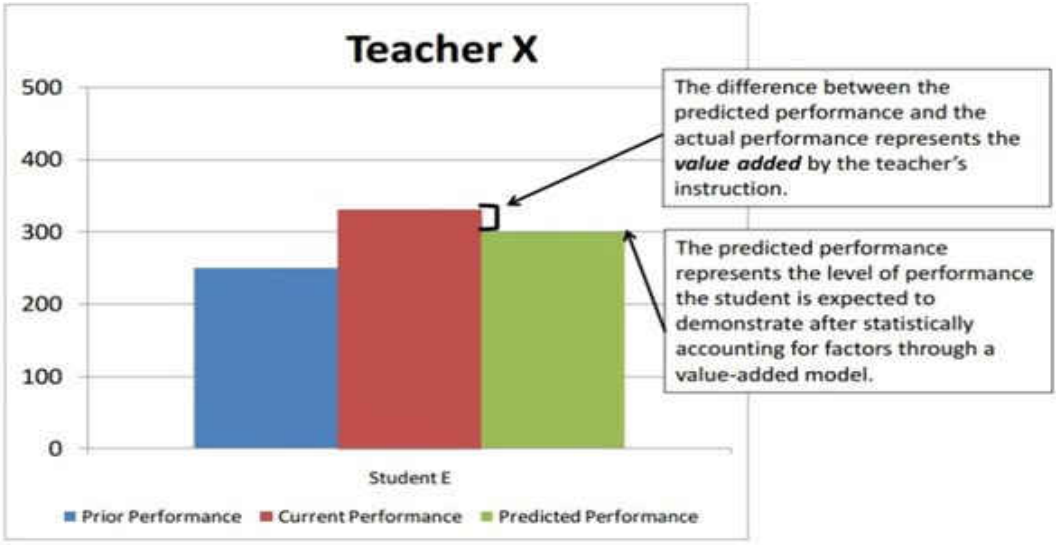
g. Student mobility. This variable is an indicator of the number of transitions a student experienced across schools within a school year.

h. Difference from modal age in grade. This variable indicates the difference in a student's age from the common age for students enrolled in the same grade across the state and is included as an indicator of retention.

i. Class size. This variable is a count of the number of students assigned to the teacher.

j. Homogeneity of students' entering test scores in the class. This variable indicates the variation within a classroom in terms of students' prior test performance, (State Board Rule 6A-5.0411, 2011).

Figure 2 provides an example of the value added by one teacher's instruction using the variables of past, current, and predicted performance of the teacher.



Source. Orange County Public Schools. (2013). *Frequently asked questions about the Florida Department of Education value added model (VAM).*

Figure 2. Value added by teacher instruction based on student performance

Population and Sample

To conduct this study, the researcher reviewed a sample of teacher observations and VAM scores in a large Florida urban school district over a one-year period spanning the 2013-2014 school year derived from a pool of more than 14,000 possible teachers. To more effectively isolate the impact of observations of student learning outcomes, the study selected the teachers based on three characteristics: (a) the number of years teaching, (b) National Board Certification, and (c) individual VAM scores in tested grade levels. Teachers with fewer than three years teaching experience were excluded from the sample, because they are expected to have lower VAM scores as newer teachers. National Board Certified Teachers were assumed to have higher VAM scores; therefore, they were also excluded. Finally, many teachers in Florida received VAM scores based

on school-level calculations. These scores are less likely to be an accurate measure of their impact on student learning than the VAM scores matched to students teachers directly instruct, which were determined to be more relevant in making a correlation between feedback and effectiveness. This enabled an assumption that observations and feedback could be directly related to effectiveness. This process reduced the original group of over 14,000 teachers to 2,718 teachers.

Data Collection

The preliminary steps of obtaining approval from the target school district (Appendix C) and from the Institutional Review Board (IRB) at the University of Central Florida (Appendix D) were completed in October 2013 and July 2014, respectively. A request for data was submitted to and approved by the Assessment Department in the large central Florida school district in the study. The data request included (a) de-identified iObservation teacher protocol data including grade level designation, scale ratings and comments and (b) de-identified VAM Data including teachers with matched scores, grade level designation, years of teaching experience, and NBCT status.

The teacher protocol and VAM data were linked through the use of a research identifier. Once these tasks central to the study were completed, the researcher created a rubric for reviewing observations with criteria and defining scales. The rubric involved the following: (a) levels of feedback alignment including point value where 1 = no feedback, 2 = unrelated or mismatched feedback; 3 = recount of observation events; 4 =

general affirmation or praise statement; 5 = reflective feedback; 6 = standardized, rote or paraphrased feedback; or 7 = specific targeted feedback for improvement.

De-identified data for the sample set of 2,718 teachers was gathered for the 2013-2014 year from iObservation by administrators observing teachers using the domains of the Marzano instructional model. The protocols were needed in order to review the comments recorded and submitted as feedback. In addition, matching de-identified VAM data was provided for the sample set of teachers for comparison, analysis and correlational study. After these data were collected, teachers were categorized as to the type of feedback received over the course of the year so as to establish the relationship between feedback and VAM.

The following rubric and accompanying definitions were used to gather data on feedback categories and levels:

Level 1-No feedback-- The observer provides no opinion in the comment section of the protocol.

Level 2-Unrelated feedback or General Statement-- The observer gives some information in the comment section but it is not relevant to the element or meaning cannot be interpreted.

Level 3-Recount of Observation Events-- This could include a narrative of what the teacher and students were doing during the observation, general statements of events, or notes the observer took to justify the rating given. In some instances the observer included statements to support the effectiveness of a strategy.

Level 4 -General Affirmation or Praise Statement--The observer either leaves a single word or phrase to indicate approval or adds a complement to the end of a recount of observation events.

Level 5-Reflective feedback--The observer asks the teacher to think about the practice or a specific element in either a general or specific way.

Level 6-Standardized feedback-- The observer uses the cut and paste option in the protocol to leave systematized feedback.

Level 7-Specific targeted feedback-- The observer leaves differentiated and meaningful statements intended to improve the impact of an instructional strategy.

Appendix E contains an example of the documentation used in completing a rubric for one teacher. The sample, shows the ratings for observation elements, feedback by category, and overall.

Data Analysis

Descriptive statistics and additional statistical tests were used to address the research questions of this study using SPSS. The following statistical procedures were used to analyze data to answer each of the research questions in the study.

Research Question 1

What is the frequency of classroom observations and comments for teachers, including formal, informal and walkthrough observations?

To answer Research Question 1 related to observation frequency, descriptive statistics were run by reviewing data for the 2,718 teachers in the group on all walkthrough, informal, and formal observations. Teachers had a value for the number of walkthrough, informal, and formal observations provided to them over the 2013-2014 school year. Descriptive statistics were used to describe the resulting data.

Research Question 2

What relationship, if any, exists between the frequency of observations by observers as measured by the number of classroom observations during a school year and student achievement outcomes as measured by teacher VAM scores?

H_{01}. There is no significant relationship between the number of classroom observations during a school year and student achievement outcomes as measured by VAM.

This research question sought to determine what relationship exists between the frequencies of observations during a school year and student achievement outcomes as measured by teacher VAM scores. The desired outcome was to determine if there was an optimum number of observations that resulted in a positive VAM score. These data were analyzed to show the relationship between the number of observations overall and in each category individually and VAM scores.

The researcher computed Pearson's r to determine the relationship between total number of classroom observations and VAM scores. In addition, statistical inference was

employed to determine if the coefficient was significantly different from zero at the 0.05 level.

Research Question 3

What is the frequency by level of feedback, defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?

To respond to Research Question 3, the aim was to calculate frequencies by level of feedback (no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback) provided by administrators to teachers during classroom observations. The desired outcome was to establish the type of feedback provided to teachers in the data set during classroom observations. A rubric was created to define the levels of feedback and can be found in the Appendix E. Descriptive statistics were run to illustrate the findings.

Research Question 4

What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

> H02. There is no significant difference between classroom observations by feedback category and student achievement outcomes as measured by VAM scores.

To respond to Research Question 4, examining the difference between the type of feedback provided to teachers and alignment with student achievement outcomes as measured by VAM scores, further analysis will be performed. The desired outcome was to learn if one of the delineated types of feedback had a greater impact on teacher VAM scores than another. After teachers were categorized, an average of all teacher VAM scores for each category was calculated and compared through a one-way analysis of variance (ANOVA).

To provide the more targeted data needed to answer Research Questions 3 and 4 related to types of feedback, all scored elements for teachers with an individual value-added score were used. Each individual walkthrough, informal and formal observation protocol was reviewed and the data were categorized, establishing levels based on the scale rating and information found in the feedback portion of the observation. Each of the feedback levels was assigned a value corresponding to the definitions assigned to the observational comments. Descriptive statistics were initially run on the data to determine overall trends in the types of feedback the set of teachers received. Feedback trends for each teacher were then reviewed to determine the predominant feedback type received. The individual ratings information was used to place each teacher in a category that described the overall pattern of feedback received. Teachers were placed in the respective feedback categories based on the feedback type that represented that majority or plurality of the feedback received.

Table 4 presents the research questions which guided the study. Also displayed in the table are the variables considered, the sources of data, and the methods of analysis.

Table 4

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

Research Questions	Qualifying/Independent Variable(s)	Dependent Variable	Source(s) of Data	Method(s) of Analysis
1. What is the frequency of classroom observations and comments for teachers, including formal, informal and walkthrough observations?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations.	Frequency of Observations	iObservation Reporting	Descriptive statistics were used to describe resulting data, which included frequency distribution for each type of observation.
2. What relationship if any exists between the frequency of observations by observers as measured by the number of classroom observations during a school year and student achievement outcomes as measured by teacher VAM scores?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Teachers with matched individual VAM scores for the 2013-2014 school year.	Value-added score results	iObservation Reporting Teacher VAM scores	Pearson's r computed to determine relationship between total number of classroom observations and VAM scores. Statistical inference conducted to determine if coefficient is significantly different from zero at 0.05 level.

Research Questions	Qualifying/Independent Variable(s)	Dependent Variable	Source(s) of Data	Method(s) of Analysis
3. What is the frequency by level of feedback, defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Implementation of the scales for protocol observation. • Data captured from the feedback section of the observation protocol.	Types of Feedback	iObservation Feedback section	Descriptive statistics, including the frequency distribution were run to illustrate findings.
4. Is there a difference between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Implementation of the scales for protocol observation. • Data captured from the feedback section of the observation protocol. • Teachers with matched individual VAM scores for the 2013-2014 school year.	Value-added score results	iObservation Feedback section Teacher VAM scores	After teachers were uniquely classified into categories, an average of all teacher VAM scores for each category was calculated and compared through one-way analysis of variance (ANOVA).

CHAPTER 4
PRESENTATION AND ANALYSIS OF DATA

Introduction

The purpose of this research study was to evaluate the observations and feedback provided in the evaluation of instructional personnel in a large, urban public school district and to examine how the frequency and quality of observations and feedback was associated with student learning growth. Chapter 4 contains descriptive and inferential analyses of quantitative and qualitative data to answer the four research questions which were used to guide the study. Narrative descriptions and supportive tables have been used to report the data analysis that was completed in response to each of the research questions.

Research Question 1

> What is the frequency of classroom observations and comments on elements scored for teachers, including formal, informal and walkthrough observations?

The school district provided information on observation and feedback provided to all teachers and a sample of teachers who also had qualifying student learning growth data. In order to be present in the sample, teachers had to (a) have three or more years of teaching experience as of the 2013-14 school year, (b) not be a Nationally Board Certified Teacher, and (c) have a student learning growth score calculated from the statewide value-added model that connected teachers to the students they directly instructed in the 2013-14 school year.

Table 5 displays the distribution of teachers by level in the entire district and the sample. For the district overall, 12,990 teachers received an instructional practice evaluation. This number may not be aligned precisely with the number of teachers in the district at any specific time because of teacher turnover during the year and administrator error in the evaluation process. These factors may have led to instructional practice results that were disregarded. Of the 12,990 teachers, 6,436 (49.55%) were elementary school teachers, 2,467 (15.99%) were middle school teachers, 2,799 (21.55%) were high school teachers, and 1,288 (9.92%) were teachers classified as other teachers. Other teachers included teachers at alternative and exceptional student education sites along with other instructional personnel who had direct contact with students outside of a traditional or charter elementary, middle, or high school.

A total of 2,718 teachers met the requirements to be included in the sample. This was 20.92% of the total population. Of the sample teachers, 1,302 (47.90%) were elementary school teachers, 949 (33.81%) were middle school teachers, 398 (14.64%) were high school teachers, and 69 (2.54%) were other teachers. The distribution of the sample teachers differed from the population primarily due to the availability of statewide student learning growth scores. Statewide student learning growth scores were only available for teachers of reading in Grades 4 through 10 and mathematics in Grades 4 through 8. This reduced the percentage of all reading and mathematics teachers available for the sample for elementary and high school years relative to middle school reading and mathematics teachers. Table 5 contains information regarding evaluated teachers in the school district and the sample by grade level.

Table 5

Evaluated Teachers in School District and Sample by Grade Level

	School District Total		Sample	
Grade Level	*f*	%	*f*	%
Elementary	6,436	49.55	1,302	47.90
Middle	2,467	18.99	949	34.92
High	2,799	21.55	393	14.64
Other	1,288	9.92	69	2.54
Total	12,990	100.00	2,718	100.00

Tables 6-12 present the comparisons of school district/all teacher results to those of the sample to determine similarity of frequency of observations and feedback. Table 6 describes the number of observations provided for all teachers and those in the sample. The number of observations for teachers ranged from 1 to 32. For all teachers and the sample, the modal number of observations was three with 40.34% of all teachers and 41.94% of sample teachers receiving three observations. The overwhelming majority of teachers in both groups received six or fewer observations. A total of 87.52% of all teachers received six or fewer observations, and 91.22% of sample teachers received six or fewer observations.

Table 6

Observations per Teacher: School District and Sample

	School District Total		Sample	
Observations	*f*	%	*f*	%
1	157	1.21	4	0.15
2	188	1.45	21	.78
3	5,240	40.34	1,140	41.94
4	2,847	21.92	724	26.64
5	1,172	9.02	362	13.32
6	1,764	13.58	228	8.39
7	840	6.47	106	3.90
8	335	2.58	52	1.91
9	151	1.16	23	0.85
10	100	0.77	13	0.48
11 or more	196	1.51	45	1.66

Table 7 presents the elements scored with comments for teachers in the school district and sample. For all teachers in the school district, 287,501 elements were scored. Of these elements, 219,446 (76.33%) were scored with comments. This compares to 59,967 elements scored and 46,849 (78.12%) comments for teachers in the sample. The highest proportion of elements scored with comments for school district teachers occurred for elementary teachers where 79.53% of elements scored were accompanied by comments. Though a slightly higher proportion of elements scored for elementary teachers in the sample received comments, the highest proportion of comments for scored elements occurred for other teachers. These teachers received comments on 85.71% of their scored elements.

Table 7

Elements Scored with Comments by Level: School District and Sample

School Level	Elements Scored *f*	Elements Scored with Comments *f*	Elements Scored with Comments %
District			
Elementary	143,148	113,846	79.53
Middle	57,007	44,306	77.72
High	61,231	42,721	69.77
Other	26,115	18,573	71.12
Total	287,501	21,446	76.33
Sample			
Elementary	30,550	24,452	80.04
Middle	12,352	9,811	79.43
High	12,900	9,016	69.89
Other	4,165	3,570	85.71
Total	59,967	46,849	78.12

In Table 8, the number and percentage of comments for elements scored are provided for all 41 elements in Domain 1 of the Marzano instructional evaluation system. Over 35% of all comments for all teachers (35.29%) and sample teachers (35.44%) were provided on five elements: "Providing clear learning goals and scales," "Tracking student progress," "Establishing classroom routines," "Reviewing content," and "Practicing skills, strategies and processes." For both the all teacher and sample teacher groups, over 10% of all comments were related to the element "Providing clear learning goals and scales." These five elements had similar percentages of total comments for all teachers and sample teachers.

Six elements accounted individually for fewer than one half of one percent of all comments. These elements were: "Using homework," "Organizing students for cognitively complex tasks," "Providing resources and guidance," "Using friendly

controversy," "Asking questions of low expectancy students," and "Probing incorrect answers with low expectancy students." As with the highest commented upon elements, the lowest commented upon elements for all teachers and sample teachers were similar.

Table 8

Total Comments by Elements Scored: School District and Sample

Comments by Element (DQ = Design Questions)	School District *f*	School District %	Sample *f*	Sample %
Communicating learning goals and feedback (DQ1)				
1. Providing clear learning goals and scales	22,085	10.06	4,744	10.13
2. Tracking student progress	11,024	5.02	2,332	4.98
3.Celebrating success	4,525	2.06	974	2.08
Establishing rules and procedures (DQ6)				
4. Establishing classroom routines	16,272	7.42	3,537	7.55
5. Organizing the classroom physical layout	6,276	2.86	1,346	2.87
Helping students interact with new knowledge (DQ2)				
6. Identifying critical information	10,815	4.93	2,235	4.77
7. Organizing students to interact with new knowledge	5,405	2.46	1,174	2.51
8. Previewing new content	5,719	2.61	1,243	2.65
9. Chunking content into 'digestible bites'	8,751	3.99	1,899	4.05
10. Processing of new information	3,762	1.71	785	1.68
11. Elaborating of new information	4,207	1.92	880	1.88
12. Recording and representing new knowledge	6,485	2.96	1,277	2.73
13. Reflecting on learning	2,533	1.15	572	1.22
Helping students practice and deepen new knowledge (DQ3)				
14. Reviewing content	14,897	6.79	3,171	6.77
15. Organizing students to practice/deepen new knowledge	10,413	4.75	2,267	4.84
16. Using homework	755	0.34	125	0.27
17. Examining similarities and differences	4,400	2.01	910	1.94
18. Examining errors in reasoning	2,301	1.05	450	0.96
19. Practicing skills, strategies, and processes	13,163	6.00	2,815	6.01
20. Revising knowledge	1,528	0.70	355	0.76
Helping students generate and test hypotheses (DQ4)				
21. Organizing students for cognitively complex tasks	689	0.31	145	0.31
22. Engaging students in cognitively complex tasks	1,248	0.57	259	0.55
23. Providing resources and guidance	879	0.40	190	0.41
Engaging students (DQ5)				
24. Noticing when students are not engaged	7,058	3.22	1,557	3.32
25. Using academic games	1,920	0.87	418	0.89
26. Managing response rates	7,078	3.23	1,434	3.06
27. Using physical movement	2,698	1,23	561	1.20
28. Maintaining a lively pace	7,162	3.26	1,577	3.37
29. Demonstrating intensity and Enthusiasm	5,196	2.37	1,079	2.30
30. Using friendly controversy	426	0.19	89	0.19
31. Providing opportunity for student talk	1,245	0.57	253	0.54
32. Presenting unusual/intriguing information	1,496	0.68	321	0.69
Recognizing adherence to rules and procedures (DQ7				
33. Demonstrating 'withitness'	8,269	3.77	1,819	3.88
34. Applying consequences for lack of adherence	4,429	2.02	960	2.05
35. Acknowledging adherence to rules/procedures	2,559	1.17	586	1.25
Establishing/maintaining effective student relationships (DQ8)				
36. Understanding students' interests/backgrounds	1,761	0.80	369	0.79
37. Using verbal and nonverbal behaviors	5,585	2.55	1,191	2.54
38. Displaying objectivity and control	1,152	0.52	253	0.54

Comments by Element (DQ = Design Questions)	School District		Sample	
	f	%	*f*	%
Communicating high expectations to students (DQ9)				
39. Demonstrating value/respect for low expectancy students	1,370	0.62	311	0.66
40. Asking questions of low expectancy students	1,099	0.50	225	0.48
41. Probing incorrect answers with Low expectancy students	811	0.39	161	0.23

The total number of comments for elements scored for elementary school, middle school, high school, and other teachers are displayed in Tables 9-12. The scored elements with the most comments differed slightly by level. For elementary teachers, the most commented on elements were the same as the overall sample: "Providing clear learning goals and scales," "Tracking student progress," "Establishing classroom routines," "Reviewing content," and "Practicing skills, strategies, and processes." For middle and high school teachers, the same elements were in the top five most commented upon, with the exception of "Tracking student progress" which was replaced by "Organizing students to practice/deepen knowledge." Other teachers had "Providing clear learning goals and scales," "Identifying critical knowledge," and "Reviewing content" in their top five most commented upon elements, similar to teachers in all other levels. However, "Identifying critical information" and "Chunking content into digestible bites" were also in the top five for these teachers.

Table 9

Total Comments by Elements Scored: All Elementary and Sample Elementary

Comments by Element (DQ = Design Questions)	All Elementary *f*	All Elementary %	Sample Elementary *f*	Sample Elementary %
Communicating learning goals and feedback (DQ1)				
1. Providing clear learning goals and scales	10,703	9.40	2,302	9.41
2. Tracking student progress	5,824	5.12	1,247	5.10
3.Celebrating success	2,523	2.22	558	2.28
Establishing rules and procedures (DQ6)				
4. Establishing classroom routines	9,655	8.48	2,109	8.63
5. Organizing the classroom physical layout	3,016	2.65	655	2.68
Helping students interact with new knowledge (DQ2)				
6. Identifying critical information	5,096	4.48	1,051	4.30
7. Organizing students to interact with new knowledge	2,936	2.58	667	2.73
8. Previewing new content	3,039	2.67	661	2.70
9. Chunking content into 'digestible bites'	4,641	4.08	1,029	4.21
10. Processing of new information	2,051	1.80	454	1.86
11. Elaborating of new information	2,254	1.98	453	1.85
12. Recording and representing new knowledge	3,334	2.93	648	2.65
13. Reflecting on learning	1,13	1.02	279	1.14
Helping students practice and deepen new knowledge (DQ3)				
14. Reviewing content	7,817	6.87	1,649	6.74
15. Organizing students to practice/deepen new knowledge	5,352	4.70	1,159	4.73
16. Using homework	204	0.18	39	0.16
17. Examining similarities and differences	2,325	2.04	514	2.10
18. Examining errors in reasoning	1,031	0.91	218	0.89
19. Practicing skills, strategies, and processes	7,430	6.53	1,565	6.40
20. Revising knowledge	638	0.56	131	0.54
Helping students generate and test hypotheses (DQ4)				
21. Organizing students for cognitively complex tasks	327	0.29	69	0.28
22. Engaging students in cognitively complex tasks	516	0.45	116	0.47
23. Providing resources and guidance	256	0.22	60	0.25
Engaging students (DQ5)				
24. Noticing when students are not engaged	3,604	3.7	797	3.26
25. Using academic games	1,041	0.91	239	0.98
26. Managing response rates	3,882	.41	739	3.02
27. Using physical movement	1,722	1.51	364	1.49
28. Maintaining a lively pace	3617	3.18	787	3.22
29. Demonstrating intensity and Enthusiasm	2,698	2.37	553	2.26
30. Using friendly controversy	150	0.13	29	0.12
31. Providing opportunity for student talk	468	0.41	102	0.42
32. Presenting unusual/intriguing information	544	0.48	132	0.54
Recognizing adherence to rules and procedures (DQ7				
33. Demonstrating 'withitness'	3,977	3.49	883	3.61
34. Applying consequences for lack of adherence	2,670	2.35	612	2.50
35. Acknowledging adherence to rules/procedures	2,001	1.76	450	1.84
Establishing/maintaining effective student relationships (DQ8)				
36. Understanding students' interests/backgrounds	500	0.44	102	0.42
37. Using verbal and nonverbal behaviors	2,610	2.29	552	2.26
38. Displaying objectivity and control	526	0.46	117	0.48

Comments by Element (DQ = Design Questions)	All Elementary		Sample Elementary	
	f	%	*f*	%
Communicating high expectations to students (DQ9)				
39. Demonstrating value/respect for low expectancy students	671	0.59	147	0.60
40. Asking questions of low expectancy students	544	0.48	117	0.48
41. Probing incorrect answers with Low expectancy students	487	0.43	100	0.41

Table 10

Total Comments by Elements Scored: All Middle and Sample Middle

Comments by Element (DQ = Design Questions)	All Middle		Sample Middle	
	f	%	*f*	%
Communicating learning goals and feedback (DQ1)				
1. Providing clear learning goals and scales	4,861	10.97	1,058	10.78
2. Tracking student progress	2,086	4.71	424	4.32
3.Celebrating success	707	1.60	143	1.46
Establishing rules and procedures (DQ6)				
4. Establishing classroom routines	3,394	7.66	789	8.04
5. Organizing the classroom physical layout	1,287	2.90	318	3.24
Helping students interact with new knowledge (DQ2)				
6. Identifying critical information	2,019	4.56	420	4.28
7. Organizing students to interact with new knowledge	1,017	2.30	232	2.36
8. Previewing new content	879	1.98	204	2.08
9. Chunking content into 'digestible bites'	1,417	3.20	300	3.06
10. Processing of new information	637	1.44	`127	1.29
11. Elaborating of new information	774	1.75	176	1.79
12. Recording and representing new knowledge	1,312	2.96	284	2.89
13. Reflecting on learning	544	1.23	121	1.23
Helping students practice and deepen new knowledge (DQ3)				
14. Reviewing content	3,051	6.89	705	7.19
15. Organizing students to practice/deepen new knowledge	2,428	5.48	543	5.53
16. Using homework	167	0.38	25	0.25
17. Examining similarities and differences	923	2.08	180	1.83
18. Examining errors in reasoning	510	1.15	100	1.02
19. Practicing skills, strategies, and processes	2,561	5.78	597	6.09
20. Revising knowledge	363	0.82	100	1.02
Helping students generate and test hypotheses (DQ4)				
21. Organizing students for cognitively complex tasks	154	0.35	33	0.34
22. Engaging students in cognitively complex tasks	335	0.76	64	0.65
23. Providing resources and guidance	176	0.40	40	0.41
Engaging students (DQ5)				
24. Noticing when students are not engaged	1,550	3.50	363	3.70
25. Using academic games	361	0.81	73	0.74
26. Managing response rates	1,371	3.09	328	3.34
27. Using physical movement	464	1.05	93	0.95
28. Maintaining a lively pace	1,598	3.61	381	3.88
29. Demonstrating intensity and Enthusiasm	941	2.12	216	2.20
30. Using friendly controversy	93	0.21	22	0.22
31. Providing opportunity for student talk	239	0.54	43	0.44
32. Presenting unusual/intriguing information	389	0.88	80	0.82
Recognizing adherence to rules and procedures (DQ7				
33. Demonstrating 'withitness'	2,118	4.78	451	4.60
34. Applying consequences for lack of adherence	813	1.83	164	1.67
35. Acknowledging adherence to rules/procedures	297	0.67	70	0.71
Establishing/maintaining effective student relationships (DQ8)				
36. Understanding students' interests/backgrounds	376	0.85	81	0.83
37. Using verbal and nonverbal behaviors	1,201	2.71	274	2.79
38. Displaying objectivity and control	250	0.56	56	0.57

Comments by Element (DQ = Design Questions)	All Middle		Sample Middle	
	f	%	*f*	%
Communicating high expectations to students (DQ9)				
39. Demonstrating value/respect for low expectancy students	245	0.55	50	0.51
40. Asking questions of low expectancy students	242	0.55	51	0.52
41. Probing incorrect answers with Low expectancy students	156	0.35	32	0.33

Table 11

Total Comments by Elements Scored: All High and Sample High

Comments by Element (DQ = Design Questions)	All High		Sample High	
	f	%	*f*	%
Communicating learning goals and feedback (DQ1)				
1. Providing clear learning goals and scales	4,332	10.14	943	10.46
2. Tracking student progress	1,788	4.19	392	4.35
3.Celebrating success	651	1.52	140	1.55
Establishing rules and procedures (DQ6)				
4. Establishing classroom routines	2,425	5.68	484	5.37
5. Organizing the classroom physical layout	1,475	3.45	279	3.09
Helping students interact with new knowledge (DQ2)				
6. Identifying critical information	2,438	5.71	521	5.78
7. Organizing students to interact with new knowledge	1,021	2.39	184	2.04
8. Previewing new content	1,045	2.45	224	2.48
9. Chunking content into 'digestible bites'	1,728	4.04	398	4.41
10. Processing of new information	646	1.51	131	1.45
11. Elaborating of new information	722	1.69	154	1.71
12. Recording and representing new knowledge	1,279	2.99	250	2.77
13. Reflecting on learning	561	1.31	133	1.48
Helping students practice and deepen new knowledge (DQ3)				
14. Reviewing content	2,845	6.66	598	6.63
15. Organizing students to practice/deepen new knowledge	2,046	4.79	443	4.91
16. Using homework	298	0.70	48	0.53
17. Examining similarities and differences	787	1.84	164	1.82
18. Examining errors in reasoning	612	1.43	114	1.26
19. Practicing skills, strategies, and processes	2,337	5.47	482	5.35
20. Revising knowledge	389	0.91	98	1.09
Helping students generate and test hypotheses (DQ4)				
21. Organizing students for cognitively complex tasks	139	0.33	34	0.38
22. Engaging students in cognitively complex tasks	279	0.65	57	0.63
23. Providing resources and guidance	351	0.82	77	0.85
Engaging students (DQ5)				
24. Noticing when students are not engaged	1,515	3.54	325	3.60
25. Using academic games	384	0.90	85	0.94
26. Managing response rates	1,433	3.35	285	3.16
27. Using physical movement	305	0.71	71	0.79
28. Maintaining a lively pace	1,426	3.34	310	3.44
29. Demonstrating intensity and Enthusiasm	1,032	2.42	219	2.43
30. Using friendly controversy	138	0.32	31	0.34
31. Providing opportunity for student talk	325	0.76	70	0.78
32. Presenting unusual/intriguing information	398	0.93	79	0.88
Recognizing adherence to rules and procedures (DQ7				
33. Demonstrating 'withitness'	1,754	4.11	391	4.34
34. Applying consequences for lack of adherence	770	1.80	146	1.62
35. Acknowledging adherence to rules/procedures	126	0.29	33	0.37
Establishing/maintaining effective student relationships (DQ8)				
36. Understanding students' interests/backgrounds	622	1.46	145	1.61
37. Using verbal and nonverbal behaviors	1,357	3.18	281	3.12
38. Displaying objectivity and control	270	0.63	52	0.58

Comments by Element (DQ = Design Questions)	All High		Sample High	
	f	%	*f*	%
Communicating high expectations to students (DQ9)				
39. Demonstrating value/respect for low expectancy students	333	0.78	89	0.99
40. Asking questions of low expectancy students	226	0.53	39	0.43
41. Probing incorrect answers with Low expectancy students	113	0.26	17	0.19

Table 12

Total Comments by Elements Scored: All Other and Sample

Comments by Element (DQ = Design Questions)	All Other *f*	All Other %	Sample Other *f*	Sample Other %
Communicating learning goals and feedback (DQ1)				
1. Providing clear learning goals and scales	2,129	11.46	441	12.35
2. Tracking student progress	1,326	7.14	269	7.54
3.Celebrating success	644	3.47	133	3.73
Establishing rules and procedures (DQ6)				
4. Establishing classroom routines	798	4.30	155	4.34
5. Organizing the classroom physical layout	498	2.68	94	2.63
Helping students interact with new knowledge (DQ2)				
6. Identifying critical information	1,262	6.79	243	6.81
7. Organizing students to interact with new knowledge	431	2.32	91	2.55
8. Previewing new content	756	4.07	154	4.31
9. Chunking content into 'digestible bites'	965	5.20	172	4.82
10. Processing of new information	428	2.30	73	2.05
11. Elaborating of new information	457	2.46	97	2.72
12. Recording and representing new knowledge	560	3.02	95	2.66
13. Reflecting on learning	265	1.43	39	1.09
Helping students practice and deepen new knowledge (DQ3)				
14. Reviewing content	1,184	6.37	219	6.13
15. Organizing students to practice/deepen new knowledge	587	3.16	125	3.50
16. Using homework	86	0.48	13	0.36
17. Examining similarities and differences	365	1.97	52	1.46
18. Examining errors in reasoning	148	0P.80	18	0.50
19. Practicing skills, strategies, and processes	835	4.50	171	4.79
20. Revising knowledge	138	0.74	26	0.73
Helping students generate and test hypotheses (DQ4)				
21. Organizing students for cognitively complex tasks	69	0.37	9	0.25
22. Engaging students in cognitively complex tasks	118	0.64	22	0.62
23. Providing resources and guidance	96	0.52	13	0.36
Engaging students (DQ5)				
24. Noticing when students are not engaged	386	2.08	72	2.02
25. Using academic games	134	0.72	21	0.59
26. Managing response rates	392	2.11	82	2.30
27. Using physical movement	207	1.11	33	0.92
28. Maintaining a lively pace	521	2.81	99	2.77
29. Demonstrating intensity and Enthusiasm	525	2.83	91	2.55
30. Using friendly controversy	45	0.24	7	0.20
31. Providing opportunity for student talk	213	1.15	38	1.06
32. Presenting unusual/intriguing information	165	0.89	30	0.84
Recognizing adherence to rules and procedures (DQ7				
33. Demonstrating 'withitness'	420	2.26	94	2.63
34. Applying consequences for lack of adherence	176	0.95	38	1.06
35. Acknowledging adherence to rules/procedures	135	0.73	33	0.92
Establishing/maintaining effective student relationships (DQ8)				
36. Understanding students' interests/backgrounds	263	1.42	41	1.15
37. Using verbal and nonverbal behaviors	417	2.25	84	2.35
38. Displaying objectivity and control	166	0.89	28	0.78

Comments by Element (DQ = Design Questions)	All Other		Sample Other	
	f	%	*f*	%
Communicating high expectations to students (DQ9)				
39. Demonstrating value/respect for low expectancy students	121	0.65	25	0.70
40. Asking questions of low expectancy students	87	0.47	18	0.50
41. Probing incorrect answers with Low expectancy students	55	0.30	12	0.30

Research Question 2

What relationship if any exists between the frequency of observations by observers as measured by the number of classroom observations during a school year and student achievement outcomes as measured by teacher VAM scores?

For inclusion in the sample, teachers were required to have a statewide value-added score. The statewide value-added scores, when standardized, describe the percentage of a year's growth more or less than average that students associated with a teacher score after an adjustment is made for student covariates. Though this information appears as a decimal when received from the state, it is presented as a percentage for ease of understanding in all tables in this chapter. For example, "2.52% below" would mean that students in a teacher's class scored 2.52% of a year's growth below comparable students across the state.

Table 13 displays average value-added scores by the number of observations per teacher. The modal number of observations for teachers in the sample was three, and over 90% of teachers received six or fewer observations. Teachers with one or two observations had below average value-added scores, 3.47% and 2.70% below average, respectively. Teachers with three through six observations had above average value-added scores, 6.21%, 6.28%, 5.52%, and 1.88% above average, respectively. There was considerably more variation in average value-added scores for teachers with more than

six observations, ranging from 26.50% below average to 24.17% above average. It is important to note that this last group consists of fewer than 10% of all teachers in the sample.

Table 13

Value-added (VAM) Scores by Observations per Teacher

Observations	Average VAM Score[a]	Teachers *f*	Teachers %
1	3.47% below	4	0.15
2	2.70% below	21	.78
3	6.21% above	1,140	41.94
4	6.28% above	724	26.64
5	5.52% above	362	13.32
6	1.88% above	228	8.39
7	3.31% below	106	3.90
8	6.95% below	52	1.91
9	8.88% above	23	0.85
10	5.06% above	13	0.48
11 or more	1.16% above	45	1.66

[a]Expressed in percentage of a year's growth above or below average of associated students' scores on statewide assessments

In order to determine if there was a relationship between the number of observations and value-added scores, Pearson's *r* correlations were calculated. To determine if statistical significance could be identified, a standard critical table was used (Steinberg 2010). This analysis is presented in Table 14.

No significant relationship existed between value-added scores and the number of observations for teachers. This lack of significance was true when examining the number of observations for all groups (school district, elementary, middle, high, or other

teachers). This indicated that there was no evidence of a relationship between the number of observations and value-added scores among teachers in the sample.

Table 14

Correlational Analysis: Pearson Correlation Coefficients between Number of Observations and Value-added Scores by School District and School Level

	Value-added Scores				
Descriptor	School District	Elementary School	Middle School	High School	Other
Number of Observations	-0.03	-0.02	-0.05	-0.04	-0.03

Note. *p < .05, **p < .01

As noted previously, the majority of scored elements (78.12%) for sample teachers contained comments. Teachers, however, varied in the percentage of comments that were scored by their observers. This information, along with average value-added scores for these teachers, are provided in Table 15. The majority of sample teachers (61.77%) received comments on between 70% and 80% of their scored elements. Only 7.29% of teachers received comments on fewer than 60% of their scored elements.

Teachers in the sample had higher value-added scores on average than all teachers in the state. As shown in Table 15, in all categories of percentage of scored elements with comments, the average value-added score was between 2.73% to 6.87% above average. The majority of teachers, those who received between 70% and 80% of their

scored elements with comments, had an average value-added score of 4.33% above average.

Table 15

Value-added Scores by Percentage of Elements with Comments

% of Elements With Comments	Average VAM Score[a]	Sample Teachers *f*	%
Less than 10%	4.81% above	6	0.22
Between 10% and 20%	4.14% above	3	0.11
Between 20% and 30%	3.27% above	13	0.48
Between 30% and 40%	6.51% above	57	2.10
Between 40% and 50%	2.73% above	37	1.36
Between 50% and 60%	5.46% above	82	3.02
Between 60% and 70%	6.87% above	345	12.69
Between 70% and 80%	4.33% above	1,679	61.77
Between 80% and 90%	5.41% above	272	10.01
Between 90% and 100%	5.71% above	224	8.24

[a]Expressed in % of a year's growth above or below average associated students scored on statewide assessments

Pearson's *r* correlations were calculated to determine the relationship between the percentage of elements with comments and value-added scores. No significant relationship was found between the value-added scores overall or at any school level and the percentage of elements scored that contained comments. The results of the correlation are displayed in Table 16.

Table 16

Correlational Analysis: Pearson Correlation Coefficients between Number of Observations and Value-added Scores by School District and School Level

	Value-added Scores				
Descriptor	School District	Elementary School	Middle School	High School	Other
Number of Observations	0.01	0.02	0.01	0.01	-0.02

Note. *p < .05, **p < .01

Research Question 3

> What is the frequency by level of feedback, defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?

For this research question, all comments in the sample were coded to determine if the feedback provided was not present, unrelated, a simple recount of observation details, general affirmations such as praise, reflective feedback, standardized feedback, or specific, targeted feedback. Additional details on these feedback types can be found in the rubrics contained in Appendix E

In order to ensure that the codes provided for comments operated reliably, an additional coder was employed to measure intercoder reliability. This additional coder received the rubric and a sample of 650 comments in order to ensure that the comment definitions were followed in coding. No scored elements where evaluators left no comments were contained in the sample to determine intercoder reliability. Of the 650

sample elements with comments, 604 (92.9%) were coded identically by both coders. This suggested that the coding executed by the researcher measured replicable findings that were clear in the comments provided by evaluators.

Table 17 describes the frequency and respective percentages of feedback types for all scored elements for the sample. Of 59,967 scored elements, 13,122 scored elements (21.88%) contained no feedback. Only 138 scored elements (0.23%) contained unrelated feedback. Over half of the scored elements, 34,682 (57.84%) contained a recount of events in the observation. Three-quarters of all scored elements received either no feedback or feedback recounting events in the observation. Almost 6% of all scored elements (3,581, 5.97%) received general affirmations that mainly took the form of praise. There were 2,778 scored elements (4.63%) that provided reflective feedback and 1,442 scored elements (2.40%) that provided standardized feedback. The most important form of feedback for improving instruction, specific targeted feedback, was provided on 4,224 scored elements which accounted for 7.04% of all scored elements.

Table 17

Frequency and Percentages of Feedback for all Elements by Feedback Type

Level	Feedback Types	*f*	% of Total Feedback
1	No feedback	13,122	21.88
2	Unrelated feedback	138	0.23
3	Recount of observation events	34,682	57.84
4	General affirmations (praise)	3,581	5.97
5	Reflective feedback	2,778	4.63
6	Standardized feedback	1,442	2.40
7	Specific targeted feedback	4,224	7.04
	Total	59,967	100.00

Table 18 displays the type of feedback provided for each element. Overall, the distribution of comments by individual elements was similar to the overall distribution for all elements. For most elements, the percentage of comments without feedback was approximately 20% with the lowest percentage of comments without feedback of 15.75% for the element, "Presenting unusual and intriguing information." The element with the highest percentage (43.47%) of scored elements without comments was "Asking questions of low expectancy students." It is important to note the proportion of scored elements overall for some elements were below 1% of all scored elements and, therefore, had little impact on total percentages. Among all scored elements, no element had more than 1% of unrelated feedback.

Feedback related to recounting observation events made up over half of all feedback on scored elements for all but two elements "Examining errors in reasoning" (47.14%) and "Asking questions of low expectancy students" (41.96%). The highest proportion of recounting feedback was providing for "Using friendly controversy" where

66.67% of all feedback recounted observation events. General affirmation feedback was fewer than 10% of all feedback for nearly all elements, ranging from a low of 3.78% for "Providing clear learning goals and scales" to 13.88% for "Probing incorrect answers with low expectancy students".

Reflective feedback was fewer than 10% of all feedback for all elements. The lowest percentage of reflective feedback was given for the "Providing resources and guidance" element at 1.14%. The highest percentage of reflective feedback was provided for "Noticing when students are not engaged" at 8.33 percent. Standardized feedback was also scarce with no element containing more than 6% of this feedback type. The percentage of standardized feedback ranged from a low of 0.6% for the "Providing opportunity for student talk" element to 5.74% for "Practicing skills, strategies and processes."

Table 18

Frequency and Percentages of Levels of Feedback by Element

Levels of Feedback by Element	*f*	%
Providing clear learning goals and scales		
Level 1 No feedback	1,262	21.02
Level 2 Unrelated feedback	21	.35
Level 3 Recount of observation events	3,497	58.23
Level 4 General affirmations (praise)	227	3.78
Level 5 Reflective feedback	296	4.93
Level 6 Standardized feedback	165	2.74
Level 7 Specific targeted feedback	537	8.94
Tracking student progress		
Level 1 No feedback	651	21.83
Level 2 Unrelated feedback	4	0.13
Level 3 Recount of observation events	1,631	54.69
Level 4 General affirmations (praise)	189	6.34
Level 5 Reflective feedback	183	6.14
Level 6 Standardized feedback	80	2.68
Level 7 Specific targeted feedback	244	8.18
Celebrating success		
Level 1 No feedback	243	19.97
Level 2 Unrelated feedback	2	0.16
Level 3 Recount of observation events	737	60.56
Level 4 General affirmations (praise)	94	7.72
Level 5 Reflective feedback	55	4.52
Level 6 Standardized feedback	36	2.96
Level 7 Specific targeted feedback	50	4.11
Establishing classroom routines		
Level 1 No feedback	940	21.00
Level 2 Unrelated feedback	11	0.25
Level 3 Recount of observation events	2813	62.83
Level 4 General affirmations (praise)	226	5.05
Level 5 Reflective feedback	146	3.26
Level 6 Standardized feedback	70	1.56
Level 7 Specific targeted feedback	271	6.05
Organizing the classroom physical layout		
Level 1 No feedback	352	20.73
Level 2 Unrelated feedback	7	0.41
Level 3 Recount of observation events	1,077	63.43
Level 4 General affirmations (praise)	90	5.30
Level 5 Reflective feedback	66	3.89
Level 6 Standardized feedback	26	1.53
Level 7 Specific targeted feedback	80	4.71

Levels of Feedback by Element	*f*	%
Identifying critical information		
Level 1 No feedback	604	21.28
Level 2 Unrelated feedback	9	0.32
Level 3 Recount of observation events	1,666	58.70
Level 4 General affirmations (praise)	181	6.38
Level 5 Reflective feedback	147	5.18
Level 6 Standardized feedback	49	1.73
Level 7 Specific targeted feedback	182	6.41
Organizing students to interact with new knowledge		
Level 1 No feedback	382	24.55
Level 2 Unrelated feedback	1	0.06
Level 3 Recount of observation events	883	56.75
Level 4 General affirmations (praise)	76	4.88
Level 5 Reflective feedback	64	4.11
Level 6 Standardized feedback	32	2.06
Level 7 Specific targeted feedback	118	7.58
Previewing new content		
Level 1 No feedback	232	15.73
Level 2 Unrelated feedback	2	0.14
Level 3 Recount of observation events	928	62.92
Level 4 General affirmations (praise)	102	6.92
Level 5 Reflective feedback	72	4.88
Level 6 Standardized feedback	23	1.56
Level 7 Specific targeted feedback	116	7.86
Chunking content into ‘digestible bites’		
Level 1 No feedback	517	21.41
Level 2 Unrelated feedback	3	0.12
Level 3 Recount of observation events	1,411	58.43
Level 4 General affirmations (praise)	157	6.50
Level 5 Reflective feedback	125	5.18
Level 6 Standardized feedback	59	2.44
Level 7 Specific targeted feedback	143	5.92
Processing of new information		
Level 1 No feedback	288	26.84
Level 2 Unrelated feedback	1	0.09
Level 3 Recount of observation events	556	51.82
Level 4 General affirmations (praise)	56	5.22
Level 5 Reflective feedback	52	4.85
Level 6 Standardized feedback	17	1.58
Level 7 Specific targeted feedback	103	9.60

Levels of Feedback by Element	*f*	%
Elaborating of new information		
Level 1 No feedback	240	21.43
Level 2 Unrelated feedback	3	0.27
Level 3 Recount of observation events	590	52.68
Level 4 General affirmations (praise)	88	7.86
Level 5 Reflective feedback	70	6.25
Level 6 Standardized feedback	26	2.32
Level 7 Specific targeted feedback	103	9.20
Recording and representing new knowledge		
Level 1 No feedback	361	22.04
Level 2 Unrelated feedback	6	0.37
Level 3 Recount of observation events	946	57.75
Level 4 General affirmations (praise)	85	5.19
Level 5 Reflective feedback	90	5.49
Level 6 Standardized feedback	36	2.20
Level 7 Specific targeted feedback	114	6.96
Reflecting on learning		
Level 1 No feedback	149	20.67
Level 2 Unrelated feedback	4	0.55
Level 3 Recount of observation events	394	54.65
Level 4 General affirmations (praise)	43	5.96
Level 5 Reflective feedback	45	6.24
Level 6 Standardized feedback	12	1.66
Level 7 Specific targeted feedback	74	10.26
Reviewing content		
Level 1 No feedback	1,063	25.11
Level 2 Unrelated feedback	5	0.12
Level 3 Recount of observation events	2,318	54.75
Level 4 General affirmations (praise)	20	4.75
Level 5 Reflective feedback	220	5.20
Level 6 Standardized feedback	109	2.57
Level 7 Specific targeted feedback	318	7.51
Organizing students to practice/deepen knowledge		
Level 1 No feedback	472	17.23
Level 2 Unrelated feedback	4	0.15
Level 3 Recount of observation events	1,652	60.31
Level 4 General affirmations (praise)	140	5.11
Level 5 Reflective feedback	152	5.55
Level 6 Standardized feedback	77	2.81
Level 7 Specific targeted feedback	242	8.84

Levels of Feedback by Element	*f*	%
Using homework		
Level 1 No feedback	33	20.89
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	81	51.27
Level 4 General affirmations (praise)	15	9.49
Level 5 Reflective feedback	10	6.33
Level 6 Standardized feedback	5	3.16
Level 7 Specific targeted feedback	14	8.86
Examining similarities and differences		
Level 1 No feedback	247	21.34
Level 2 Unrelated feedback	7	0.61
Level 3 Recount of observation events	637	55.06
Level 4 General affirmations (praise)	65	5.62
Level 5 Reflective feedback	67	5.79
Level 6 Standardized feedback	42	3.63
Level 7 Specific targeted feedback	92	7.95
Examining errors in reasoning		
Level 1 No feedback	161	26.35
Level 2 Unrelated feedback	3	0.49
Level 3 Recount of observation events	288	47.14
Level 4 General affirmations (praise)	46	7.53
Level 5 Reflective feedback	27	4.42
Level 6 Standardized feedback	32	5.24
Level 7 Specific targeted feedback	54	8.84
Practicing skills, strategies, and processes		
Level 1 No feedback	1,034	26.86
Level 2 Unrelated feedback	3	0.08
Level 3 Recount of observation events	2,003	52.04
Level 4 General affirmations (praise)	167	4.34
Level 5 Reflective feedback	286	4.83
Level 6 Standardized feedback	221	5.74
Level 7 Specific targeted feedback	235	6.11
Revising knowledge		
Level 1 No feedback	104	22.66
Level 2 Unrelated feedback	3	0.00
Level 3 Recount of observation events	115	61.50
Level 4 General affirmations (praise)	8	4.28
Level 5 Reflective feedback	4	2.14
Level 6 Standardized feedback	3	1.60
Level 7 Specific targeted feedback	15	8.02

Levels of Feedback by Element	*f*	%
Organizing students for cognitively complex tasks		
Level 1 No feedback	42	22.46
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	115	61.50
Level 4 General affirmations (praise)	8	4.28
Level 5 Reflective feedback	4	2.14
Level 6 Standardized feedback	3	1.60
Level 7 Specific targeted feedback	15	8.02
Engaging students in cognitively complex tasks		
Level 1 No feedback	73	21.99
Level 2 Unrelated feedback	3	0.90
Level 3 Recount of observation events	189	56.93
Level 4 General affirmations (praise)	23	6.93
Level 5 Reflective feedback	6	1.81
Level 6 Standardized feedback	4	1.20
Level 7 Specific targeted feedback	34	10.24
Providing resources and guidance		
Level 1 No feedback	74	28.03
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	152	57.58
Level 4 General affirmations (praise)	21	7.95
Level 5 Reflective feedback	3	1.14
Level 6 Standardized feedback	4	1.52
Level 7 Specific targeted feedback	10	3.79
Noticing when students are not engaged		
Level 1 No feedback	291	15.75
Level 2 Unrelated feedback	2	0.11
Level 3 Recount of observation events	1,048	56.71
Level 4 General affirmations (praise)	74	5.00
Level 5 Reflective feedback	154	8.33
Level 6 Standardized feedback	55	2.98
Level 7 Specific targeted feedback	224	12.12
Using academic games		
Level 1 No feedback	123	22.74
Level 2 Unrelated feedback	2	0.37
Level 3 Recount of observation events	315	58.23
Level 4 General affirmations (praise)	42	7.76
Level 5 Reflective feedback	18	3.33
Level 6 Standardized feedback	5	0.92
Level 7 Specific targeted feedback	36	6.65

Levels of Feedback by Element	*f*	%
Managing response rates		
Level 1 No feedback	420	22.65
Level 2 Unrelated feedback	2	0.11
Level 3 Recount of observation events	936	50.49
Level 4 General affirmations (praise)	139	7.5
Level 5 Reflective feedback	111	5.99
Level 6 Standardized feedback	33	1.78
Level 7 Specific targeted feedback	213	11.49
Using physical movement		
Level 1 No feedback	141	20.09
Level 2 Unrelated feedback	2	0.28
Level 3 Recount of observation events	459	65.38
Level 4 General affirmations (praise)	45	.41
Level 5 Reflective feedback	19	2.71
Level 6 Standardized feedback	6	0.85
Level 7 Specific targeted feedback	30	4.27
Maintaining a lively pace		
Level 1 No feedback	568	26.48
Level 2 Unrelated feedback	1	0.05
Level 3 Recount of observation events	1,137	53.01
Level 4 General affirmations (praise)	179	8.34
Level 5 Reflective feedback	75	3.50
Level 6 Standardized feedback	41	1.91
Level 7 Specific targeted feedback	144	6.71
Demonstrating intensity and enthusiasm		
Level 1 No feedback	328	23.31
Level 2 Unrelated feedback	7	0.50
Level 3 Recount of observation events	854	60.70
Level 4 General affirmations (praise)	133	9.45
Level 5 Reflective feedback	39	2.77
Level 6 Standardized feedback	16	1.14
Level 7 Specific targeted feedback	30	2.13
Using friendly controversy		
Level 1 No feedback	22	19.82
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	74	66.67
Level 4 General affirmations (praise)	5	4.50
Level 5 Reflective feedback	4	3.60
Level 6 Standardized feedback	1	0.90
Level 7 Specific targeted feedback	5	4.50

Levels of Feedback by Element	*f*	%
Providing opportunity for student talk		
Level 1 No feedback	75	22.87
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	205	62.50
Level 4 General affirmations (praise)	21	6.40
Level 5 Reflective feedback	17	5.18
Level 6 Standardized feedback	2	0.60
Level 7 Specific targeted feedback	8	2.43
Presenting unusual/intriguing information		
Level 1 No feedback	60	15.75
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	248	65.09
Level 4 General affirmations (praise)	41	10.76
Level 5 Reflective feedback	15	3.94
Level 6 Standardized feedback	8	2.10
Level 7 Specific targeted feedback	9	2.36
Demonstrating 'withitness'		
Level 1 No feedback	321	15.00
Level 2 Unrelated feedback	8	0.37
Level 3 Recount of observation events	1,399	65.37
Level 4 General affirmations (praise)	199	9.30
Level 5 Reflective feedback	68	3.18
Level 6 Standardized feedback	45	2.10
Level 7 Specific targeted feedback	100	4.67
Applying consequences for lack of adherence		
Level 1 No feedback	280	22.58
Level 2 Unrelated feedback	4	0.32
Level 3 Recount of observation events	701	56.53
Level 4 General affirmations (praise)	47	3.79
Level 5 Reflective feedback	59	4.76
Level 6 Standardized feedback	22	1.77
Level 7 Specific targeted feedback	127	10.24
Acknowledging adherence to rules/procedures		
Level 1 No feedback	160	21.45
Level 2 Unrelated feedback	2	0.27
Level 3 Recount of observation events	469	62.87
Level 4 General affirmations (praise)	49	6.57
Level 5 Reflective feedback	16	2.14
Level 6 Standardized feedback	12	1.61
Level 7 Specific targeted feedback	38	5.09

Levels of Feedback by Element	*f*	%
Understanding students' interests/backgrounds		
Level 1 No feedback	101	21.49
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	312	66.38
Level 4 General affirmations (praise)	35	7.45
Level 5 Reflective feedback	10	2.13
Level 6 Standardized feedback	7	1.49
Level 7 Specific targeted feedback	5	1.06
Using verbal and nonverbal behaviors		
Level 1 No feedback	337	22.05
Level 2 Unrelated feedback	3	0.20
Level 3 Recount of observation events	000	75/39
Level 4 General affirmations (praise)	122	8/09
Level 5 Reflective feedback	25	1.64
Level 6 Standardized feedback	26	1.70
Level 7 Specific targeted feedback	16	1.05
Displaying objectivity and control		
Level 1 No feedback	69	21.43
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	195	60.56
Level 4 General affirmations (praise)	23	7.14
Level 5 Reflective feedback	11	3.42
Level 6 Standardized feedback	5	1.55
Level 7 Specific targeted feedback	19	5.90
Demonstrating value and respect for low expectancy students		
Level 1 No feedback	81	20.66
Level 2 Unrelated feedback	2	0.51
Level 3 Recount of observation events	233	59.44
Level 4 General affirmations (praise)	38	9.69
Level 5 Reflective feedback	15	3.83
Level 6 Standardized feedback	9	2.30
Level 7 Specific targeted feedback	14	3.57
Asking questions of low expectancy students		
Level 1 No feedback	173	43.47
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	167	41.96
Level 4 General affirmations (praise)	25	6.28
Level 5 Reflective feedback	14	3.52
Level 6 Standardized feedback	3	0.75
Level 7 Specific targeted feedback	16	4.02

Levels of Feedback by Element	*f*	%
Probing incorrect answers with low expectancy students		
Level 1 No feedback	48	22.97
Level 2 Unrelated feedback	0	0.00
Level 3 Recount of observation events	107	51.20
Level 4 General affirmations (praise)	29	13.88
Level 5 Reflective feedback	7	3.35
Level 6 Standardized feedback	4	1.91
Level 7 Specific targeted feedback	14	6.70

Because of its importance, the frequencies and percentages of specific, targeted feedback for all elements are shown in Table 19 along with the frequency of all scored elements, and the total number of scored elements with comments. For no element did the frequency of specific, targeted feedback exceed the frequency of no comments. The number of elements with no comments was most often two to four times greater than the number of elements with specific, targeted feedback. Across all elements, the percentage of specific, targeted feedback ranged from a low of 1.05% for "Using verbal and nonverbal behaviors" to 12.12% for "Noticing when students are not engaged."

Table 19

Elements Scored With Comments and Specific, Targeted Feedback

Element #	Scored Elements	Elements with Comments	Elements With Specific, Targeted Feedback	% Elements With Specific, Targeted Feedback
1	6,005	4,744	537	8.94
2	2,982	2,332	244	8.18
3	1,217	974	50	4.11
4	4,477	3,537	271	6.05
5	1,698	1,346	80	4.71
6	2,838	2,235	182	6.41
7	1,556	1,174	118	7.58
8	1,475	1,243	116	7.86
9	2,415	1,899	143	5.92
10	1,073	785	103	9.60
11	1,120	880	103	9.20
12	1,638	1,277	114	6.96
13	721	572	74	10.26
14	4,234	3,171	318	7.51
15	2,739	2,267	2242	8.84
16	158	125	14	8.86
17	1,157	910	92	7.95
18	61	450	54	8.84
19	3,849	2,815	235	6.11
20	459	355	28	6.10
21	187	145	15	8.02
22	332	259	34	10.24
23	264	190	10	3.79
24	1,848	1,557	224	12.12
25	541	418	36	6.65
26	1,854	1,434	213	11.49
27	702	561	30	4.27
28	2,145	1,577	144	6.71
29	1,407	1,079	30	2.13
30	111	89	5	4.50
31	328	253	8	2.44
32	381	321	8	2.10
33	2,140	1,819	100	4.67
34	1,240	960	127	10.24
35	746	586	38	5.09
36	470	369	5	1.06
37	1,528	1,191	16	1.05
38	322	253	19	5.90
39	393	311	14	3.57
40	298	225	16	4.02
41	209	161	14	6.70

Research Question 4

What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

In order to determine if there was a difference between the type of feedback and student achivement oucomes, each teacher in the sample was placed into one of seven categories based on the predominant feedback type provided by the evaluator. These data are displayed in Table 20. A high majority of teachers (2,165, 79.65%) received recount of observation feedback. The next largest group of teachers (282, 10.38%) received primarily no feedback. Of the remaining feedback categories, 104 (3.83%) received predominantly general affirmation feedback; 91 (3.35%) received predominantly specific targeted feedback; 53 (1.95%) received predominantly reflective feedback; and 23 (0.85%) received predominantly standardized feedback. No teachers received predominantly unrelated feedback.

Table 20

Teachers by Predominant Feedback Type

Predominant Feedback Type	f	% of Sample
Predominantly no feedback	282	10.38
Predominantly unrelated feedback	0	0.00
Predominantly recount of observation events	2,165	79.65
Predominantly general affirmations (praise)	104	3.83
Predominantly reflective feedback	53	1.95
Predominantly standardized feedback	23	0.85
Predominantly specific targeted feedback	91	3.35
Total Teachers	2,718	100.00

Table 21 displays value-added scores for teachers based on predominant feedback type. There was little variation among predominant feedback types, and the value-added scores ranged from 4.30% above average to 7.01% above average. Teachers who received predominantly standardized feedback had the highest average value-added scores (7.01% above average). Teachers who received predominantly no feedback received the lowest average value-added scores (4.30% above average).

Table 21

Value-added (VAM) Scores by Predominant Feedback Type (N = 2,718)

Predominant Feedback Type	*f*	Average VAM Score*
Predominantly no feedback	282	4.30% above
Predominantly unrelated feedback	0	Not applicable
Predominantly recount of observation events	2,165	5.15% above
Predominantly general affirmations (praise)	104	5.18% above
Predominantly reflective feedback	53	6.88% above
Predominantly standardized feedback	23	7.01% above
Predominantly specific targeted feedback	91	5.15% above

*Expressed in percentage of a year's growth above or below average associated students' score on statewide assessments

Because very few teachers were located in categories that received a majority of comments other than recounts of the observation, a one-way ANOVA was performed to determine if the differences in value-added scores by feedback type were statistically significant. Table 22 displays these results. The summary table reveals that there was no significant relationship between the predominant type of feedback provided to teachers and their value-added scores.

Table 22

One-way ANOVA Results: Value-added Scores for Predominant Types of Feedback Groups

Descriptor	*SS*	*df*	*MS*	*f*	*p-value*
Between feedback groups	236.0114	5	47.20230	0.71	0.4243
Within feedback groups	2752.126	2,712	1.0145		
Total	2988.137	2,717			

Summary

In this chapter, data were analyzed to respond to four research questions associated with the analysis of classroom observations and comments as they related to value-added scores. Descriptive and inferential statistics were used in the analysis.

Research Question 1 addressed the frequency of classroom observations and comments for teachers. The population of all teachers and the sample teachers produced similar distributions for all data in this study. The most common number of observations for teachers was three observations. A total of 87.52% of all teachers received six or fewer observations, and 91.22% of sample teachers received six of fewer observations. Most elements that were scored by observers received comments. Nearly 80% of scored elements were commented upon. High school teachers were less likely to receive comments on scored elements than teachers at other levels.

Over one-third of scored elements came from the following five elements: "Providing clear learning goals and scales," "Tracking student progress," "Establishing

classroom routines," "Reviewing content," and "Practicing skills, strategies and processes." The percentage of total scored elements by element varied little by level.

Research Question 2 focused on the relationship between the number of observations or percentage of elements scored with comments and student achievement outcomes as measured by teacher value-added scores. For the relationship between number of observations and value-added scores, the null hypothesis was not rejected. This indicated that there was no significant relationship between the number of classroom observations and teacher value-added scores. Similarly, the null hypothesis was not rejected for the relationship between the percentage of elements scored with comments and teacher value-added scores. No significant relationship existed between the percentage of elements with comments and teacher value-added scores.

Research Question 3 addressed the frequency by level of feedback provided by observers to teachers during classroom observations. Nearly 60,000 scored elements were coded to determine what type of feedback was provided to teachers. This feedback was separated in categories determining whether the content was (a) not present, (b) unrelated, (c) a recount of events, (d) general affirmations, (e) reflective, (f) standardized, or (g) specific and targeted. Intercoder reliability checks indicated that the rubric for scoring these comments measured replicable and clear categories. As a result of coding, it was found that the most common form of feedback was a recount of observation events. The next most common form of feedback was not providing feedback. More complex feedback that could potentially provide additional support for instruction was rare across all elements. There were few differences in the distribution of feedback types

by elements. Where differences existed, it was common for elements to have comments provided sparingly.

Research Question 4 was used to determine if there was a difference between the type of feedback provided to teachers and student achievement outcomes as measured by value-added scores. The descriptive and inferential analyses did not yield any significant association between the predominant types of feedback provided and value-added scores. Most teachers received a majority of recount feedback in their comments. Further analysis, using an analysis of variance, indicated that there was no significant relationship between classroom observations and value-added scores based on feedback categories, and the null hypothesis was not rejected.

Table 23 presents an overall summary of the study, including the research questions, variables, sources of data, methods used to conduct the research, and the results. Chapter 5 contains an elaborated summary and discussion of the findings of the study along with implications for practice and future research.

Table 23

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

Research Questions	Qualifying/Independent Variable(s)	Dependent Variable	Source(s) of Data	Method(s) of Analysis	Results
1. What is the frequency of classroom observations and comments or elements scored for teachers, including formal, informal and walkthrough observations?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations.	Frequency of Observations	iObservation Reporting	Descriptive statistics were used to describe resulting data, which included frequency distribution for each type of observation.	• 35% of scores are in 5 elements. • Goals and Scales received 10% of scores. • Teachers received up to 32 observations. • Teacher Observations 1-0.15% 6-8.39% 2-0.78% 7-3.90% 3-41.94% 8-1.91% 4-26.64% 9-0.85% 5-13.32% 10-0.48% 11 or More-1.66%
2. What relationship if any exists between the frequency of observations by observers as measured by the number of classroom observations during a school year and student achievement outcomes as measured by teacher VAM scores?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Teachers with matched individual VAM scores for the 2013-2014 school year.	Value-added score results	iObservation Reporting Teacher VAM scores	Pearson's r computed to determine relationship between total number of classroom observations and VAM scores. Statistical inference conducted to determine if coefficient is significantly different from zero at 0.05 level.	• No significant relationship existed between VAM scores and number of observations or percentage of comments for teachers at any grade level band.

Research Questions	Qualifying/Independent Variable(s)	Dependent Variable	Source(s) of Data	Method(s) of Analysis	Results
3. What is the frequency by level of feedback, defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Implementation of the scales for protocol observation. • Data captured from the feedback section of the observation protocol.	Types of Feedback	iObservation Feedback section	Descriptive statistics, including the frequency distribution were run to illustrate findings.	Frequency of FB types L1-No FB 21.88% L2-Unrelated 0.23% L3-Recount 57.84% L4-Praise 5.97% L5-Reflective? 4.63% L6-Copy,Cut,Paste 2.40% L7-Targeted FB 7.04% • Specific Targeted Feedback ranges from 1% to 12% across the elements. • Independent element distribution is similar to overall distribution and held true within the elements.
4. What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?	• Teachers with more than three years of teaching experience and not holding National Board Certification. • Observation data for the 2013-2014 school year. • Number of walkthrough, informal and formal observations. • Implementation of the scales for protocol observation. • Data captured from the feedback section of the observation protocol. • Teachers with matched individual VAM scores for the 2013-2014 school year.	Value-added score results	iObservation Feedback section Teacher VAM scores	After teachers were uniquely classified into categories, an average of all teacher VAM scores for each category was calculated and compared through one-way analysis of variance (ANOVA).	• 80% of teachers' predominantly had recount of events for comments. • 10% of teachers' predominantly had no comments. • No significant relationship existed between predominant comment for teachers and VAM scores.

CHAPTER 5
SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Introduction

This chapter contains a summary of the research study findings and a discussion of how the findings may influence policy. The summary includes a restatement of the problem, the purpose of this study, a review of the research questions, the conceptual framework, and the research design. The remainder of the chapter is devoted to a discussion of the findings along with policy implications and recommendations for future research. The chapter is concluded with a summative statement about the research study.

Summary of the Study

Over the past decade, there has been a widespread shift in the process and purpose of teacher evaluation. The goal has been to transform the effort from an annual exercise for compliance to a process conducted for the ultimate purpose of improving student achievement. Nationwide, considerable time, energy, and money have been spent revamping teacher evaluation systems in hopes that changes in instruction will reverse years of over-inflated teacher evaluation scores and stagnant student achievement data. This study was conducted to evaluate the relationships between the number of observations and types of feedback provided in the evaluations of instruction personnel in a large, urban public school district and student learning growth. Also examined were the extent to which frequency and quality of observations and feedback were associated with student learning growth. It has been said that "good feedback can significantly improve

learning processes and outcomes, if delivered correctly" (Shute, 2008, p. 154). The pivotal portion of affecting change is the assurance that the feedback is aligned, targeted, and delivered appropriately to teachers. This is where the transition from prior evaluation systems needs to occur.

During the 2013-2014 school year, the Marzano instructional framework was used as the evaluation system to evaluate teacher performance in the targeted school district. The sample of 2,718 was drawn from a population of 14,000 teachers in the school district. The observation frequency and additional comments they received in each of the elements rated were investigated in relation to their value-added (VAM) scores, to determine if there was a relationship between frequency of observations or type of feedback received and student achievement. To date there has been little research into the influence on student achievement of the frequency of classroom observations or the relationship to forms of feedback provided following classroom observations. The conceptual framework upon which the study was built was reinforcement theory. The findings of this study were intended to inform policy on current teacher observation practices and the impact of feedback on instruction.

Following are the research questions used to guide the study:

1. What is the frequency of classroom observations and comments for teachers, including formal, informal, and walkthrough observations?
2. What relationship, if any, exists between the frequency of observations by observers as measured by the number of classroom observations and elements

scored during a school year and student achievement outcomes as measured by teacher VAM scores?

H_{01}. There is no significant relationship between the number of classroom observations during a school year and student achievement outcomes as measured by VAM.

3. What is the frequency by level of feedback defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?
4. What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

   H_{02}. There is no significant difference between classroom observations by feedback category and student achievement outcomes as measured by VAM scores.

Data for the study were collected to respond to the research questions. For Research Questions 1 and 2, descriptive data were organized and analyzed to provide a snapshot of how many observations teachers received during the 2013-2014 school year and if VAM scores were impacted. For Research Questions 3 and 4, a rubric was designed and used to clearly define, sort and evaluate the levels of feedback provided to teachers in the sample. Nearly 60,000 scored elements were reviewed and leveled, using a rubric which indicated 46,850 scored elements had some form of feedback. In some

cases, when there were multiple levels represented in one comment, the higher level was chosen. For example, if a praise statement and a reflective question were both present, the latter was chosen. The rubric with the working definitions and sample language for each of the levels is presented in Appendix E.

Discussion of the Findings

Research Question 1

> What is the frequency of classroom observations and comments for teachers, including formal, informal and walkthrough observations?

Descriptive statistics provided a wealth of information about the sample, the number of observations sample teachers received and the elements they were observed using. It is noteworthy that a higher percentage of middle school teachers and a lower percentage of high school teachers existed in the sample group than in the population, perhaps indicating the group of middle school teachers was more stable, but that the population of high school teachers had less experience.

Not surprisingly, the majority of teachers in the sample received the minimum number of three observations required by the school district. This may be due in part to the time consuming nature of the requirements for completing the observation process. Formal observations require a pre-conference, to be completed over an entire class period, and to be followed up by a post-conference. The time consuming nature of the process may also explain the small number of teachers who had between 12-33 observations. The numbers of teachers in each group decreased as the number of

observations increased. This could be a result of the time intensive nature of performing the increased numbers of observations or that observers were providing this type of oversight to teachers who may have needed additional support.

Also of interest, when looking at both all high school and sample high school teachers received 10% fewer comments for rated elements. There were no data available to explain this, though it is possible that high school observers give feedback verbally during conferences more so than in electronic format.

It is important to note that during the first two years of the implementation of the new teacher evaluation system, the focus for training and observation was placed heavily on Design Question 1, Communicating learning goals and feedback, Element 1, "Providing clear learning goals and scales." The 2013-2014 school year was the third year of implementation of the new evaluation system and the first year the focus was on the content elements in Design Question 2, Helping students interact with new knowledge, and Design Question 3, Helping students practice and deepen new knowledge. However, observation data continued to show a heavy focus on the goals and scales element at every level.

Quantitative data from this research study revealed that administrators and observers at all levels observed teachers using (or not using) similar strategies and engaging students in a limited number of activities with high levels of complexity and cognitive demand. Tables 4-8 provided an overview of the comments provided in each element in Domain 1 of the Marzano instructional model. The model is divided into nine Design Questions (DQ). Each of them is organized into three columns known as

Classroom Routine Events (DQ1 and DQ6), outlining the actions teachers take to organize their classrooms, Content Strategies (DQ2, DQ3 and DQ4) which indicate whether instruction is based on teaching new information, review of information or involving hypothesis, and engagement or relationship strategies enacted On the Spot (DQ5, DQ7, DQ8 and DQ9).

The results indicated that observers, overall, rated and provided comments on similar elements and were consistent at all levels. Element 1, "Providing clear learning goals and scales," had the highest percentage of overall rating and comments. This may be due, in part, to the school district focus and training on this particular element for the first three years of the implementation of the Marzano instructional model. There was also a high percentage of ratings with comments on the low risk element, DQ6, Establishing Routines and Procedures. Although they are both important elements and can easily be scored during each classroom visit, neither speaks to improving instruction in content.

Within the content strategies, the following elements had the highest percentages of comments associated with them: Element 6, "Identifying critical information," Element 9, "Chunking content into 'digestible bites'," Element 14, "Reviewing content," Element 15, "Organizing students to practice and deepen knowledge," and Element 19, "Practicing skills, strategies, and processes." This was surprising due to the training in recognizing and rating Marzano's " elements. Only Element 6, "Identifying critical information", is one of the Super 7. This indicates that observers at all levels were either not seeing these strategies being used or they were not skilled at scoring them. In

addition, only Element 1, “Providing clear learning goals and scales” and Element 19, “Practicing skills, strategies and process” are part of the 11 highest yield strategies outlined by Marzano.

The following elements from Design Question 4, Helping Students Generate and Test Hypotheses had extremely low percentages of elements scored with comments: Element 20, “Organizing students for cognitively complex tasks,” Element 21, “Engaging students in cognitively complex tasks,” and Element 22, “Providing resources and guidance.” This was consistent with national findings (Marzano & Toth, 2014) and was true even at the high school level where it would be the expectation to see activities with higher cognitive demands. However, this was not the case. A review of the tabular data told a story of a lack of high cognitive demand strategies being used by teachers and a lack of elements being scored with comments by observers to provide them with feedback to increase the effective use of the strategies.

These data are a confirmation of the research shared at the June 2014 Marzano International Conference (Marzano & Toth, 2014). At that conference, presenters elaborated on the findings resulting from the analysis of more than two million data points nationwide. The researchers found that only 6% of observed lessons were devoted to high cognitive complexity tasks involving hypothesis generation and testing found in Design Question 4, Helping Students Generate and Test Hypotheses. In the present study, the figure was lower yet--only 3%. Summary data are presented in Table 23 for the present study contrasted with that of Marzano and Toth, revealing the congruency of results of both studies. The results serve to confirm previous studies and add new

information to show the need for more focused observations at all levels from elementary to high school. Based on the findings from this study, the recommendation is for observers to focus on providing feedback to teachers at all levels in the elements known to promote high cognitive demands. This evidence-based guidance is intended to inform practice as well as define teaching practices that are predictive of student learning.

Table 24

Comparison of Teachers' Most and Least Used Instructional Strategies

Strategies	Marzano & Toth	Rafalski
Most Used Strategies (Lecture, Practice, and Review)		
Identifying critical information	12.5%	11.8%
Practicing skills, strategies and processes	12.0%	13.4%
Chunking information into digestible bites	10.7%	9.4%
Reviewing content	11.8%	15.2%
Least Used Strategies (Critical to develop cognitive complexity)		
Engaging students in cognitively complex tasks involving hypothesis generating and testing	1.2%	1.1%
Revising knowledge	1.1%	1.5%
Organizing students for cognitively complex tasks	.9%	.7%

Research Question 2

> What relationship if any exists between the frequency of observations by observers as measured by the number of classroom observations and elements scored during a school year and student achievement outcomes as measured by teacher VAM scores?

The findings in the present research study showed that there was no significant relationship between the number of observations and teachers' VAM scores. This was shown in tabular data and through correlational analyis. This was not surprising, based on the lack of high cognitive complexity elements scored and the lack of substantive feedback provided to teachers at all levels. Therefore the null hypothesis, H01, was not rejected. There was no significant relationship between the number of classroom observations during a school year and student achievement outcomes as measured by VAM.

Data have been presented as to VAM scores according to the number of observations conducted for teachers. Teachers with between three and six observations had higher than average VAM scores, and teachers with fewer than three, or either seven or eight observations had lower than average VAM scores. Although the number was small, teachers with nine or 10 observations had above average growth scores. Few teachers in the sample school district received more than 10 observations.

Although just fewer than 80% of all the elements scored had comments included with the ratings, there was no significance to the comments associated with the scored elements as they related to teachers' VAM scores. No significant relationship was found between the value-added scores overall or at any level and the percentage of elements scored that contained comments.

Based on the findings from this study, one recommendation is for observers to provide more than the minimum number of observations. The mere fact that comments exist does not have a bearing on student achievement, and therefore it is recommended that the comments include feedback aligned according to the research model.

Research Question 3

> What is the frequency by level of feedback, defined as no feedback, unrelated feedback, recount of observation events, general affirmations, reflective feedback, standardized feedback, or specific targeted feedback, provided by observers to teachers during classroom observations?

The findings from this research study showed that the majority of the feedback provided was a retelling of the classroom events during a given observation rather than feedback that was informative, constructive, objective, actionable, and focused on specific classroom strategies and behaviors during a set time interval, (Florida RTTT glossary). Even the no feedback level, which meant the comment section was left blank, had a higher percentage than did standardized feedback and reflective questions combined. It should be noted, however, that observers may have given "off the record" feedback in the form of written or verbal communication which is not part of a data source for the study.

Three main themes emerged in the analysis of the qualitative data. First, as previously discussed, the elements observers chose to score were those, according to evidence-based guidance, that will not result in the highest levels of cognitively complex thinking by students. Second, the comments provided by observers were predominantly just a recap of the events of the observation and not considered feedback. Lastly, when

feedback was provided, it was most often provided in the area of "Providing clear learning goals and scales" which will not likely result in improvements in student learning that other strategies provide. These themes resonated throughout the study results.

Comments were coded into levels for all elements rated for the sample group of teachers in the study. Overall, 22% of the ratings in the sample group had no comment recorded. A dismal example of the guidance teachers are receiving to improve their practice, only 7% of the ratings teachers in the sample group received were considered to be feedback. Of the comments in the elements scored, 5% and 2% respectively were considered reflective questions either created by the observer or standardized by using the copy, cut and paste feature within the iObservation system. Reflective questions encourage teachers to think about their practice or ways to improve their craft. However, without knowledge of the outcomes of the follow-up discussion, the effectiveness of feedback intervention is unknown. A total of 6% of the comments contained a praise statement intended to be an affirmation of what the teacher was doing well or the overall opinion of the class or strategy use. A full 58% of the comments fell into the category of simply a recount of classroom events. This may be attributed to a lack of training on what feedback is, how to provide it or as a result of observer training. Some observers may have been trained to provide comments as a recap of what happened during the observation or a previous culture within the system which only provided a menu of comments. At the onset of the implementation of the Marzano instructional model, observers were instructed to use the comment box to write notes of the events during the

classroom visit or use later for scoring the element or to justify the rating when shared with the teacher. This, coupled with a previous evaluation system that discouraged recommendations or other feedback in writing, may have led to the continuance of past practice.

A closer look at each of the elements reflects the level of comments provided and also sheds some light on the areas about which observers were most comfortable in providing feedback. Feedback as defined in this study, when given, was found most often for Elements 1, 10, 11, 13, 18, 24, 26, and 34 as detailed in Table 25

Table 25

Observers' Most Commented Upon Elements

#	Element	%
1	Providing clear learning goals and scales	8.94
10	Processing of new information	9.60
11	Elaborating new informtion	9.20
13	Reflecting on learning	10.26
18	Examining errors in reasoning	8.84
24	Noticing when students are not engaged	12.12
26	Managing response rates	11.49
34	Applying consequences for lack of adherance	10.24

Unfortunately, as described in the third resonating theme, only Elements 11 and 18 are considered part of the Super 7 high yield instructional strategies. Most of the feedback that was being provided to teachers was in areas that do not maximize the cognitive resources and efforts of the students. Reflective questions were used most frequently in Element 19, "Practicing skills, strategies and processes" and Element 24, "Noticing when students are not engaged."

Affirmation statements (Praise) were used most often in the following elements: Element 11, "Elaborating of new information," Element 28, "Maintaining a lively pace," Element 29, "Demonstrating Intensity and enthusiasm," Element 33, "Demonstrating 'withitness'," and Element 37, "Using verbal and nonverbal behavior." Surprisingly, middle school emerged as having the highest percentage of feedback and high school had the highest percentage of praise. The importance of affirmation was noted by Fernandez-Toro et al. (1989) in the following statement:

> Students appreciate motivating comments. The most effective comments for helping students to understand inadequacies in their work are those that offer an explanation that is designed to help them bridge the gap between their current knowledge, understanding and skills and those expected of them. (p. 818).

Based on the findings of this study, a recommendation would be to reaffirm the commitment to not only observe the Super 7 or Essential 13 but also provide actionable feedback in these elements, predictive of student learning.

Research Question 4

What difference, if any, exists between the type of feedback alignment provided to teachers and student achievement outcomes as measured by VAM scores?

The null hypothesis indicating that there was no significant difference between classroom observations by feedback category and student achievement outcomes as measured by VAM scores could not be rejected. This finding was expected due to the low percentage of teachers receiving predominantly higher-level feedback. There was considerably more variation among the large numbers of teachers who received no or recount feedback than between them and the small number of teachers who received more complex feedback. This suggests that the relationship between feedback and student achievement outcomes is likely to remain difficult to measure in an environment where limited targeted and actionable feedback is provided.

Implications for Policy and Practice

It is well known that the movement to new standards necessitates teachers becoming more student-centered and proficient in actively engaging students in learning through cooperative, hands-on, and cognitively complex experiences. Educational leaders everywhere have been trying to adjust to meet the rigorous demands of these new teaching standards. The standards provide a new framework for what students are expected to know and be able to do. It moves students away from memorizing information to application of knowledge. There will be growing pains and some unknowns as school districts determine the best ways to prepare teachers to incorporate the new standards into their teaching with the goal of improving student readiness for the demands of college and career.

Three main policy and practice implications emerged from the results of this study. School and district leaders, as well as educational researchers designing teacher evaluation systems, should consider these implications:

1. It has been advised that observers should score elements based on evidence-based guidance that will result in the highest levels of cognitively complex thinking by students. The elements being scored and the levels of feedback observers provide should be monitored and analyzed for ongoing formative planning. In this study, observers did provide teachers with comments for approximately 80% of the elements they scored; however, the ratings and comments were for a limited number of low cognitive demand elements and the comments were predominantly only a recount of the events of observations. Because these types of comments are not informative, constructive, objective, actionable, and focused on specific classroom strategies or behaviors that result in student learning, the feedback is not appropriately aligned within the evaluation model to impact student learning.
2. "Formative feedback needs to take into consideration instructional context as well as characteristics of the learner to provide effective feedback for complex learning tasks" (Shute, 2008, p. 172), and "Feedback enables individuals to understand and improve their judgments, improve their expertise in the judgment task, and reduce commitment to incorrect judgment strategies" (Balzer et al., 1989, p. 412). Thus, feedback should not be used as a recap of the events of an observation. It is recommended that observers receive more

professional development in the value of providing feedback, what specific targeted feedback is, and how to provide feedback aligned with goals of the evaluation model.

3. To avoid misinterpretation, the box titled "Comments" in the online observation system of the school district should either (a) be renamed or have language added to indicate an expectation that observers provide specific targeted feedback for rated elements or (b) another box titled "Feedback" should be added for observers to leave informative, constructive, objective, actionable, and focused feedback. This would provide a distinction in the evaluation between a description of the evaluation details (e.g. time, room, length, evidence observed) and feedback provided to improve instruction. Subsequent professional development should also be provided in the use of the online space(s) provided for feedback.

Recommendations for Future Research

The goal of this research study was to determine if the number of observations and feedback provided in the evaluations of instructional personnel in a large, urban public school district was associated with student learning growth. Also examined were the frequency of observations and quality of feedback associated with student learning growth. Following are recommendations for future research based on the findings of the current study.

Due to the limitations of this study, which included the policy context for the use of the Marzano instructional model as an evaluation system, results were derived from the comments section within the online instrument. Although a limitation, overall the mandates as they pertain to fidelity of program implementation were followed, and all observers attending professional development on rating elements and completing observations within the system. Observers also were required to attend inter-rater reliability training. However, this did not discount that variability was expected in the interpretation of the elements and the way feedback was provided. Additional training needs to occur to go beyond just training in the technical aspects of the evaluation system and include professional development specifically created on how to give effective feedback for instructional improvement. A follow-up study after this professional development occurs is recommended to interpret the impact of feedback on teacher practice and ultimately student academic achievement.

Recommendations for other next steps include reviewing the specific targeted feedback and categorizing it based on common language. The verbiage in each category can then be analyzed to see if there is a relationship to improved instructional practice and to inform professional development. Formal and informal observation data can also be disaggregated to determine if there is a difference in rating results or type of feedback provided between the two. This may reveal a contrast based on the highly structured and bureaucratic conditions placed on observations through bargaining requirements versus the impromptu and more realistic picture an informal observation provides.

Although only electronic comments were reviewed from the 2013-2014 school year as part of this study, all feedback may not have been recorded within the iObservation tool as prescribed by the program. Many administrators may have given "off the record" feedback in the form of written or verbal communication which was not part of a data source for the study. More research is needed to discover if any of this feedback should be included in observations. A survey of both observers and teachers may be one way to gather information about the amount and kind of non-recorded feedback that may exist. Related to this idea, future research could also include a review of how reflective questions that have some degree of follow-up dialog can be recorded.

Further investigation could also address how the ideas of practice, questioning, and reflection interface with feedback research. There is some data to suggest that in order for someone to become an expert, it takes 10 years and 10,000 hours of focused practice (Gladwell, 2008), but the relationship with feedback is unclear. In addition, the concept that talent is not born but it can be grown through deep practice, ignition, and master coaching is interesting to think about in relation to feedback (Coyle, 2009).

As a final recommendation, research can be conducted to improve how feedback is received by the learner, which is the opposite of training observers on how to give feedback. In *Thanks for the Feedback*, the authors discussed three reasons for giving feedback: (a) appreciation, (b) coaching, and (c) evaluation, but they are all dependent on relationships and the receptivity of the receiver (Stone & Heen, 2014). Booth-Butterfield, as early as 1989, succinctly identified the problem as follows: "The influence of feedback may depend on receivers' interpretation of that information--this is

why the way it is provided is so important" (p. 119). Furthermore, the context of feedback to the learner was outlined by Shute (2008), stating "formative feedback needs to take into consideration instructional context as well as characteristics of the learner to provide effective feedback for complex learning tasks" (p. 172).

Summary

Because improving the quality of instruction is paramount to student achievement, improvement is needed in training observers to provide aligned feedback to teachers. Not only does this feedback need to be given in regard to effective instructional strategies, it also needs to be specific and targeted. By completing this study, the researcher has shared research on teacher evaluation, effectiveness, value-added models, and feedback; and some new insights have been shared as a contribution to the field of study.

When learning any discipline or craft, the ability to discover and improve is critical. As Shute (2008) described, "Imagine trying to learn something new in the absence of any feedback." In any sport whether baseball, tennis or golf, the arts from music or dance, or in any professional field from construction to the sciences, coaches, instructors and mentors do not simply recount observations from their subjects with the understanding that this recounting will allow them to improve. It would be unusual for a baseball coach to say, "You stood on first base," or for a basketball coach to say, "I love how you ran down the court." Similarly, one would not be impressed with a piano teacher who emphasized how a student "placed your fingers on the white and black keys

really well while you read those notes" or a construction foreman training new employees who commented to a trainee, "You hit the wood when you swung that hammer."

Individuals who provide training to new practitioners must understand that their roles are to give feedback aligned with where their subjects currently are, offering advice or direction on how to advance subjects to where they need to be. They close the performance gap by presenting a feedback intervention. That is why people take lessons, join a team or listen to an expert-- to improve their craft. The acknowledgement of events or a simple recounting is often needed in order to provide perspective. It is, however, only the first step to progress or improvement. If specific, targeted feedback is not provided to improve, an aspiring athlete, musician, builder, or teacher is left with nothing more than a reflection of their current strengths and weaknesses

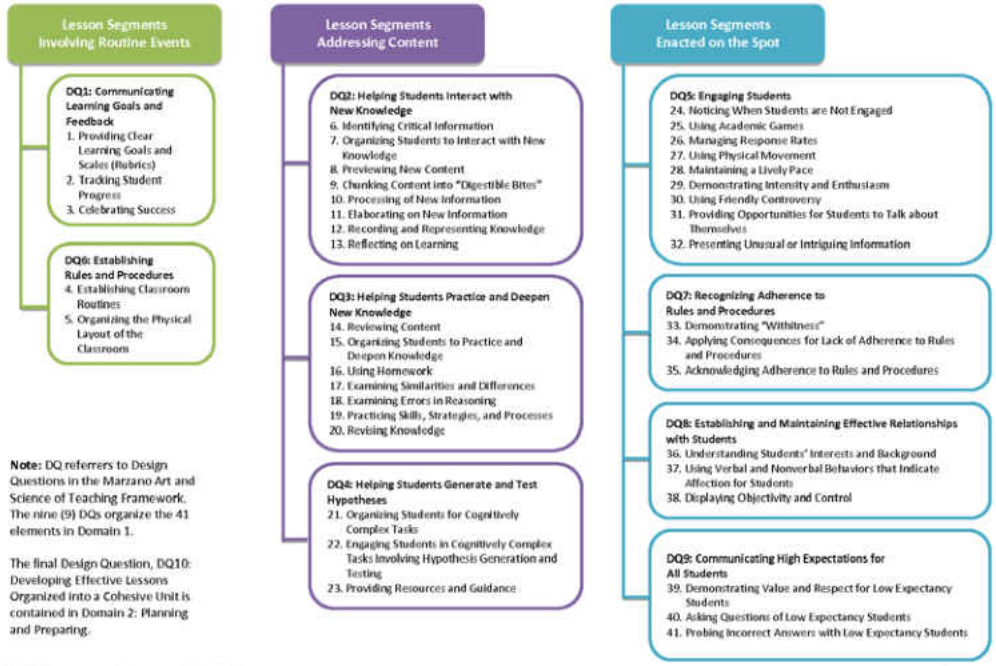
Observations and feedback should be a tool for improvement of instruction, but the data confirm that this process continues to be compliance-based with inflated scores that do not match the level of performance of students. Changing this is strongly linked to the provision of feedback associated with improving instruction and holding teachers accountable in meeting the standards outlined in the feedback. Observers are in need of professional development on how to provide effective feedback in the areas of instruction that will make the biggest impact on student achievement. Without this, the work and training on improving evaluation systems will not impact student achievement outcomes and may even undermine the role of observers in providing support to teachers.

APPENDIX A
MARZANO INSTRUCTIONAL MODEL LEARNING MAP

Marzano Art and Science of Teaching Teacher Evaluation Model
Learning Map

LearningSciencesInternational
LEARNING AND PERFORMANCE MANAGEMENT

Domain 1: Classroom Strategies and Behaviors



Note: DQ referrers to Design Questions in the Marzano Art and Science of Teaching Framework. The nine (9) DQs organize the 41 elements in Domain 1.

The final Design Question, DQ10: Developing Effective Lessons Organized into a Cohesive Unit is contained in Domain 2: Planning and Preparing.

FORMATIVE FEEDBACK GUIDELINES TO ENHANCE LEARNING (THINGS TO DO)

Prescription	Description and Reference
Focus feedback on the task, not the learner.	Feedback to the learner should address specific features of his or her work in relation to the task, with suggestions on how to improve (e.g., Butler, 1987; Corbett & Anderson, 2001; Kluger & DeNisi, 1996; Narciss & Huth, 2004).
Provide elaborated feedback to enhance learning.	Feedback should describe the what, how, and why of a given problem. This type of cognitive feedback is typically more effective than verification of results (e.g., Bangert-Drowns et al, 1991; Gilman, 1969; Mason & Bruning, 2001; Narciss & Huth, 2004).
Present elaborated feedback in manageable units.	Provide elaborated feedback in small enough prieces fo that it is not overwhelming and discarded (Bransfoed et a., 2000; Sweller et al., 1998). Presenting too much information may not only result in superficial learning but may also invoke cognitive overload (e.g. Mayer & Moreno, 2002; Phye & Bender, 1989). A stepwise presentation of feedback offers the possibility to control for mistakes and gives learners sufficient information to correct errors on their own.
Be specific and clear with feedback message.	If feedback ins not specific or clear, it can impede learning and can frustrate learners (e.g. Moreno, 2004; Williams, 1997). If possible, try to link feedback clearly and specifically to goals and performance (Hoska, 1993; Song & Keller, 2001).
Keep feedback as simple as possible (based on learners needs and instructional constraints).	Simple feedback is generally based on one cue (e.g., verification or hint and complex feedback on multiple cues (e.g. verification, correct response, error analysis). Keep feedback as simple and focused as possible. Generate only enough information to help students and not more. Kulhavy et al. (1985) found that feedback that was too complex did not promote learning compared to simpler feedback.
Reduce uncertainty between performance and goals.	Formative feedback should clarify goals and seek to reduce or remove uncertainty in relation to how well learners are performing on task, and what needs to be accomplished to attain the goal(s) (e.g., Ashford et al., 2003; Bangert-Drowns et al., 1991).
Give unbiased, objective feedback, written or via computer.	Feedback from a trustworthy source will be considered more seriously than other feedback, which may be disregarded. This may explain why computer-based feedback is often better than human-delivered in some experiments in that perceived biases are eliminated (see Kluger & DeNisi, 1996).
Promote a "learning" goal orientation via feedback.	Formative feedback can be used to alter goal orientation-from a focus on performance to a focus on learning (Hoska, 1993). This can be facilitated by crafting feedback emphasizing that effort yields increased learning and performance, and mistakes are an important part of the learning process (Dweck, 1986).
Provide feedback after learners have attempted a solution.	Do not let learners see answers before trying to solve a problem on their own . Several studies that have controlled presearch availability show a benefit of feedback, whereas studies without such control show inconsistent results (Bangert-Drowns et al., 1991).

Source. Shute, V. (2008, March). Focus on Formative Feedback. *Review of Educational Research*. 78(1), p. 177.

FORMATIVE FEEDBACK GUIDELINES TO ENHANCE LEARNING (THINGS TO AVOID)

Prescription	Description and Reference
Do not give normative comparisons.	Feedback should avoid comparisons with other students-directly or indirectly (e.g., "grading on the curve"). In general, do not draw attention to "self" during learning (Kluger & NeNisi, 1996; Wiliam, 2007).
Be cautious about providing overall grades.	Feedback should note areas of strength and provide information on how to improve, as warranted and without overall grading. Wiliam (2007) summarized the following findings: (a) students receiving just grades showed no learning gains, (b) those getting just comments showed large gains, and (c) those with grades and comments showed no gains (likely due to focusing on the grade and ignoring comments). Effective feedback relates to the content of the comments (Butler, 1987; McColskey & Leary, 1985).
Do not present feedback that discourages the learner or threatens the learner's self-esteem.	This prescription is based not only on common sense but also on research reported in Kluger and DeNisi (1996) citing a list of feedback interventions that undermine learning as it draws focus to the "self" and away from the task at hand. In addition, do not provide feedback that is either too controlling or critical of the learner (Barron, 1993; Fedor et al., 2001)
Use "praise" sparingly, if at all.	Kluger & DeNisi (1996), Butler (1987) and others have noted that use of praise as feedback directs the learner's attention to "self", which distracts from the task and consequently from learning.
Try to avoid delivering feedback orally.	This also was addressed in Kluger & DeNisi (1991). When feedback is delivered in a more neutral manner (e.g., written or computer delivered), it is construed as less biased.
Do not interrupt learner with feedback if learner is actively engaged.	Interrupting a student who is immersed in a task-trying to solve a problem or task on his or her own-can be disruptive to the student and impede learning (Como & Snow, 1986)
Avoid using progressive hints that always terminate with the correct answer.	Although hints can be facilitative, they can also be abused, so if they are employed to scaffold learners, provisions to prevent their abuse should be made (e.g., Aleven & Koedinger, 2000; Shute, Woltz, & Regian, 1989). Consider using prompts and cues (i.e., more specific kinds of hints).
Do not limit the mode of feedback presentation to text.	Exploit the potential of multimedia to avoid cognitive overload due to presenting feedback messages as text. Instead, consider alternative modes of presentation (e.g., acoustic, visual).
Minimize use of extensive error analysis and diagnosis.	In line with findings by Sleeman et al. (1989) and VanLehn et al. (2005), the cost of conducting extensive error analyses and cognitive diagnosis may not provide sufficient benefit to learning. Furthermore, error analyses are rarely complete and not always accurate, thus only helpful in a subset of circumstances.

Source. Shute, V. (2008, March). Focus on Formative Feedback. *Review of Educational Research*. 78(1), p. 178.

APPENDIX B
TEACHER DOMAIN 1 OBSERVATIONAL PROTOCOL

LearningSciencesInternational
LEARNING AND PERFORMANCE MANAGEMENT

Marzano Protocol: Lesson Segments Involving Routine Events

Design Question #1: What will I do to establish and communicate learning goals, track student progress, and celebrate success?

1. Providing Clear Learning Goals and Scales (Rubrics)

The teacher provides a clearly stated learning goal accompanied by scale or rubric that describes levels of performance relative to the learning goal.

Teacher Evidence

- ❒ Teacher has a learning goal posted so that all students can see it
- ❒ The learning goal is a clear statement of knowledge or information as opposed to an activity or assignment
- ❒ Teacher makes reference to the learning goal throughout the lesson
- ❒ Teacher has a scale or rubric that relates to the learning goal posted so that all students can see it
- ❒ Teacher makes reference to the scale or rubric throughout the lesson

Student Evidence

- ❒ When asked, students can explain the learning goal for the lesson
- ❒ When asked, students can explain how their current activities relate to the learning goal
- ❒ When asked, students can explain the meaning of the levels of performance articulated in the scale or rubric

Scale Levels: *(choose one)*
❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Providing clear learning goals and scales (rubrics)	Adapts and creates new strategies for unique student needs and situations.	Provides a clearly stated learning goal accompanied by a scale or rubric that describes levels of performance and monitors students understanding of the learning goal and the levels of performance.	Provides a clearly stated learning goal accompanied by a scale or rubric that describes levels of performance.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Providing clear learning goals and scales (rubrics)	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for providing clearly stated learning goals and rubrics that address the unique student needs and situations?	In addition to providing a clearly stated learning goal accompanied by a scale or rubric that describes levels of performance, how can you monitor students understanding of the learning goal and the levels of performance?	How can you provide a clearly stated learning goal accompanied by a scale or rubric that describes levels of performance?	How can you begin to incorporate some aspects of this strategy into your instruction?

2. Tracking Student Progress

The teacher facilitates tracking of student progress on one or more learning goals using a formative approach to assessment.

Teacher Evidence

- ❒ Teacher helps student track their individual progress on the learning goal
- ❒ Teacher uses formal and informal means to assign scores to students on the scale or rubric depicting student status on the learning goal
- ❒ Teacher charts the progress of the entire class on the learning goal

Student Evidence

- ❒ When asked, students can describe their status relative to the learning goal using the scale or rubric
- ❒ Students systematically update their status on the learning goal

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Tracking student progress	Adapts and creates new strategies for unique student needs and situations.	Facilitates tracking of student progress using a formative approach to assessment and monitors the extent to which students understand their level of performance.	Facilitates tracking of student progress using a formative approach to assessment.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Tracking student progress	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for facilitating tracking of student progress using a formative approach to assessment, that address unique student needs and situations?	In addition to facilitating tracking of student progress using a formative approach to assessment, how can you monitor the extent to which students understand their level of performance?	How can you facilitate tracking of student progress using a formative approach to assessment?	How can you begin to incorporate some aspects of this strategy into your instruction?

3. Celebrating Success

The teacher provides students with recognition of their current status and their knowledge gain relative to the learning goal.

Teacher Evidence

- ❒ Teacher acknowledges students who have achieved a certain score on the scale or rubric
- ❒ Teacher acknowledges students who have made gains in their knowledge and skill relative to the learning goal
- ❒ Teacher acknowledges and celebrates the final status and progress of the entire class
- ❒ Teacher uses a variety of ways to celebrate success
  - Show of hands
  - Certification of success
  - Parent notification
  - Round of applause

Student Evidence

- ❒ Student show signs of pride regarding their accomplishments in the class
- ❒ When asked, students say they want to continue to make progress

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Celebrating success	Adapts and creates new strategies for unique student needs and situations.	Provides students with recognition of their current status and their knowledge gain relative to the learning goal and monitors the extent to which students are motivated to enhance their status.	Provides students with recognition of their current status and their knowledge gain relative to the learning goal.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Celebrating success	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for providing students with recognition of their current status and their knowledge gain relative to the learning goal that address unique student needs and situations?	In addition to providing students with recognition of their current status and their knowledge gain relative to the learning goal, how can you monitor the extent to which students are motivated to enhance their status?	How can you provide students with recognition of their current status and their knowledge gain relative to the learning goal?	How can you begin to incorporate some aspects of this strategy into your instruction?

Student Interviews
Student Questions: • What learning goal did today's lesson focus on? • How well are you doing on that learning goal? • Describe the different levels you can be at on the learning goal.

Design Question #6: What will I do to establish and maintain classroom rules and procedures?

4. Establishing Classroom Routines

The teacher reviews expectations regarding rules and procedures to ensure their effective execution.

Teacher Evidence

- ❒ Teacher involves students in designing classroom routines
- ❒ Teacher uses classroom meetings to review and process rules and procedures
- ❒ Teacher reminds students of rules and procedures
- ❒ Teacher asks students to restate or explain rules and procedures
- ❒ Teacher provides cues or signals when a rule or procedure should be used

Student Evidence

- ❒ Students follow clear routines during class
- ❒ When asked, students can describe established rules and procedures
- ❒ When asked, students describe the classroom as an orderly place
- ❒ Students recognize cues and signals by the teacher
- ❒ Students regulate their own behavior

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Establishing classroom routines	Adapts and creates new strategies for unique student needs and situations.	Establishes and reviews expectations regarding rules and procedures and monitors the extent to which students understand the rules and procedures.	Establishes and reviews expectations regarding rules and procedures.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Establishing classroom routines	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create strategies for establishing and reviewing expectations, rules, and procedures that address unique student needs and situations?	In addition to establishing and reviewing expectations regarding rules and procedures, how can you monitor the extent to which students understand the rules and procedures?	How can you establish and review expectations regarding rules and procedures?	How can you begin to incorporate some aspects of this strategy into your instruction?

5. Organizing the Physical Layout of the Classroom

The teacher organizes the physical layout of the classroom to facilitate movement and focus on learning.

Teacher Evidence

- ❒ The physical layout of the classroom has clear traffic patterns
- ❒ The physical layout of the classroom provides easy access to materials and centers
- ❒ The classroom is decorated in a way that enhances student learning:
  - Bulletin boards relate to current content
  - Students work is displayed

Student Evidence

- ❒ Students move easily about the classroom
- ❒ Students make use of materials and learning centers
- ❒ Students attend to examples of their work that are displayed
- ❒ Students attend to information on the bulletin boards
- ❒ Students can easily focus on instruction

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Organizing the physical layout of the classroom	Adapts and creates new strategies for unique student needs and situations.	Organizes the physical layout of the classroom to facilitate movement and focus on learning and monitors the impact of the environment on student learning.	Organizes the physical layout of the classroom to facilitate movement and focus on learning.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Organizing the physical layout of the classroom	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for organizing the physical layout of the classroom to facilitate movement and focus on learning that address unique student needs and situations?	In addition to organizing the physical layout of the classroom to facilitate movement and focus on learning, how can you monitor the impact of the environment on student learning?	How can you organize the physical layout of the classroom to facilitate movement and focus on learning?	How can you begin to incorporate some aspects of this strategy into your instruction?

Student Interviews

Student Questions:

- What are the regular rules and procedures you are expected to follow in class?
- How well do you do at following the rules and procedures and why?

LearningSciencesInternational
LEARNING AND PERFORMANCE MANAGEMENT

Marzano Protocol: Lesson Segments Enacted on the Spot

Design Question #5: What will I do to engage students?

24. Noticing when Students are Not Engaged

The teacher scans the room making note of when students are not engaged and takes overt action.

Teacher Evidence

- ❒ Teacher notices when specific students or groups of students are not engaged
- ❒ Teacher notices when the energy level in the room is low
- ❒ Teacher takes action to re-engage students

Student Evidence

- ❒ Students appear aware of the fact that the teacher is taking note of their level of engagement
- ❒ Students try to increase their level of engagement when prompted
- ❒ When asked, students explain that the teacher expects high levels of engagement

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Noticing when students are not engaged	Adapts and creates new strategies for unique student needs and situations.	Scans the room making note of when students are not engaged and takes action and monitors the extent to which students re-engage.	Scans the room making note of when students are not engaged and takes action.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Noticing when students are not engaged	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for noticing when students are not engaged that address unique student needs and situations?	In addition to scanning the room, making note of when students are not engaged and taking action, how can you monitor the extent to which students re-engage?	How can you scan the room making note of when students are not engaged and take action to engage students?	How can you begin to incorporate some aspects of this strategy into your instruction?

25. Using Academic Games

The teacher uses academic games and inconsequential competition to maintain student engagement.

Teacher Evidence

- ❒ Teacher uses structured games such as Jeopardy, family feud, and the like
- ❒ Teacher develops impromptu games such as making a game out of which answer might be correct for a given question
- ❒ Teacher uses friendly competition along with classroom games

Student Evidence

- ❒ Students engage in the games with some enthusiasm
- ❒ When asked, students can explain how the games keep their interest and help them learn or remember content

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Using academic games	Adapts and creates new strategies for unique student needs and situations.	Uses academic games and inconsequential competition to maintain student engagement and monitors the extent to which students focus on the academic content of the game.	Uses academic games and inconsequential competition to maintain student engagement.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Using academic games	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for using academic games and inconsequential competition to maintain student engagement that address unique student needs and situations?	In addition to using academic games and inconsequential competition to maintain student engagement, how can you monitor the extent to which students focus on the academic content of the game?	How can you use academic games and inconsequential competition to maintain student engagement?	How can you begin to incorporate this strategy into your instruction?

26. Managing Response Rates

The teacher uses response rate techniques to maintain student engagement in questions.

Teacher Evidence

- ❒ Teacher uses wait time
- ❒ Teacher uses response cards
- ❒ Teacher has students use hand signals to respond to questions
- ❒ Teacher uses choral response
- ❒ Teacher uses technology to keep track of students' responses
- ❒ Teacher uses response chaining

Student Evidence

- ❒ Multiple students or the entire class responds to questions posed by the teacher
- ❒ When asked, students can describe their thinking about specific questions posed by the teacher

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Managing response rates	Adapts and creates new strategies for unique student needs and situations.	Uses response rate techniques to maintain student engagement in questions and monitors the extent to which the techniques keep students engaged.	Uses response rate techniques to maintain student engagement in questions.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Managing response rates	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new response rate techniques to maintain student engagement in questions that address unique student needs and situations?	In addition to using response rate techniques to maintain student engagement in questions, how can you monitor the extent to which the techniques keep students engaged?	How can you use response rate techniques to maintain student engagement in questions?	How can you begin to incorporate this strategy into your instruction?

27. Using Physical Movement

The teacher uses physical movement to maintain student engagement.

Teacher Evidence

- ❒ Teacher has students stand up and stretch or related activities when their energy is low
- ❒ Teacher uses activities that require students to physically move to respond to questions
  - Vote with your feet
  - Go to the part of the room that represents the answer you agree with
- ❒ Teacher has students physically act out or model content to increase energy and engagement
- ❒ Teacher use give-one-get-one activities that require students to move about the room

Student Evidence

- ❒ Students engage in the physical activities designed by the teacher
- ❒ When asked, students can explain how the physical movement keeps their interest and helps them learn

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Using physical movement	Adapts and creates new strategies for unique student needs and situations.	Uses physical movement to maintain student engagement and monitors the extent to which these activities enhance student engagement.	Uses physical movement to maintain student engagement.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Using physical movement	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new physical movement techniques to maintain student engagement that address unique student needs and situations?	In addition to using physical movement to maintain student engagement, how can you monitor the extent to which these activities enhance student engagement?	How can you use physical movement to maintain student engagement?	How can you begin to incorporate this strategy into your instruction?

28. Maintaining a Lively Pace

The teacher uses pacing techniques to maintain students' engagement.

Teacher Evidence
- ❒ Teacher employs crisp transitions from one activity to another
- ❒ Teacher alters pace appropriately (i.e. speeds up and slows down)

Student Evidence
- ❒ Students quickly adapt to transitions and re-engage when a new activity is begun
- ❒ When asked about the pace of the class, students describe it as not too fast or not too slow

Scale Levels: *(choose one)*
☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Maintaining a lively pace	Adapts and creates new strategies for unique student needs and situations.	Uses pacing techniques to maintain students' engagement and monitors the extent to which these techniques keep students engaged.	Uses pacing techniques to maintain students' engagement.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Maintaining a lively pace	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new pacing techniques that address unique student needs and situations?	In addition to pacing techniques to maintain students' engagement, how can you monitor the extent to which students keep engaged?	How can you use pacing techniques to maintain students' engagement?	How can you begin to incorporate this strategy into your instruction?

29. Demonstrating Intensity and Enthusiasm

The teacher demonstrates intensity and enthusiasm for the content in a variety of ways.

Teacher Evidence

- ❒ Teacher describes personal experiences that relate to the content
- ❒ Teacher signals excitement for content by:
  - Physical gestures
  - Voice tone
  - Dramatization of information
- ❒ Teacher overtly adjusts energy level

Student Evidence

- ❒ When asked, students say that the teacher "likes the content" and "likes teaching"
- ❒ Students' attention levels increase when the teacher demonstrates enthusiasm and intensity for the content

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Demonstrating intensity and enthusiasm	Adapts and creates new strategies for unique student needs and situations.	Demonstrates intensity and enthusiasm for the content in a variety of ways and monitors the extent to which students' engagement increases.	Demonstrates intensity and enthusiasm for the content in a variety of ways.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Demonstrating intensity and enthusiasm	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new techniques for demonstrating intensity and enthusiasm for the content that address unique student needs and situations?	In addition to demonstrating intensity and enthusiasm for the content in a variety of ways, how can you monitor the extent to which students keep engaged?	How can you demonstrate intensity and enthusiasm for the content in a variety of ways?	How can you begin to incorporate this strategy into your instruction?

30. Using Friendly Controversy

The teacher uses friendly controversy techniques to maintain student engagement.

Teacher Evidence

- ❒ Teacher structures mini-debates about the content
- ❒ Teacher has students examine multiple perspectives and opinions about the content
- ❒ Teacher elicits different opinions on content from members of the class

Student Evidence

- ❒ Students engage in friendly controversy activities with enhanced engagement
- ❒ When asked, students describe friendly controversy activities as "stimulating," "fun," and so on.
- ❒ When asked, students explain how a friendly controversy activity helped them better understand the content

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Using friendly controversy	Adapts and creates new strategies for unique student needs and situations.	Uses friendly controversy techniques to maintain student engagement and monitors the effect on students' engagement.	Uses friendly controversy techniques to maintain student engagement.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Using friendly controversy	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new techniques for using friendly controversy to maintain student engagement that address unique student needs and situations?	In addition to using friendly controversy techniques to maintain student engagement, how can you monitor the extent to which students keep engaged?	How can you use friendly controversy techniques to maintain student engagement?	How can you begin to incorporate this strategy into your instruction?

31. Providing Opportunities for Students to Talk about Themselves

The teacher provides students with opportunities to relate what is being addressed in class to their personal interests.

Teacher Evidence

- ❒ Teacher is aware of student interests and makes connections between these interests and class content
- ❒ Teacher structures activities that ask students to make connections between the content and their personal interests
- ❒ When students are explaining how content relates to their personal interests, the teacher appears encouraging and interested

Student Evidence

- ❒ Students engage in activities that require them to make connections between their personal interests and the content
- ❒ When asked, students explain how making connections between content and their personal interests engages them and helps them better understand the content

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Providing opportunities for students to talk about themselves	Adapts and creates new strategies for unique student needs and situations.	Provides students with opportunities to relate what is being addressed in class to their personal interests and monitors the extent to which these activities enhance student engagement.	Provides students with opportunities to relate what is being addressed in class to their personal interests.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Providing opportunities for students to talk about themselves	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new techniques for providing students with opportunities to relate what is being addressed in class to their personal interests that address unique student needs and situations?	In addition to providing students with opportunities to relate what is being addressed in class to their personal interests, how can you monitor the extent to which these activities enhance student engagement?	How can you provide students with opportunities to relate what is being addressed in class to their personal interests?	How can you begin to incorporate this strategy into your instruction?

32. Presenting Unusual or Intriguing Information

The teacher uses unusual or intriguing information about the content in a manner that enhances student engagement.

Teacher Evidence

- ❒ Teacher systematically provides interesting facts and details about the content
- ❒ Teacher encourages students to identify interesting information about the content
- ❒ Teacher engages students in activities like "Believe it or not" about the content
- ❒ Teacher uses guest speakers to provide unusual information about the content

Student Evidence

- ❒ Students' attention increases when unusual information is presented about the content
- ❒ When asked, students explain how the unusual information makes them more interested in the content

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Presenting unusual or intriguing information	Adapts and creates new strategies for unique student needs and situations.	Uses unusual or intriguing information about the content and monitors the extent to which this information enhances students' interest in the content.	Uses unusual or intriguing information about the content.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Presenting unusual or intriguing information	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new techniques for using unusual or intriguing information about the content that address unique student needs and situations?	In addition to using unusual or intriguing information about the content, how can you monitor the extent to which this information enhances students' interest in the content?	How can you use unusual or intriguing information about the content?	How can you begin to incorporate this strategy into your instruction?

Student Interviews

Student Questions:

- How engaged were you in this lesson?
- What are some things that keep your attention?
- What are some things that made you bored?

Design Question #7: What will I do to recognize and acknowledge adherence or lack of adherence to rules and procedures?

33. Demonstrating "Withitness"

The teacher uses behaviors associated with "withitness" to maintain adherence to rules and procedures.

Teacher Evidence

- ❒ Teacher physically occupies all quadrants of the room
- ❒ Teacher scans the entire room making eye contact with all students
- ❒ Teacher recognizes potential sources of disruption and deals with them immediately
- ❒ Teacher proactively addresses inflammatory situations

Student Evidence

- ❒ Students recognize that the teacher is aware of their behavior
- ❒ When asked, students describe the teacher as "aware of what is going on" or "has eyes on the back of his/her head"

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Demonstrating "withitness"	Adapts and creates new strategies for unique student needs and situations.	Uses behaviors associated with "withitness" and monitors the effect on students' behavior.	Uses behaviors associated with "withitness".	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Demonstrating "withitness"	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new techniques for using behaviors associated with "withitness" that address unique student needs and situations?	In addition to, using behaviors associated with "withitness," how can you monitor the effect on students' behavior?	How can you use behaviors associated with "withitness"?	How can you begin to incorporate this strategy into your instruction?

34. Applying Consequences for Lack of Adherence to Rules and Procedures

The teacher applies consequences for not following rules and procedures consistently and fairly.

Teacher Evidence

❒ Teacher provides nonverbal signals when students' behavior is not appropriate
- Eye contact
- Proximity
- Tap on the desk
- Shaking head, no

❒ Teacher provides verbal signals when students' behavior is not appropriate
- Tells students to stop
- Tells students that their behavior is in violation of a rule or procedure

❒ Teacher uses group contingency consequences when appropriate (i.e. whole group must demonstrate a specific behavior)
❒ Teacher involves the home when appropriate (i.e. makes a call home to parents to help extinguish inappropriate behavior)
❒ Teacher uses direct cost consequences when appropriate (e.g. student must fix something he or she has broken)

Student Evidence

❒ Students cease inappropriate behavior when signaled by the teacher
❒ Students accept consequences as part of the way class is conducted
❒ When asked, students describe the teacher as fair in application of rules

Scale Levels: *(choose one)*
□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Applying consequences for lack of adherence to rules and procedures	Adapts and creates new strategies for unique student needs and situations.	Applies consequences for not following rules and procedures consistently and fairly and monitors the extent to which rules and procedures are followed.	Applies consequences for not following rules and procedures consistently and fairly.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Applying consequences for lack of adherence to rules and procedures	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for applying consequences for not following rules and procedures consistently and fairly that address unique student needs and situations?	In addition to, applying consequences for not following rules and procedures consistently and fairly, how can you monitor the extent to which rules and procedures are followed?	How can you apply consequences for not following rules and procedures consistently and fairly?	How can you begin to incorporate this strategy into your instruction?

35. Acknowledging Adherence to Rules and Procedures

The teacher consistently and fairly acknowledges adherence to rules and procedures.

Teacher Evidence

- ❒ Teacher provides nonverbal signals that a rule or procedure has been followed:
  - Smile
  - Nod of head
  - High Five
- ❒ Teacher gives verbal cues that a rule or procedure has been followed:
  - Thanks students for following a rule or procedure
  - Describes student behaviors that adhere to rule or procedure
- ❒ Teacher notifies the home when a rule or procedure has been followed
- ❒ Teacher uses tangible recognition when a rule or procedure has been followed:
  - Certificate of merit
  - Token economies

Student Evidence

- ❒ Students appear appreciative of the teacher acknowledging their positive behavior
- ❒ When asked, students describe teacher as appreciative of their good behavior
- ❒ The number of students adhering to rules and procedures increases

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Acknowledging adherence to rules and procedures	Adapts and creates new strategies for unique student needs and situations.	Acknowledges adherence to rules and procedures consistently and fairly and monitors the extent to which new actions affect students' behavior.	Acknowledges adherence to rules and procedures consistently and fairly.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Acknowledging adherence to rules and procedures	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for acknowledging adherence to rules and procedures consistently and fairly that address unique student needs and situations?	In addition to, acknowledging adherence to rules and procedures consistently and fairly, how can you monitor the extent to which new actions affect students' behavior?	How can you acknowledge adherence to rules and procedures consistently and fairly?	How can you begin to incorporate this strategy into your instruction?

Student Interviews

Student Questions:

- How well did you do at following classroom rules and procedures during this lesson?
- What are some things that helped you follow the rules and procedures?
- What are some things that didn't help you follow the rules and procedures?

Design Question #8: What will I do to establish and maintain effective relationships with students?

36. Understanding Students' Interests and Background

The teacher uses students' interests and background to produce a climate of acceptance and community.

Teacher Evidence

- ❒ Teacher has side discussions with students about events in their lives
- ❒ Teacher has discussions with students about topics in which they are interested
- ❒ Teacher builds student interests into lessons

Student Evidence

- ❒ When asked, students describe the teacher as someone who knows them and/or is interested in them
- ❒ Students respond when teacher demonstrates understanding of their interests and background
- ❒ When asked students say they feel accepted

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Understanding students' interests and background	Adapts and creates new strategies for unique student needs and situations.	Uses students' interests and background during interactions with students and monitors the sense of community in the classroom.	Uses students' interests and background during interactions with students.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Understanding students' interests and background	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for using students' interests and backgrounds during interactions with students that address unique student needs and situations?	In addition to using students' interests and background during interactions with students, how can you monitor the extent to which a sense of community is formed in the classroom?	How can you use students' interests and background during interactions with students?	How can you begin to incorporate this strategy into your instruction?

37. Using Verbal and Nonverbal Behaviors that Indicate Affection for Students

When appropriate, the teacher uses verbal and nonverbal behavior that indicates caring for students.

Teacher Evidence

- ❒ Teacher compliments students regarding academic and personal accomplishments
- ❒ Teacher engages in informal conversations with students that are not related to academics
- ❒ Teacher uses humor with students when appropriate
- ❒ Teacher smiles, nods, (etc) at students when appropriate
- ❒ Teacher puts hand on students' shoulders when appropriate

Student Evidence

- ❒ When asked, students describe teacher as someone who cares for them
- ❒ Students respond to teachers verbal interactions
- ❒ Students respond to teachers nonverbal interactions

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Using verbal and nonverbal behaviors that indicate caring for students	Adapts and creates new strategies for unique student needs and situations.	Uses verbal and nonverbal behaviors that indicate caring for students and monitors the quality of relationships in the classroom.	Uses verbal and nonverbal behaviors that indicate caring for students.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Using verbal and nonverbal behaviors that indicate caring for students	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for using verbal and nonverbal behaviors that indicate caring for students that address unique student needs and situations?	In addition to using verbal and nonverbal behaviors that indicate caring for students how can you monitor the quality of relationships in the classroom?	How can you use verbal and nonverbal behaviors that indicate caring for students?	How can you begin to incorporate this strategy into your instruction?

38. Displaying Objectivity and Control

The teacher behaves in an objective and controlled manner.

Teacher Evidence

- ❒ Teacher does not exhibit extremes in positive or negative emotions
- ❒ Teacher addresses inflammatory issues and events in a calm and controlled manner
- ❒ Teacher interacts with all students in the same calm and controlled fashion
- ❒ Teacher does not demonstrate personal offense at student misbehavior

Student Evidence

- ❒ Students are settled by the teacher's calm demeanor
- ❒ When asked, the students describe the teacher as in control of himself/herself and in control of the class
- ❒ When asked, students say that the teacher does not hold grudges or take things personally

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Displaying emotional objectivity and control	Adapts and creates new strategies for unique student needs and situations.	Behaves in an objective and controlled manner and monitors the effect on the classroom climate.	Behaves in an objective and controlled manner.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Displaying emotional objectivity and control	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for behaving in an objective and controlled manner that address unique student needs and situations?	In addition to behaving in an objective and controlled manner, how can you monitor the effect on the classroom climate?	How can you behave in an objective and controlled manner?	How can you begin to incorporate this strategy into your instruction?

Student Interviews

Student Questions:

- How much did you feel accepted and welcomed in the class today?
- What are some things that made you feel accepted and welcomed?
- What are some things that did not make you feel accepted and welcomed?

Design Question #9: What will I do to communicate high expectations for all students?

39. Demonstrating Value and Respect for Low Expectancy Students

The teacher exhibits behaviors that demonstrate value and respect for low expectancy students.

Teacher Evidence

❒ When asked, the teacher can identify the students for whom there have been low expectations and the various ways in which these students have been treated differently from high expectancy students

❒ The teacher provides low expectancy with nonverbal indications that they are valued and respected:

- Makes eye contact
- Smiles
- Makes appropriate physical contact

❒ The teacher proves low expectancy students with verbal indications that they are valued and respected:

- Playful dialogue
- Addressing students in a manner they view as respectful

❒ Teacher does not allow negative comments about low expectancy students

Student Evidence

❒ When asked, students say that the teacher cares for all students

❒ Students treat each other with respect

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Communicating value and respect for low expectancy students	Adapts and creates new strategies for unique student needs and situations.	Exhibits behaviors that demonstrate value and respect for low expectancy students and monitors the impact on low expectancy students.	Exhibits behaviors that demonstrate value and respect for low expectancy students.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Communicating value and respect for low expectancy students	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for behaviors that demonstrate value and respect for low expectancy students that address unique student needs and situations?	In addition to exhibiting behaviors that demonstrate value and respect for low expectancy students, how can you monitor the impact on low expectancy students?	How can you exhibit behaviors that demonstrate value and respect for low expectancy students?	How can you begin to incorporate this strategy into your instruction?

40. Asking Questions of Low Expectancy Students

The teacher asks questions of low expectancy students with the same frequency and depth as with high expectancy students.

Teacher Evidence

- ❒ Teacher makes sure low expectancy students are asked questions at the same rate as high expectancy students
- ❒ Teacher makes sure low expectancy students are asked complex questions at the same rate as high expectancy students

Student Evidence

- ❒ When asked, students say the teacher expects everyone to participate
- ❒ When asked, students say the teacher asks difficult questions of every student

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Asking questions of low expectancy students	Adapts and creates new strategies for unique student needs and situations.	Asks questions of low expectancy students with the same frequency and depth with high expectancy students and monitors the quality of participation of low expectancy students.	Asks questions of low expectancy students with the same frequency and depth as with high expectancy students.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Asking questions of low expectancy students	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies and techniques for asking questions of low expectancy students that address unique student needs and situations?	In addition to asking questions of low expectancy students with the same frequency and depth as with high expectancy students, how can you monitor the quality of participation of low expectancy students?	How can you ask questions of low expectancy students with the same frequency and depth as with high expectancy students?	How can you begin to incorporate this strategy into your instruction?

41. Probing Incorrect Answers with Low Expectancy Students

The teacher probes incorrect answers of low expectancy students in the same manner as he/she does with high expectancy students.

Teacher Evidence

- ❒ Teacher asks low expectancy students to further explain their answers when they are incorrect
- ❒ Teacher rephrases questions for low expectancy students when they provide an incorrect answer
- ❒ Teacher breaks a question into smaller and simpler parts when a low expectancy student answers a question incorrectly
- ❒ When low expectancy students demonstrate frustration, the teacher allows them to collect their thoughts but goes back to them at a later point in time

Student Evidence

- ❒ When asked, students say that the teacher won't "let you off the hook"
- ❒ When asked, students say that the teacher "won't give up on you"
- ❒ When asked, students say the teacher helps them answer questions successfully

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Probing incorrect answers by low expectancy students	Adapts and creates new strategies for unique student needs and situations.	Probes incorrect answers of low expectancy students in the same manner as with high expectancy students and monitors the level and quality responses of low expectancy students.	Probes incorrect answers of low expectancy students in the same manner as with high expectancy students.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Probing incorrect answers by low expectancy students	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for probing incorrect answers of low expectancy students in the same manner as with high expectancy students that address their unique student needs and situations?	In addition to probing incorrect answers of low expectancy students in the same manner as with high expectancy students, how can you monitor the level and quality responses of low expectancy students?	How can you probe incorrect answers of low expectancy students in the same manner as with high expectancy students?	How can you begin to incorporate this strategy into your instruction?

Student Interviews
Student Questions: • How does your teacher demonstrate that they care and respect you? • How does your teacher communicate that everyone is expected to participate and answer difficult questions? • What are some ways that your teacher helps you answer questions successfully?

LearningSciencesInternational
LEARNING AND PERFORMANCE MANAGEMENT

Marzano Protocol: Lesson Segments Addressing Content

Design Question #2: What will I do to help students effectively interact with new knowledge?

6. Identifying Critical Information

The teacher identifies a lesson or part of a lesson as involving important information to which students should pay particular attention.

Teacher Evidence

- ❒ Teacher begins the lesson by explaining why upcoming content is important
- ❒ Teacher tells students to get ready for some important information
- ❒ Teacher cues the importance of upcoming information in some indirect fashion
  - Tone of voice
  - Body position
  - Level of excitement

Student Evidence

- ❒ When asked, students can describe the level of importance of the information addressed in class
- ❒ When asked, students can explain why the content is important to pay attention to
- ❒ Students visibly adjust their level of engagement

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Identifying critical information	Adapts and creates new strategies for unique student needs and situations.	Signals to students which content is critical versus non-critical and monitors the extent to which students are attending to critical information.	Signals to students which content is critical versus non-critical.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Identifying critical information	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for identifying critical information that address unique student needs and situations?	In addition to signaling to students which content is critical versus non-critical, how might you monitor the extent to which students attend to critical information?	How can you signal to students which content is critical versus non-critical?	How can you begin to incorporate some aspect of this strategy in your instruction?

7. Organizing Students to Interact with New Knowledge

The teacher organizes students into small groups to facilitate the processing of new information.

Teacher Evidence

- ❒ Teacher has established routines for student grouping and student interaction in groups
- ❒ Teacher organizes students into ad hoc groups for the lesson
  - Diads
  - Triads
  - Small groups up to about 5

Student Evidence

- ❒ Students move to groups in an orderly fashion
- ❒ Students appear to understand expectations about appropriate behavior in groups
  - Respect opinions of others
  - Add their perspective to discussions
  - Ask and answer questions

Scale Levels: *(choose one)*
❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students to interact with new knowledge	Adapts and creates new strategies for unique student needs and situations.	Organizes students into small groups to facilitate the processing of new knowledge and monitors group processing.	Organizes students into small groups to facilitate the processing of new knowledge.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students to interact with new knowledge	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for organizing students to interact with new knowledge that address unique student needs and situations?	In addition to organizing students into small groups to facilitate the processing of new knowledge, how can you monitor group processes?	How can you organize students into small groups to facilitate the processing of new knowledge?	How can you begin to incorporate some aspect of this strategy in your instruction?

8. Previewing New Content

The teacher engages students in activities that help them link what they already know to the new content about to be addressed and facilitates these linkages.

Teacher Evidence

- ❒ Teacher uses preview question before reading
- ❒ Teacher uses K-W-L strategy or variation of it
- ❒ Teacher asks or reminds students what they already know about the topic
- ❒ Teacher provides an advanced organizer
  - Outline
  - Graphic organizer
- ❒ Teacher has students brainstorm
- ❒ Teacher uses anticipation guide
- ❒ Teacher uses motivational hook/launching activity
  - Anecdotes
  - Short selection from video
- ❒ Teacher uses word splash activity to connect vocabulary to upcoming content

Student Evidence

- ❒ When asked, students can explain linkages with prior knowledge
- ❒ When asked, students make predictions about upcoming content
- ❒ When asked, students can provide a purpose for what they are about to learn
- ❒ Students actively engage in previewing activities

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Previewing new content	Adapts and creates new strategies for unique student needs and situations.	Engages students in learning activities that require them to preview and link new knowledge to what has been addressed and monitors the extent to which students are making linkages.	Engages students in learning activities that require them to preview and link new knowledge to what has been addressed.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Previewing new content	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for previewing new content that address unique student needs and situations?	In addition to engaging students in learning activities that require them to preview and link new knowledge to what has been addressed, how can you also monitor the extent to which students are making linkages?	How can you engage students in learning activities that require them to preview and link new knowledge to what has been addressed?	How can you begin to incorporate some aspect of this strategy in your instruction?

9. Chunking Content into "Digestible Bites"

Based on student needs, the teacher breaks the content into small chunks (i.e. digestible bites) of information that can be easily processed by students.

Teacher Evidence

- ❒ Teacher stops at strategic points in a verbal presentation
- ❒ While playing a video tape, the teacher turns the tape off at key junctures
- ❒ While providing a demonstration, the teacher stops at strategic points
- ❒ While students are reading information or stories orally as a class, the teacher stops at strategic points

Student Evidence

- ❒ When asked, students can explain why the teacher is stopping at various points
- ❒ Students appear to know what is expected of them when the teacher stops at strategic points

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Chunking content into digestible bites	Adapts and creates new strategies for unique student needs and situations.	Breaks input experiences into small chunks based on student needs and monitors the extent to which chunks are appropriate.	Breaks input experiences into small chunks based on student needs.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Chunking content into digestible bites	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for chunking content into digestible bites that address unique student needs and situations?	In addition to breaking input experiences into small chunks based on student needs, how can you also monitor the extent to which chunks are appropriate?	How can you break input experiences into small chunks based on student needs?	How can you begin to incorporate some aspect of this strategy in your instruction?

10. Processing New Information

During breaks in the presentation of content, the teacher engages students in actively processing new information.

Teacher Evidence

- ❒ Teacher has group members summarize new information
- ❒ Teacher employs formal group processing strategies
  - Jigsaw
  - Reciprocal Teaching
  - Concept attainment

Student Evidence

- ❒ When asked, students can explain what they have just learned
- ❒ Students volunteer predictions
- ❒ Students voluntarily ask clarification questions
- ❒ Groups are actively discussing the content
  - Group members ask each other and answer questions about the information
  - Group members make predictions about what they expect next

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Processing new information	Adapts and creates new strategies for unique student needs and situations.	Engages students in summarizing, predicting, and questioning activities and monitor the extent to which the activities enhance students' understanding.	Engages students in summarizing, predicting, and questioning activities.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Processing new information	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for processing new information that address unique student needs and situations?	In addition to engaging students in summarizing, predicting, and questioning activities, how can you monitor the extent to which the activities enhance students' understanding?	How can you engage students in summarizing, predicting, and questioning activities?	How can you begin to incorporate some aspect of this strategy in your instruction?

11. Elaborating on New Information

The teacher asks questions or engages students in activities that require elaborative inferences that go beyond what was explicitly taught.

Teacher Evidence

- ❒ Teacher asks explicit questions that require students to make elaborative inferences about the content
- ❒ Teacher asks students to explain and defend their inferences
- ❒ Teacher presents situations or problems that require inferences

Student Evidence

- ❒ Students volunteer answers to inferential questions
- ❒ Students provide explanations and "proofs" for inferences

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Elaborating on new information	Adapts and creates new strategies for unique student needs and situations.	Engages students in answering inferential questions and monitors the extent to which students elaborate on what was explicitly taught.	Engages students in answering inferential questions.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Elaborating on new information	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for elaborating on new information that address unique student needs and situations?	In addition to engaging students in answering inferential questions, how can you monitor the extent to which students elaborate on what was explicitly taught?	How can you engage students in answering inferential questions?	How can you begin to incorporate some aspect of this strategy in your instruction?

12. Recording and Representing Knowledge

The teacher engages students in activities that help them record their understanding of new content in linguistic ways and/or represent the content in nonlinguistic ways.

Teacher Evidence

- ❒ Teacher asks students to summarize the information they have learned
- ❒ Teacher asks students to generate notes that identify critical information in the content
- ❒ Teacher asks students to create nonlinguistic representations for new content
  - Graphic organizers
  - Pictures
  - Pictographs
  - Flow charts
- ❒ Teacher asks students to create mnemonics that organize the content

Student Evidence

- ❒ Students' summaries and notes include critical content
- ❒ Students' nonlinguistic representations include critical content
- ❒ When asked, students can explain main points of the lesson

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Recording and representing knowledge	Adapts and creates new strategies for unique student needs and situations.	Engages students in activities that help them record their understanding of new content in linguistic ways and/or in nonlinguistic ways and monitors the extent to which this enhances students' understanding.	Engages students in activities that help them record their understanding of new content in linguistic ways and/or in nonlinguistic ways.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Recording and representing knowledge	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for recording and representing knowledge that address unique student needs and situations?	In addition to engaging students in activities that help them record their understanding of new content in linguistic ways and/or in nonlinguistic ways, how can you monitor the extent to which this enhances students' understanding?	How can you engage students in activities that help them record their understanding of new content in linguistic ways and/or in nonlinguistic ways?	How can you begin to incorporate some aspect of this strategy in your instruction?

13. Reflecting on Learning

The teacher engages students in activities that help them reflect on their learning and the learning process.

Teacher Evidence
- ❒ Teacher asks students to state or record what they are clear about and what they are confused about
- ❒ Teacher asks students to state or record how hard they tried
- ❒ Teacher asks students to state or record what they might have done to enhance their learning

Student Evidence
- ❒ When asked, students can explain what they are clear about and what they are confused about
- ❒ When asked, students can describe how hard they tried
- ❒ When asked, students can explain what they could have done to enhance their learning

Scale Levels: *(choose one)*
❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Reflecting on learning	Adapts and creates new strategies for unique student needs and situations.	Engages students in reflecting on their own learning and the learning process and monitors the extent to which students self-assess their understanding and effort.	Engages students in reflecting on their own learning and the learning process.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Reflecting on learning	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for reflecting on learning that address unique student needs and situations?	In addition to engaging students in reflecting on their own learning and the learning process, how can you monitor the extent to which students self-assess their understanding and effort?	How can you engage students in reflecting on their own learning and the learning process?	How can you begin to incorporate some aspect of this strategy in your instruction?

Student Interviews

Student Questions:
- Why is the information that you are learning today important?
- How do you know what are the most important things to pay attention to?
- What are the main points of this lesson?

Design Question #3: What will I do to help students practice and deepen their understanding of new knowledge?

14. Reviewing Content

The teacher engages students in a brief review of content that highlights the critical information.

Teacher Evidence

❒ Teacher begins the lesson with a brief review of content
❒ Teacher uses specific strategies to review information
- Summary
- Problem that must be solved using previous information
- Questions that require a review of content
- Demonstration
- Brief practice test or exercise

Student Evidence

❒ When asked, students can describe the previous content on which new lesson is based
❒ Student responses to class activities indicate that they recall previous content

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Reviewing content	Adapts and creates new strategies for unique student needs and situations.	Engages students in a brief review of content that highlights the critical information and monitors the extent to which students can recall and describe previous content.	Engages students in a brief review of content that highlights the critical information.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Reviewing content	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for reviewing content that address unique student needs and situations?	In addition to, engaging students in a brief review of content, how can you monitor the extent to which students can recall and describe previous content?	How can you engage students in a brief review of content that highlights the critical information?	How can you begin to incorporate some aspect of this strategy in your instruction?

15. Organizing Students to Practice and Deepen Knowledge

The teacher uses grouping in ways that facilitate practicing and deepening knowledge.

Teacher Evidence
❒ Teacher organizes students into groups with the expressed idea of deepening their knowledge of informational content
❒ Teacher organizes students into groups with the expressed idea of practicing a skill, strategy, or process

Student Evidence
❒ When asked, students explain how the group work supports their learning
❒ While in groups students interact in explicit ways to deepen their knowledge of informational content or, practice a skill, strategy, or process
- Asking each other questions
- Obtaining feedback from their peers

Scale Levels: *(choose one)*
□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students to practice and deepen knowledge	Adapts and creates new strategies for unique student needs and situations.	Organizes students into groups to practice and deepen their knowledge and monitors the extent to which the group work extends their learning.	Organizes students into groups to practice and deepen their knowledge.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students to practice and deepen knowledge	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for organizing students to practice and deepen knowledge that address unique student needs and situations?	In addition to organizing students into groups to practice and deepen their knowledge, how can you also monitor the extent to which the group work extends their learning?	How can you organize students into groups to practice and deepen their knowledge?	How can you begin to incorporate some aspect of this strategy in your instruction?

16. Using Homework

When appropriate (as opposed to routinely) the teacher designs homework to deepen students' knowledge of informational content or, practice a skill, strategy, or process.

Teacher Evidence

- ❒ Teacher communicates a clear purpose for homework
- ❒ Teacher extends an activity that was begun in class to provide students with more time
- ❒ Teacher assigns a well crafted homework assignment that allows students to practice and deepen their knowledge independently

Student Evidence

- ❒ When asked, students can describe how the homework assignment will deepen their understanding of informational content or, help them practice a skill, strategy, or process
- ❒ Students ask clarifying questions of the homework that help them understand its purpose

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Using homework	Adapts and creates new strategies for unique student needs and situations.	When appropriate (as opposed to routinely) assigns homework that is designed to deepen knowledge of informational content or, practice a skill, strategy, or process and monitors the extent to which students understand the homework.	When appropriate (as opposed to routinely) assigns homework that is designed to deepen knowledge of informational content or, practice a skill, strategy, or process.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Using homework	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for assigning homework that address unique student needs and situations?	In addition to assigning homework that is designed to deepen knowledge of informational content or practice a skill, strategy, or process, how can you also monitor the extent to which the group work extends their learning?	How can you assign homework that is designed to deepen knowledge of informational content or practice a skill, strategy, or process?	How can you begin to incorporate some aspect of this strategy in your instruction?

17. Examining Similarities and Differences

When the content is informational, the teacher helps students deepen their knowledge by examining similarities and differences.

Teacher Evidence

❒ Teacher engages students in activities that require students to examine similarities and differences between content
- Comparison activities
- Classifying activities
- Analogy activities
- Metaphor activities

❒ Teacher facilitates the use of these activities to help students deepen their understanding of content
- Ask students to summarize what they have learned from the activity
- Ask students to explain how the activity has added to their understanding

Student Evidence

❒ Student artifacts indicate that their knowledge has been extended as a result of the activity
❒ When asked about the activity, student responses indicate that they have deepened their understanding
❒ When asked, students can explain similarities and differences
❒ Student artifacts indicate that they can identify similarities and differences

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Examining similarities and differences	Adapts and creates new strategies for unique student needs and situations.	When content is informational, engages students in activities that require them to examine similarities and differences, and monitors the extent to which the students are deepening their knowledge.	When content is informational, engages students in activities that require them to examine similarities and differences.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Examining similarities and differences	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for examining similarities and differences that address unique student needs and situations?	In addition to engaging students in examining similarities and differences, how can you monitor the extent to which the students are deepening their knowledge?	How can you engage students in activities that require them to examine similarities and differences?	How can you begin to incorporate some aspect of this strategy in your instruction?

18. Examining Errors in Reasoning

When content is informational, the teacher helps students deepen their knowledge by examining their own reasoning or the logic of the information as presented to them.

Teacher Evidence

❒ Teacher asks students to examine information for errors or informal fallacies
- Faulty logic
- Attacks
- Weak reference
- Misinformation

❒ Teacher asks students to examine the strength of support presented for a claim
- Statement of a clear claim
- Evidence for the claim presented
- Qualifiers presented showing exceptions to the claim

Student Evidence

❒ When asked, students can describe errors or informal fallacies in information
❒ When asked, students can explain the overall structure of an argument presented to support a claim
❒ Student artifacts indicate that they can identify errors in reasoning.

Scale Levels: *(choose one)*
□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Examining errors in reasoning	Adapts and creates new strategies for unique student needs and situations.	When content is informational, engages students in activities that require them to examine their own reasoning or the logic of information as presented to them and monitors the extent to which students are deepening their knowledge.	When content is informational, engages students in activities that require them to examine their own reasoning or the logic of information as presented to them.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Examining errors in reasoning	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for examining their own reasoning or the logic of information that address unique student needs and situations?	In addition to engaging students in examining their own reasoning or the logic of information as presented to them, how can you monitor the extent to which the students are deepening their knowledge?	How can you engage students in activities that require them to examine their own reasoning or the logic of information as presented to them?	How can you begin to incorporate some aspect of this strategy in your instruction?

19. Practicing Skills, Strategies, and Processes

When the content involves a skill, strategy, or process, the teacher engages students in practice activities that help them develop fluency.

Teacher Evidence

❒ Teacher engages students in massed and distributed practice activities that are appropriate to their current ability to execute a skill, strategy, or process

- Guided practice if students cannot perform the skill, strategy, or process independently
- Independent practice if students can perform the skill, strategy, or process independently

Student Evidence

❒ Students perform the skill, strategy, or process with increased confidence
❒ Students perform the skill, strategy, or process with increased competence

Scale Levels: *(choose one)*
□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Practicing skills, strategies, and processes	Adapts and creates new strategies for unique student needs and situations.	When content involves a skill, strategy, or process, engages students in practice activities and monitors the extent to which the practice is increasing student fluency.	When content involves a skill, strategy, or process, engages students in practice activities.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Practicing skills, strategies, and processes	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create practice activities that increase fluency and address unique student needs and situations?	In addition to engaging students in practice activities, how can you monitor the extent to which the practice is increasing student fluency?	How can you engage students in practice activities when content involves a skill, strategy, or process?	How can you begin to incorporate some aspect of this strategy in your instruction?

20. Revising Knowledge

The teacher engages students in revision of previous knowledge about content addressed in previous lessons.

Teacher Evidence

- ❒ Teacher asks students to examine previous entries in their academic notebooks or notes
- ❒ The teacher engages the whole class in an examination of how the current lesson changed perceptions and understandings of previous content
- ❒ Teacher has students explain how their understanding has changed

Student Evidence

- ❒ Students make corrections to information previously recorded about content
- ❒ When asked, students can explain previous errors or misconceptions they had about content

Scale Levels: *(choose one)*

☐ Innovating ☐ Applying ☐ Developing ☐ Beginning ☐ Not Using ☐ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Revising knowledge	Adapts and creates new strategies for unique student needs and situations.	Engages students in revision of previous content and monitors the extent to which these revisions deepen students' understanding.	Engages students in revision of previous content.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Revising knowledge	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for revising content that address unique student needs and situations?	In addition to engaging students in revision of previous content, how can you monitor the extent to which these revisions deepen students' understanding?	How can you engage students in the revision of previous content?	How can you begin to incorporate some aspect of this strategy in your instruction?

Student Interviews

Student Questions:

- How did this lesson add to your understanding of the content?
- What changes did you make in your understanding of the content as a result of the lesson?
- What do you still need to understand better?

Design Question #4: What will I do to help students generate and test hypotheses about new knowledge?

21. Organizing Students for Cognitively Complex Tasks

The teacher organizes the class in such a way as to facilitate students working on complex tasks that require them to generate and test hypotheses.

Teacher Evidence

- ❒ Teacher establishes the need to generate and test hypotheses
- ❒ Teacher organizes students into groups to generate and test hypotheses

Student Evidence

- ❒ When asked, students describe the importance of generating and testing hypotheses about content
- ❒ When asked, students explain how groups support their learning
- ❒ Students use group activities to help them generate and test hypotheses

Scale Levels: *(choose one)*

❒ Innovating ❒ Applying ❒ Developing ❒ Beginning ❒ Not Using ❒ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students for cognitively complex tasks	Adapts and creates new strategies for unique student needs and situations.	Organizes students into groups to facilitate working on cognitively complex tasks and monitors the extent to which group processes facilitate generating and testing hypotheses.	Organizes students into groups to facilitate working on cognitively complex tasks.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Organizing students for cognitively complex tasks	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for organizing students to complete cognitively complex tasks?	In addition to organizing students in groups for cognitively complex tasks, how can you monitor the extent to which group processes facilitate generating and testing hypotheses?	How can you organize students in groups to facilitate working on cognitively complex tasks?	How can you begin to incorporate some aspect of this strategy in your instruction?

22. Engaging Students in Cognitively Complex Tasks Involving Hypothesis Generation and Testing

The teacher engages students in complex tasks (e.g. decision making, problem solving, experimental inquiry, investigation) that require them to generate and test hypotheses.

Teacher Evidence

❒ Teacher engages students with an explicit decision making, problem solving, experimental inquiry, or investigation task that requires them to generate and test hypotheses

❒ Teacher facilitates students generating their own individual or group task that requires them to generate and test hypotheses

Student Evidence

❒ Students are clearly working on tasks that require them to generate and test hypotheses

❒ When asked, students can explain the hypothesis they are testing

❒ When asked, students can explain whether their hypothesis was confirmed or disconfirmed

❒ Student artifacts indicate that they can engage in decision making, problem solving, experimental inquiry, or investigation

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Engaging students in cognitively complex tasks involving hypothesis generation and testing	Adapts and creates new strategies for unique student needs and situations.	Engages students in cognitively complex tasks (e.g. decision making, problem solving, experimental inquiry, investigation) and monitors the extent to which students are generating and testing hypotheses.	Engages students in cognitively complex tasks (e.g. decision making, problem solving, experimental inquiry, investigation).	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Engaging students in cognitively complex tasks involving hypothesis generation and testing	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for organizing students to complete cognitively complex tasks?	In addition to engaging students in groups for cognitively complex tasks, involving hypothesis generation and testing, how can you monitor the extent to which students are generating and testing hypotheses?	How can you engage students in cognitively complex tasks involving hypothesis generation and testing?	How can you begin to incorporate some aspect of this strategy in your instruction?

23. Providing Resources and Guidance

The teacher acts as resource provider and guide as students engage in cognitively complex tasks

Teacher Evidence

❒ Teacher makes himself/herself available to students who need guidance or resources
- Circulates around the room
- Provides easy access to himself/herself

❒ Teacher interacts with students during the class to determine their needs for hypothesis generation and testing tasks
❒ Teacher volunteers resources and guidance as needed by the entire class, groups of students, or individual students

Student Evidence

❒ Students seek out the teacher for advice and guidance regarding hypothesis generation and testing tasks
❒ When asked, students can explain how the teacher provides assistance and guidance in hypothesis generation and testing tasks

Scale Levels: *(choose one)*

□ Innovating □ Applying □ Developing □ Beginning □ Not Using □ Not Applicable

Scale

	Innovating	Applying	Developing	Beginning	Not Using
Providing resources and guidance	Adapts and creates new strategies for unique student needs and situations.	Acts as a guide and resource provider as students engage in cognitively complex tasks and monitors the extent to which students request and use guidance and resources.	Acts as a guide and resource provider as students engage in cognitively complex tasks.	Uses strategy incorrectly or with parts missing.	Strategy was called for but not exhibited.

Reflection Questions

	Innovating	Applying	Developing	Beginning	Not Using
Providing resources and guidance	What are you learning about your students as you adapt and create new strategies?	How might you adapt and create new strategies for providing resources and guidance?	In addition to acting as a guide and resource provider, how can you monitor the extent to which students request and use guidance and resources?	How can you act as a guide and resource provider as students engage in cognitively complex tasks?	How can you begin to incorporate some aspect of this strategy in your instruction?

Student Interviews

Student Questions:
- How did this lesson help you apply or use what you have learned?
- What change has this lesson made about your understanding of the content?

APPENDIX C
LARGE URBAN SCHOOL DISTRICT APPROVAL OF RESEARCH

Orange County Public Schools
445 West Amelia Street • Orlando, FL 32801-1129 • Phone 407.317.3200 • www.ocps.net

July 22, 2014

Office of Research and Commercialization
12201 Research Parkway
Suite 501
Orlando, FL 32826-3246

IRB Committee Members:

Orange County Public Schools would like to formally express its intent to support dissertation research to be conducted by Shana Rafalski entitled 'Policy Implication of a Teacher Evaluation System: The Relationship between Frequency of Teacher Observation, Forms of Feedback and Student Achievement'.

This research is aligned directly with district goals in improving the Instructional Practice portion of our teacher evaluations. This analysis supports multiple strategies identified in our district business plan as well. We look forward to using the results of this analysis to inform district practice and improve instruction.

If we can provide any additional information to expedite the review, please contact the Accountability, Research and Assessment office.

Sincerely,

J. Brandon McKelvey
Senior Director
Accountability, Research, and Assessment

APPENDIX D
UCF INSTITUTIONAL REVIEW BOARD APPROVAL OF RESEARCH

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

NOT HUMAN RESEARCH DETERMINATION

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

To : **Shana Rafalski**

Date : **July 28, 2014**

Dear Researcher:

On 7/28/2014 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:	Policy Implication of a Teacher Evaluation System: The Relationship Between Frequency of Teacher Observation, Forms of Feedback and Student Achievement
Investigator:	Shana Rafalski
IRB ID:	SBE-14-10456
Funding Agency:	
Grant Title:	
Research ID:	na

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 Joanne Muratori on 07/28/2014 12:09:22 PM EDT

Joanne Muratori

IRB Coordinator

APPENDIX E
SAMPLE RUBRICS

RUBRIC FOR GENERALIZING THE TYPES OF OVERALL FEEDBACK INDIVIDUAL TEACHERS RECEIVED

VAM impact	No Feedback Provided (Level 1)	Unrelated Feedback or General Statement Provided (Level 2)	Recount of Classroom Events (Level 3)- Justification for rating	General Affirmation Statement (Level 4)	Reflective Question (Level 5)	Standardized Feedback Provided (Level 6)	Specific Targeted Feedback Provided (Level 7)
Predominant feedback type provided to teachers in sample.							

TYPES OF FEEDBACK PROVIDED IN THE COMMENT SECTION OF EACH ELEMENT

No Feedback Provided (Level 1)	Unrelated Feedback or General Statement Provided (Level 2)	Recount of Classroom Events (Level 3)- Justification for rating	General Affirmation Statement (Level 4)	Reflective Question (Level 5)	Standardized Feedback Provided (Level 6)	Specific Targeted Feedback Provided (Level 7)
	2-The message is Unintelligible	3-Recap has several different components (sometimes statement of percent of students being monitored or desired effect). Sometimes actually gives examples of what is wrong with no suggestion. You did this or that, teacher did this or	4-General praise. Good job, great job, excellent job, I liked, I loved, WOW!	5- Asks the teacher a question.	6- Examples: How might you adapt and create new strategies for chunking content into digestible bites that address unique student needs and situations? How might you expand your monitoring to involve more	7-Language like: 1. Reference to Resource Library or Reflective Teacher 2. Maybe try…. Or You might want to try…. 3. Consider…. 4. Recommendation…. Or I would recommend…. 5. Suggestion…. Or I suggest…. 6. It might be a good idea… 7. You should…. 8. This would have been good or great if…

No Feedback Provided (Level 1)	Unrelated Feedback or General Statement Provided (Level 2)	Recount of Classroom Events (Level 3)- Justification for rating	General Affirmation Statement (Level 4)	Reflective Question (Level 5)	Standardized Feedback Provided (Level 6)	Specific Targeted Feedback Provided (Level 7)
		that, students did this or that, I observed this or that…			students? What are you learning about your students as you adapt and create new strategies? In addition to monitoring students by the use of choral responses, how else can you monitor students when chunking information?	9. To move to a higher level, do this_____. 10. Think about….. 11. I want you to…. 12. Next time…. 13. Always…. 14. Be sure to… or Make sure you….. 15. Doing this_____would have been more effective. 16. You need to…. 17. Continue to… 18. Remember this_____. 19. Do this_____. 20. This_____is a good strategy. 21. I would like to see…. 22. Coaching idea…

REFERENCES

Asay, D., & Schafer, M. J. (2013). *Value-added teacher assessment in Louisiana*. LSU AgCenter. Retrieved from http://www.lsuagcenter.com/en/communications/publications/agmag/Archive/2013/Winter/ValueAdded-Teacher-Assessment-in-Louisiana.htm

Balcazar, E., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management, 7*, 65-89.

Ballou, D. (2010). *Value-added assessment: Lessons from Tennessee*. Retrieved from dpi.state.nc.us/docs/superintendents/quarterly/2010-11/201009228/ballou-lessons.pdf

Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*, 410-433.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Reviews*. *84*, 191-215.

Bill & Melinda Gates Foundation. (January, 2013). Retrieved from: http://www.metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf

Black, P., & Wiliam, D. (2010, September). Inside the black box: Raising standards through classroom assessment. *Kappan*, *92*(1), 81-90.

Booth-Butterfield, M. (1989, April). The interpretation of classroom performance feedback: an attributional approach. *Communication Education*, 38. 119-131.

Brauer, J. R. & Tittle, C. R. (2012) Social learning theory and human reinforcement. *Sociological Spectrum: Mid-South Sociological Association*, *32*(2), 157-177, DOI:10.1080/027321173.2012.646160.

Carless, D. (2006). Differing perceptions of the feedback process. *Studies in Higher Education, 31*(2), 219-33.

Cavalluzzo, L. C., (2004). *Is National Board Certification an effective signal of teacher quality?* (Report No. (IRP) 11204). Alexandria, VA: CNA. Retrieved from http://www.cna.org/documents/cavaluzzostudy.pdf

Chapter 1012.34 (2013). Title 48.K-20 Education Code.

Chetty, R., Friedman, J. F., & Rockoff, J. E. (2011). *Long-term impact of teachers: teacher valued and student outcomes in adulthood*. National Bureau of Economic Research, Working Paper #wl7699.

Chetty, R., Friedman, J. F., & Rockoff, J. E. (2013). *Measuring the impacts of teachers II: Teacher valued and student outcomes in adulthood.* National Bureau of Economic Research, Working Paper #wl9424.

Conklin, J. (2005). Book review of a taxonomy for learning, teaching, and assessing: A Revision of Bloom's taxonomy of educational objectives. *Educational Horizons 83*(3) 154-159.

Coyle, D. (2009) *The Talent Code*. New York: Random House, Inc.

CPALMS. Cognitive Complexity/Depth of Knowledge Rating. Retrieved from http://www.cpalms.org/textonly.aspx?ContentID=23&UrlPath=/page23.aspx

Cruickshank, D. R., & Haefele, D. L. (1990). Research-based indicators: Is the glass half full or half empty? *Journal of Personnel Evaluation in Education, 4*(1), 33-39.

Darling-Hammond, L. (1986). A proposal for evaluation in the teaching profession. *The Elementary School Journal, 86*(4), 530-551.

Darling-Hammond, L, Amrein-Beardsley, A., Haertel, E., & Rothstein,J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6). 8-13.

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

DeNisi, A. S., & Kluger, A. N. (2000). Feedback effectiveness: Can360-degree appraisals be improved? *Academy of Management Executive, 14*(1).

Doran, H. C. (2003.) Adding value to accountability. *Educational Leadership, 61*(3), 55-59.

Doran, H. C., & Izumi, L. T. (2004). *Putting evaluation to the test: A value-added model for California.* San Francisco, CA: Pacific Research Institute.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15-18, 20-24.

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

Fedor, D. B., & Buckley, M. R. (1987). Providing feedback to organizational members: A reconsideration. *Journal of Business Psychology, 2*(2), 171-181.

Fernandez-Toro, M., Truman, M., & Walker, M. (2013). Are the principles of effective feedback transferable across disciplines? A comparative study of written assignment feedback in language and technology. *Assessment and Evaluation in Higher Education, 38*(7), 816-830.

Florida Department of Education (2014). Educator quality update on teacher evaluation. Retrieved from http://www.fldoe.org/profdev/pdf/OverviewFloridasTeacherEvaluationSystem.pdf

Florida RTTT great teachers and leaders: Marzano evaluation framework glossary. Retrieved from http://www.columbia.k12.fl.us/personnel/Marzano%20Glossary.pdf

Florida State Board of Education Rule 6A-5.030, Instructional Personnel and School Administrator Evaluations (2011).

Florida State Statute § 1008.22(8).

Florida State Statute § Section 1012.34, Personnel Evaluation Procedures and Criteria, and State Board of Education Rule (2011).

Florida State Statute § Section 1012.34(3)(a)1.

Flowers, C. P., & Hancock, D. R. (2003). An interview protocol and scoring rubric for evaluating teacher performance. *Assessment in Education, 10*(2), 161-168.

Frase, L. E., & Streshly, W. (1994). Lack of accuracy, feedback and commitment in teacher evaluation. *Journal of Personnel Evaluation in Education*, *79*(4), 47-57.

Fusarelli, L. (2002). Tightly coupled policy in loosely coupled systems: institutional capacity and organizational change. *Journal of Educational Administration*, *40*(6), xx.

Gagne, J. (2011). *Measuring a teacher's value and effectiveness in SREB states*. (Policy Brief). Atlanta, GA: SREB.

Geddes, D., & Linnehan, F. (1996). Exploring the dimensionality of positive and negative performance feedback. *Communication Quarterly, 44*, 326-344.

Gladwell, M. (2008). *Outliers: The Story of Success*. 1st ed. New York: Little, Brown and Company.

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

Goldhaber, D. and Anthony, E. (2004). *Can teacher quality be effectively assessed?* Urban Institute. Washington, DC.

Goldhaber, D., & Theobald, R. (2009). Do Different Value-Added Models Tell Us the Same Things? Carnegie Knowledge Network, Last modified October 2012. Retrieved from http://carnegieknowledgenetwork.org/briefs/value-added/different-growth-models/.

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

Hanushek, E. A. (2011, Summer). Valuing teachers: How much is a good teacher worth? *Education Next, 11*(3), 40-45.

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

Hattie, J., & Timperiey, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Herold, D. M, & Greller, M. M, (1977). Feedback: The definition of a construct. *Academy of Management Journal, 20*(1), 142-147.

Hundt, N. E., Brown, L. H., Kimbrel, N. A., Walsh, M. A., Nelson-Gray, R., & Kwapil, T. R. (2012). Reinforcement sensitivity theory predicts positive and negative affect in daily life. *Personality and Individual Differences, 54*, 350-354.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organization. *Journal of Applied Psychology, 64*, 349-371.

Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, *26*(1), 101-136.

Jerald, C. (2012). *Ensuring accurate feedback from observations: Perspectives on Practice.* Seattle, WA: Bill and Melinda Gates Foundation.

Kane, T. J., & Staiger, D. O. (2008). *Estimate teacher impacts on student achievement: An experimental evaluation.* Working Paper 14601. Cambridge, MA: National Bureau of Economic Research.

Kane, T., & Staiger, D. (2012). *Gathering feedback for teaching: combining high-quality observations with student surveys and achievement gains* (Research Report). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Towards the understanding of a double-edge sword. *Current Direction in Psychological Science*, *7*, 67-72.

Kyriakides, L., Demetriou, D., & Charalmbous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research, 48*(1), 1-20.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes, 50*, 212-247.

Learning Sciences International. (2010). *Marzano teacher evaluation*. Retrieved from http://www.marzanoevaluation.com/#sthash.bnhH7bvq.dpuf

Little, O. (2009). *Teacher evaluation systems: The window for opportunity and reform.* Washington, DC: National Education Association.

London, M. (1997). *Job feedback: Giving, seeking and using feedback for performance improvement.* Mahwah, NJ: Erlbaum.

Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). *Investigating the links to improved student learning: Final report of research findings.* New York, NY: The Wallace Foundation.

Lunenburg, F., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in social and behavioral sciences.* Thousand Oaks, CA: Corwin Press.

Marshall, K. (2011). *Teacher evaluation rubric*. Retrieved from https://www.google.com/?gws_rd=ssl#q=Marshall%2C+K.+(2011%2C+August).+Teacher+Evaluation+Rubrics.+s.

Marshall, K. (2012). Fine Tuning Teacher Evaluation. *Educational Leadership*, *79*(7).

Marzano, R. J. (2007). *The art and science of teaching*. Alexandria, VA: ASCD

Marzano, R. J. (2011). *Research base and validation studies on the Marzano evaluation model.* Retrieved from www.MarzanoEvaluation.com

Marzano, R.J. (2013). *The Marzano teacher evaluation model.* Retrieved from http://tpep-wa.org/wp-content/uploads/Marzano_Teacher_Evaluation_Model.pdf

Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching.* Alexandria, VA: ASCD.

Marzano, R. J., & Toth, M. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD.

Marzano, R. J., & Toth, M. (2014, June). *Building expertise*. Powerpoint presentation at the meeting of the Marzano International Conference, Orlando, FL.

National Board for Professional Teaching Standards. Retrieved from http://www.nbpts.org

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

Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, *79*(4), 126-137.

Orange County Public Schools. (2013). *Frequently asked questions about the Florida Department of Education value added model (VAM).* Retrieved from https://www.ocps.net/cs/services/accountability/Documents/FAQ_About_VAM.pdf

Ovando, M. N. (1992). *Constructive feedback: A key to successful teaching and learning*. Retrieved from http://files.eric.ed.gov/fulltext/ED404291.pdf

Ovando, M. N. (2004). Building instructional leaders' capacity to deliver constructive feedback to teachers. *Journal of Personal Evaluation in Education 18*(3), 171-184.

Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal, 48*(l), 163-193. doi: 10.3102/0002831210362589.

Pate, L. E. (1977). Cognitive versus reinforcement views of intrinsic motivation. *Academy of Management Review, 6,* 75-88.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science 28,* 4-13.

Ravitch, D. (2010). *The death and life of the American school system.* New York, NY: Perseus Books.

Ready, D. D. (2013). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy 27*(1) 92-120.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2): 247 - 252.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on student achievement. *Sociology of Education, 70*, 256-284.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, *8*, 299-311.

Sanders, W. L., Ashton, J. J., & Wright, S. P. (2005). *Comparison of the effects of NBPTS certified teachers with other teachers on the rate of student academic progress*. Arlington, VA: National Board for Professional Teaching Standards. Retrieved from http://www.nbpts.org/UserFiles/File/SAS_final_NBPTS_report_D_-_Sanders.pdf

Sawyer, T., Hollis-Sawyer, L., & Pokryfke, A. (2002). Personality and social-evaluative anxieties associated with rating discomfort in anticipated positive and negative feedback conditions. *Social Behavior and Personality, 30*(4), 359-372.

Scherrer, J. (2011). Measuring teaching using value-added modeling: The imperfect panacea. *NAASP Bulletin, 95*(2), 122-140.

Scherrer, J. (2012). What's the value of VAM (value-added modeling)? *Kappan. 93*(8).

Senate Bill 736 (2011).

Shute, V. (2008, March). Focus on Formative Feedback. *Review of Educational Research*. 78(1), 153-189.

Stone, D. & Heen, S. (2014). *Thanks for the Feedback, The Science and Art of Receiving Feedback Well*. Penguin Group, New York, New York.

Stronge, J. H., & Hindman, J. L. (2003). Hiring the best teachers. *Educational Leadership, 60*(8), 48-52.

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

Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2008). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education 20*(3-4), 165-184.

The Student Success Act, H.R. 3989 (2010).

Taylor, E. S., & Tyler, J. H. (2011, March). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-year teachers*. NBER Working Paper Series. Cambridge, MA: Bureau of Economic Research.

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

Toth, M. (2014, June). *Feedback data*. Presentation at the Building Expertise Marzano International Conference, Orlando, FL.

Vandevoort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National board certified teachers and their students' achievement. *Education Policy Analysis Archives, 12*(46). Retrieved from http://epaa.asu.edu/epaa/v12n46/.

Wallace Perspective. (2013). *The school principal as leader: Guiding schools to better teaching and learning*. Retrieved from http://www.wallacefoundation.org/knowledge-center/school-leadership/effective-principal-leadership/Pages/The-School-Principal-as-Leader-Guiding-Schools-to-Better-Teaching-and-Learning.aspx

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* Brooklyn, NY: The New Teacher Project, 2009.

Wiggins, G. (2012). 7 Keys to effective feedback: Advice, evaluation, grades-none of these provide the descriptive information that students need to reach their goals. What is true feedback- and how can it improve learning? *Educational Leadership*, *70*(5), 11-16

Williams, C. (2014). *Effective Management* (6th ed.). South-Western Cengage Learning, Mason, OH.

Yeh, S. S. (2007). The cost-effectiveness of raising teacher quality. *Educational Research Review,* 220–232.

Yeh, S., & Ritter, J. (2009). The cost-effectiveness of replacing the bottom quartile of novice teachers through value-added teacher assessment. *Journal of Education Finance, 34*(4), 426-51.