Spring April 28th, 2016

Reflective Assessment, Feedback and Academic Achievement in High School Mathematics

Nalline S. Baliram
Seattle Pacific University

Reflective Assessment, Feedback and Academic Achievement in High School Mathematics

Nalline Baliram

Dissertation

Presented to the Faculty of the

Graduate School of Education at

Seattle Pacific University

In Partial Fulfillment of the Requirements for the

Doctor of Education Degree

Seattle Pacific University

2016

Reflective Assessment, Feedback and Academic Achievement in High School Mathematics

by

Nalline Baliram

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Education

Seattle Pacific University

2016

Approved by Arthur K. Ellis

Arthur Ellis, Ed.D., Professor, Chairperson

Rick Eigenbrood, Ph.D., Dean, School of Education, Committee Member

David Denton, Ed.D., Assistant Professor, Committee Member

John Bond

John Bond, Ed.D., Associate Professor, Committee Member

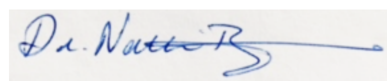
Program authorized to offer degree

SCHOOL OF EDUCATION

Date May 25, 2016

Rick Eigenbrood, Dean, School of Education

Copyright Page

Signature 

Date June 18, 2016

Table of Contents

List of Figures

List of Tables

List of Appendices

Dedication

This dissertation is dedicated to the friendship and memory of Ms. Diane DeMarco. She was my mentor and spiritual advisor who encouraged me to pursue my life-long dream of earning a Ph.D., which will allow me to serve in a greater capacity. I am grateful for the example Diane led as she trusted God through faith, hope and love. During the last year of her life, Diane gave me a new appreciation for the meaning and importance of friendship. She lived her life well, acting upon her spiritual beliefs diligently by assisting both friends and strangers in need. I will be forever grateful for her friendship.

Acknowledgements

The undertaking of this dissertation would not have been possible without the support of my family, friends, instructors, committee members and my advisor.

To my parents (Sam and Zorida), brother (Nigel) and sister (Nichell), I am grateful for your continued support and especially your prayers. To my nephews, Trey and Colin, thank you for trusting me to be a positive influence and example you can follow after. To the Singh family, *I choose family*.

To my many friends who traveled this journey by my side: Leanna Aker, Eric Boyer, Kimberly Jensen, Alex Johns, Heidi Monroe, Jennifer Norton, Jason Profit, and Jeffery Youde. Your choosing to meet with me, review my dissertation, listen to my rants, and believe in me is humbling. To Steven Zelenty and the students at Fr. Lopez High School, I am in deep gratitude for your support.

To my amazing instructors who have paved the way for my success, your expert knowledge and encouragement served as my motivation. Thank you for believing in my potential and helping me develop the confidence and expertise to accomplish this undertaking.

To a group of dedicated mentors who served on my dissertation committee, Dr. John Bond, Dr. David Denton and Dr. Rick Eigenbrood. You inspired me by your continued support and expertise in every component involved with my dissertation.

To my caring committee chair, Dr. Arthur Ellis, who guided me through the process. I am humbled by your brilliance, kindness and consistent encouragement.

Finally, because of God's unconditional love, he has lovingly provided me with the best support system a doctoral student could ask for! Thank you Lord!

Seattle Pacific University

Abstract

Reflective Assessment, Feedback and Academic Achievement in High School Mathematics

by

Nalline Baliram

Chairperson of Dissertation Committee: Arthur K. Ellis

School of Education

The purpose of this study was to determine the effects of metacognitive strategies and content-specific feedback on student achievement in high school mathematics. Participants in the study consisted of a convenience sample of honors geometry students in grades 9 and 10 in a private high school located in Daytona Beach, Florida. Beyond answering the specific research questions raised in the study, an additional aim was to contribute to the growing body of knowledge pertaining to effective ways to use metacognitive instruction and provide effective content-specific feedback to improve student achievement and learning.

A quasi-experimental, nonequivalent control-group design with repeated-measures was employed in the study. Descriptive and inferential statistics were computed to address the research questions. Specifically, an ANOVA with repeated measures, two-tailed test, was performed. For this purpose, a single within-subject factor, termed Assessment, was defined. Three levels were allocated to this factor, Pre-test, Post-test and Retention Test. Group was defined as a between-subjects factor and

the two levels allocated to this factor were Comparison and Experimental. Tests of statistical significance were analyzed at the .05 level.

There was a statistically significant main effect of the variable Group ($F(1, 73) = 7.27$, $p = .009$, $\eta_p^2 = .091$). Students in the experimental group outperformed the students in the comparison group. According to the effect size estimate, about 9% of variance in the Testing variable was attributable to the Group variable. Specifically, there was a statistically significant difference in the post-test ($p = .02$, Cohen's $d = .57$). This effect size calculated using Cohen's *d* formula is considered medium in magnitude (Cohen, 1988, 1992).

There was statistically significant time effect ($F(1, 73) = 1185$, $p = .000$, $\eta_p^2 = .942$). The retention test scores were lower than the post-test scores, however, students in the experimental group significantly outperformed the students in the comparison group in the retention test scores ($p = .00$, Cohen's $d = .69$).

The findings of this study offer a modest contribution to the body of empirical research on the impact of metacognitive practice and content-specific feedback on academic achievement at the high school level. Further studies are warranted to add to the body of literature and more specifically to provide great clarity regarding the magnitude of the current investigation.

Chapter One

Introduction

Purpose of the Study

The purpose of this study is to determine the effects of metacognitive strategies and teacher feedback on the academic achievement of high school students in mathematics. Specifically, this study examines the effects of situated metacognition, in the form of reflective assessment, and teacher feedback on high school students studying geometry. The reflective prompts utilized in this study are based on those articulated by Ellis (2001) and Mevarech and Kramarski (1997), which have been incorporated into prior studies (e.g., Bianchi, 2007; Bond, 2003; Evans, 2009; Kramarski & Mevarech, 2003).

Reflective practice, a type of formative assessment can be a diagnostic approach used to provide feedback to both the teachers and students over the course of instruction. As defined by Black and Wiliam (1998), assessment includes all activities that teachers and students undertake to get information that can be used to alter teaching and learning. By allowing the opportunity for students to practice reflection, teachers should be able to identify areas where they are struggling and further provide feedback in attempt to ameliorate the situation. Furthermore, Black and Wiliam (1998), pioneers of assessment, in their numerous research reviews concluded formative assessment raises academic standards in the classroom and produce significant learning gains as measured in test scores.

High quality studies involving feedback as a component of formative assessment have suggested students are able to regulate their own progress by recognizing where the gaps between their desired goal and current knowledge may lie and work toward obtaining the goal (Sadler, 1998). In a study conducted by Bangert-Drowns, Kulik, Kulik, and Morgan (1991), feedback provided on tests and homework were helpful to lower achieving students because comments focused on errors made along with specific suggestions for improvement. Students felt encouraged to focus their attention thoughtfully on the task rather than simply getting the right answer. Research suggests formative assessment such as reflective practice and feedback are tightly linked with instructional practices. Therefore, teachers must consider how their classroom activities, assignments and tests support student learning and allow students to freely communicate what they know, what they can do and areas in which they continue to struggle. Teachers must then use this information to improve teaching and learning.

This study presents a careful and critical analysis of previous work and theory along with the practical aim of providing insights and rationale to educators supporting the use of metacognitive strategies, such as reflective assessment, that accompanies feedback in their lessons. A central goal of this study is to further advance the growing body of knowledge regarding effective ways to use metacognitive instruction and provide effective feedback to improve student achievement and learning.

Background

We are in an era of high-stakes testing and heightened pressure to improve student achievement. Teachers are increasingly expected to help their students produce favorable outcomes on high-stakes standardized tests (Guth et al., 1999). Therefore, they continue to examine methods, concepts, and strategies that will help their students acquire, make sense of, and retain knowledge. Donald Schön (1987), whose work has been influential in developing the theory and practice of reflective thinking, argued that teachers' work is complex and often requires profound reflective practices to achieve positive outcomes. Similarly, students' reflective thinking, a crucial component of meta-cognitive practice, should be considered vital in achievement and learning. Noted by Ellis (2001), reflective assessment is for everyone and that includes students and educators.

The term 'metacognition' was coined by Flavell (1979); however, reflection, the term associated with metacognition predates Flavell. Reflective practice dates back to ancient Eastern and Western philosophies and religions (Marzano, Boogren, Heflebower, Kanold-McIntyre, & Pickering, 2012). For example, Socrates emphasized to his students the value of examination of self. Also, Buddhists have traditionally used reflection to individually search for insight and truth (Marzano et al., 2012).

An extension of reflective assessment is feedback, one of the most powerful influences on learning and achievement (Hattie & Timperley, 2007). Feedback can be perceived to be positive or negative, therefore, the type of feedback and the way it is given can be differentially effective. Both reflective thinking and teacher feedback can be characterized as highly esteemed and widely used techniques that are utilized in a variety of professions to aide in adapting and making decisions. Extensive literature on opinions and philosophy with respect to the value of these two approaches continues to emerge (Bandura, 1997; Dewey, 1910; Flavell, 1977; Hattie & Timperley, 2007; Vygotsky, 1978). An examination of the history and different perceptions of this construct are presented in Chapter Two.

Significance of the Study

This study bridges the gap of information with respect to metacognitive practice and teacher feedback. There appear to be a limited number of the empirical research studies in the area of reflective assessment and teacher feedback, specifically in secondary school mathematics, that describes a diverse population. One goal in this study is to focus on a diverse population that will further generalize the impact and applicability of metacognitive practices and teacher feedback. Second, a number of studies that examine the impact of reflective assessment or metacognition with the element of teacher feedback on student learning has done so in collegiate-age and elementary to middle school aged participants. This study focuses on high school achievement, specifically high school students who are taking geometry, a required course for graduation.

The body of knowledge regarding metacognition or reflective assessment continues to grow. Additionally, feedback plays a vital role in student progress toward learning. There are few studies that link student reflection with teacher feedback and examine the effects of both practices on academic achievement. Finally, this study seeks to shift the priority of instructional delivery by validating the need for reflection and feedback. The more evidence that teachers receive about the benefits of these two approaches, more likely that they will enhance their learning environment by integrating the approaches in their daily practice. Research conducted in a realistic classroom environment will provide teachers with helpful information that will inform their instructional practice.

In order to build on the existing body of research, this study uses a quasi-experimental design to examine the impact of metacognitive practice or reflective assessment as defined by Ellis (2001), utilized by Evans (2009) in high school English Literature, and Mevarech and Kramarski (1997), utilized by Kramarski and Mevarech (2003) on mathematics of eighth grade mathematics students.

Research Questions

This study examines the use of metacognitive practice (reflective assessment) and teacher feedback during geometry instruction at the high school level. The null and experimental hypotheses derive from the research questions presented:

Research Question 1: Is there a statistically significant difference on achievement of high school geometry students who practice metacognition or reflective assessment and receive teacher feedback, when compared to those who are provided with the same instruction but do not explicitly practice reflective techniques nor explicitly receive teacher feedback?

H_0 = There is a statistically non-significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

H_1 = There is a statistically significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

Research Question 2: Does the use of metacognitive strategies enhance student retention of Geometry concepts over time?

H_0 = There is a statistically non-significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

H_1 = There is a statistically significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

Structure of Dissertation

The body of this dissertation is organized into four subsequent chapters titled Literature Review, Research Methods, Results, and Discussion of Results.

Chapter Two defines metacognition in terms of reflective assessment and teacher feedback as it provides a thorough examination of the theoretical construct of metacognition and teacher feedback. A summary of quantitative and qualitative research related to both metacognition referred to as reflective assessment and teacher feedback are analyzed and critiqued. This summary also touches upon the lack of research that examines the impact of both metacognition and teacher feedback on academic achievement.

Chapter Three provides a description of the methodological approach employed in this study. The research hypotheses are presented and the specific research design, including participant selection and assignment, validity and reliability of the instrument utilized, and procedural elements are discussed. Additionally, the specific data analysis and statistical methods utilized in this study are thoroughly analyzed.

Chapter Four provides a detailed summary of the results for the study. The descriptive and inferential statistics related to the research questions are summarized in both narrative and table format. A review of the assumptions underlying the statistical procedures are provided. Major findings along with trends in the data are identified and further discussed in the final chapter.

Chapter Five contains a discussion of both the statistical and practical significance of the findings of this study. A comparison to findings reported in prior empirical studies are presented. Additionally, the limitations, the threats to internal and external validity will be discussed. The chapter closes with suggestions for improvement to the study and recommendations for future studies that examine the impact of metacognition and teacher feedback on academic achievement.

Chapter Two

Review of Literature

Introduction

Assessment, a component of the tri-part model of instruction in education, plays an integral role in classroom life (Pellegrino, 2010). It helps govern whether or not goals in education are met. It affects decisions about grades, placement, curriculum, and in some cases, funding. Assessment can answer such questions as, are teachers teaching what they should be teaching? Or, are students learning what they should be learning? Assessment can also address the question, how can educators become better teachers and students become better learners? It is argued that assessment ultimately leads to student achievement (Borich, 2014; Costa, 2001; Ellis, 2001). Keeping in mind, whatever form assessment takes, issues of validity, reliability, and authenticity remain.

Much of the current consensus on how schools can use assessment to inform academic achievement of students and promote positive social and emotional development is through a learner-centered environment (McCombs, 2010). One needed element, often absent, that constitutes a learner-centered approach is metacognitive practice (McCombs, 2010). And accompanying this approach, is another type of formative assessment called feedback. It is suggested that teachers use this strategy with the intention to improve teaching and learning (Black & Wiliam, 1998).

What is the potential impact of metacognitive practice and teacher feedback on student achievement in secondary schools? In recent years, there has been an exponential increase in the number of journal articles as well as books that discuss reflective assessment and teacher feedback as two isolated approaches (Hattie, 2012). While the evidence of the effects of reflective assessment and feedback are substantial, further studies that link metacognition with teacher feedback are warranted. In this chapter, definitions, theoretical underpinnings, and empirical research investigating the use of reflective assessment and teacher feedback are presented.

Definitions

Reflective assessment falls within the paradigm of metacognition drawing its theoretical origins from both cognitive psychology and constructivist theory. Metacognition, which essentially means thinking about thinking (Flavell, 1977), and reflective assessment, which is an applied form of practice based on metacognitive theory, have in common the idea that opportunity for growth is enhanced when students are given time not only to learn, but also to thoughtfully consider what they are learning. This literature review explores elements of constructivism and cognitive psychology related to both terms.

Metacognition. The term metacognition appeared as an interesting and promising new area of study based on psychologist John Flavell's work several decades ago (Flavell, 1979). He pointed out that, "ideas about metacognition are beginning to make contact with similar ideas in the areas of social learning theory, cognitive behavior modification, personality development, and education" (p. 906). In education, metacognition refers to the way teachers and students plan, monitor and assess understanding and performance. According to Bandura (1997), metacognition involves thinking about one's cognitive activities, and this skill allows the individual to organize, monitor, evaluate, and regulate the thinking process. It includes thinkers being aware of how they think and learners being aware of how they learn. Ultimately, as literature suggests, metacognition is defined as "thinking about thinking" (Costa, 2001; Flavell, 1979; Schoenfeld, 1987). In this sense, student reflection represents a value-added component often missing in teaching and learning.

Reflective assessment. John Dewey (1910) wrote that reflection has the potential to happen when there is a feeling of doubt or perplexity. In his book, *How We Think* (1910), Dewey defined reflection as involving a consecutive order so that each idea determines the next outcome. He further proposed that, “the successive portions of reflective thought grow out of one another and support one another” (p. 3). He defined the term *thought* when each phase is a step from one form of thinking to another. It streams or flows, and becomes a train, chain, or thread of reflective thought. Successive portions of reflection start from uncertainty of an idea, then lead to inquiry that corroborates or nullifies the belief.

As defined by Leung and Kember (2003), *reflection* is described as an attempt to understand an issue or question within a personal context or going beyond learning to assimilate information to make meaning (p. 64). *Assessment* is a valid measure of learning that provides feedback to both the teacher and the learner for the purpose of improving teaching and learning (Popham, 2014; Stiggins, 1996). Therefore, *reflective assessment* implies active contemplation on the cognitive process of knowledge, skills, situations or experiences with some kind of measurement, typically formative. In this sense, reflective assessment by students and teachers is assessment for purposes of learning and growth. Thus a distinction exists between summative assessment and formative assessment of learning.

Feedback. Formative assessment refers to assessment that is specifically intended to generate feedback on a student’s performance with the intent to improve learning (Black & Wiliam, 1998; Sadler, 1998). John Hattie (2012) theorized that the most powerful strategy that enhances achievement is feedback. Naturally, the effects of feedback depend on the nature of the feedback. It can provide specific information through written conversation or conversations about the learning that is happening in the classroom. Defined by Wiggins (1998), it is information about how a person did in light of what he or she attempted.

Feedback is a crucial form of formative assessment that should be used to help learners understand what they need to do to improve their learning as well as what was done well (Brookhart, 2008). Additionally, it should provide students with sufficient information so they know what to do next. In other words, it goes beyond, positive reinforcements such as writing ‘good job’ or stamping happy faces. Irons (2008) defined feedback as “any information, process or activity which affords or accelerates student learning based on comments relating to either formative or summative assessment activities” (p. 7). According to Brookhart (2008), effective feedback should be clear, age-appropriate, content specific, timely, and of high quality. Typically, it comes from teacher to student; however, effective feedback can also come from student to student as well as student to teacher.

Theoretical Underpinnings

Metacognition and reflective assessment. Reflective thinking became a vital theme during the progressive movement in American education. Dewey (1910) considered reflection an “active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends” (p. 6). He further theorized reflection as a process that enables the learner to move from one experience to the next and which involves a deeper understanding of its relationships with prior experiences and ideas. It is a rigorous and systematic way of thinking with its roots in scientific inquiry. Dewey (1910) additionally outlined steps to reflective thinking which established the foundation for the connection between reflection and learning in modern education.

Flavell (1979), who defined metacognition as ‘thinking about thinking’, acknowledged the significance of metacognition in a wide range of applications which included reading, oral skills, writing, language acquisition, memory, attention, social interactions, self-instruction, personality development and education. Components of metacognition can be activated intentionally; this could be through a memory search with the purpose of retrieving specific information (Flavell, 1979). Such components can help the individual make meaning and discover behavioral implications of metacognitive experiences.

Socrates complained that teachers spend far too much time telling and too little time allowing students to think about what they are learning (Plato, 1952). The concept of “good teaching” is achieved by providing students with opportunities to learn and practice the art of inquiry, deep order learning and reflective learning (Ramsden, 2003). As noted by Jerome Bruner (1961), students should spend more time studying problems in depth and less time covering a wide range of topics; meaning students be allowed opportunities to consistently practice formative self-assessment as a means of clarifying their thinking about what they are learning.

Cognitive psychology, in contrast to behaviorism, focuses on how the mind works with processing, representing, organizing, and retrieving information (Bandura, 1997);

therefore, metacognitive practice offers a natural link between educational practice and psychology. Many theorists have explored the role of reflection in the learning environment. Piaget (1976), an advocate of learning through discovery, wrote that principles of cognitive psychology involve learning as an ongoing process in which learners are continually assimilating and accommodating new information to that which they already know. This process involves the concept of reflection and is essential for integration and assimilation of new information. Bandura (1997), an advocate of the significance of self-efficacy in learning, proposed that effective intellectual functioning requires metacognitive skill to organize, monitor, evaluate, and regulate the thinking process. He wrote that "Metacognition involves thoughts about one's cognitive activities rather than simply higher order cognitive skills" (p. 223).

Based on these theoretical perspectives, it is clear that the idea of teaching students to think about their own thinking has been in existence for a long time, eventually taking its place as a vital theme of the progressive movement. Reflective thinking relates to constructivism in that learners are conscious of how they learn and can therefore regulate their progress (Joyce, Weil, & Calhoun, 2009). This suggests when students are taught science for example, teaching the scientific thinking processes is not enough. A value-added component occurs when students can demonstrate to themselves their academic achievement, when they are able to reflect on their own learning and the learning process (Marzano et al., 2012). The theoretical works of Dewey (1910), Bruner (1961), Bandura (1997), and Piaget (1976) continue to inform education today specifically in the area of metacognitive practice.

Feedback. In the 1960s, psychologists argued that schools could improve instruction by adopting a more systematic approach and therefore, borrowed the idea of feedback from engineering systems theory. However, through intervention, they discovered feedback must be designed to be a part of a system instead of just telling students whether their responses were correct or incorrect (Wiliam, 2012). Therefore, when feedback is given, it should inform the students that the current performance falls short of the learning goal or the goal has already been reached.

Hattie and Timperley (2007) in their article, *The Power of Feedback*, hypothesized feedback as one of the most powerful influences on learning and achievement. Feedback can be perceived to be positive or negative, therefore, the type of feedback and the way it is given can be differentially effective. The authors described feedback as a “consequence” of performance (p. 81) suggesting successful outcomes can occur when student are able to make meaning of the feedback they receive. Irons (2008) implied quality feedback can be utilized as a constructive leaning tool to improve teaching and learning. He inferred that it can empower students to become self-regulated learners with the ability to develop self-efficacy and confidence.

According to Irons (2008), feedback can benefit students only if it indicates clear expectations that can be attainable, compares their current level of performance with the intended level, and provides specific actions students must follow in order to improve their learning or to close the gap between their current to intended levels of performance. In other words, feedback should explicitly describe what students learned and did not learn based on what was communicated through either written work or oral discussions. Therefore, it should be conveyed in a way that enables students to become actively engage with the aim to improve their learning and understanding of the concepts, knowledge, and skills.

Hattie and Gan (2011) explained that feedback potentially serves different functions depending on how it is perceived and the underlying assumptions about the learning context on which research in these areas are based (p. 250). The four philosophical perspectives of learning and the nature of feedback proposed by Hattie and Gan (2011) include objectivism, information processing, sociocultural, and visible learning theory. The framework for this study will incorporate all four theoretical perspectives that are linked to reflective assessment.

Objectivism. When feedback is viewed as a process for reinforcing knowledge in a sequential and hierarchical fashion, it falls under the construct of objectivism (Hattie & Gan, 2011). In other words, feedback comes from an external source, for example, the teacher who is identified as the expert. The teacher uses the feedback mechanism to strengthen knowledge and understanding. Kulhavy and Wager (1993) argued that feedback in the form of reinforcement while benefits novice learners, its effects can be limited and confusing in that feedback that focuses on incentives can distract the learner from the content of the feedback. Anderson, Kulhavy, and Andres (1972) confirmed this argument when they found that students tend to bypass the feedback when they are aware answers will be readily available for them.

Additionally, Deci, Koestner, and Ryan (1999), in their meta-analysis review, found that when teachers provide tangible reward in a form of feedback, intrinsic motivation is significantly undermined and students are less inclined to take responsibility for motivating or regulating themselves (p. 639). Keeping this in mind, feedback should be provided as a consequence of performance rather than prior to completion of any task. Kulhavy and Wager (1993) further suggested to isolate motivational variables from the feedback so students can focus on the instructional content of feedback.

Informational processing. Feedback helps learners when it comes to processing information (Hattie & Gan, 2011). Students' prior knowledge, mental structures and beliefs can be linked through feedback in that each learner constructs his own truth through process and interpretation of their experiences. A feature of information processing as noted by Hattie and Gan (2011), is that students' cognitive ability to use information can be activated when they are engaged with the learning task. This means feedback functions not only reinforce correct answers but serves as a tool to help learners to correct their own errors. This approach provides feedback messages in two ways: through verification and elaboration. Verification indicates that the response is either right or wrong and elaboration contains relevant information that guides the students to recognizing their error and correcting their mistake.

Feedback that is elaborative would include restating the correct answer or adding multiple choice responses as alternatives to lead the students to the correct answer. This strategy is identified as task-specific. An instruction-based approach provides explanations of why a certain response is correct. The information can also be presented again in a manner this time that contains the correct answer. Further, extra-instructional elaboration refers to additional examples or analogies used to help the students with the knowledge or content (Hattie & Gan, 2011).

Sociocultural constructivism. The socio-cultural perspective is derived from the works of Vygotsky's theory of social interaction. The purpose of feedback in this case is designed to interact through meaningful use of language (Hattie & Gan, 2011). Knowledge and understanding constructed are shared through social interaction rather than individual experience. According to Vygotsky (1978), learning happens during social interaction and linguistic practices; and the interaction between the learner and the teacher becomes internalized as the basis for reflection and logical reasoning (Hattie & Gan, 2011).

Confirmed by Mercer and Littleton (2007), this approach is viewed as a mediation through a dialectical relationship between interpersonal and intrapersonal process. As opposed to objectivism which affirms knowledge exists, in a socio-cultural environment, knowledge is constructed by learners through experience and actively participating in meaningful dialogues.

Visible learning theory. The visible learning theory is an advancement on the three models discussed earlier. It constitutes as feedback viewed at different levels (Hattie & Timperley, 2007). Feedback moves from a predominantly transcribed process to a dialogic and elaborative process in a social environment (Hattie & Gan, 2011). This suggests meaningful feedback can be conveyed with peers, with adults, or alone at varying stages of proficiency and understanding through different levels of regulation. What might appear to standout with this approach is the effects seen when the teacher receives feedback and adapts instructional strategies in order to improve learning (Hattie & Gan, 2011).

Identified in the visible learning theory, feedback is most powerful when it makes learning visible to the teacher. In fact, this could lead to the teacher creating an environment and activities with the intent to optimize student learning and make it visible to the teacher. For teachers, this would mean investing the time to make learning transparent to the learner and promote a successful outcome through feedback. It is worthwhile to note that feedback can make learning visible, lead to error detection, and enhance self-regulation about learning (Hattie & Gan, 2011). Simply stated by Wiliam (2012), "just as a thermostat adjusts room temperature, effective feedback helps maintain a supportive environment for learning" (p. 31).

Summary of theoretical underpinnings. The theoretical background demonstrates that reflective assessment or metacognitive practice as well as feedback are significant components of formative assessment that have serious implications for teaching and learning. Students should not just learn; they must be encouraged to reflect on how they learn and implications for generating success (Marzano et al., 2012). Furthermore, teachers must be able to recognize valuable insights in their students' metacognitive practice and provide them with meaningful feedback that will enhance their learning (Hattie & Timperley, 2007). Establishing a teacher-student relationship, rooted in these theoretical underpinnings, can foster a classroom climate in which feedback and reflective assessment are prevalent and highly valued.

Both types of formative assessments offer promise of success when they are focused on students' performance and their ability to accomplish their academic goals. A number of studies have investigated strategies that allow students to reflect on their thinking and learning as well as stimulate the metacognitive process to occur.

Additionally, a number of studies have examined the impact of teacher feedback on student learning at a variety of ability levels (Butler & Nisan, 1986; Nunez et al., 2013; Siewert, 2011).

Empirical Research

Effects of metacognitive practice on achievement. Recent research has been conducted to explore evidence of effectiveness of educational interventions to close the gap between student learning and achievement (Hamre & Pianta, 2010). One such intervention, reflective assessment, is designed to help students determine for themselves what they are learning and what they are not learning. It can also assist teachers to consider instructional methods that will better accommodate learning needs and therefore improve achievement (Hamre & Pianta, 2010).

Keeping in mind Albert Bandura's (1977) argument of the powerful effects of modelling behaviors, teachers who want students to practice reflective thinking are encouraged first to model it and demonstrate the value of its worth (Ellis, Denton, & Bond, 2013). Borich (2014) called this mental modeling, comprised of three steps: showing students how to reason, making students aware of their own reasoning, and helping students apply their reasoning. The key question is: How do teachers know what students are learning in any given lesson? By the end of a particular lesson, students should be able to explain what they learned, identify parts of the lesson they found most interesting, expound on the value of learning the specific content, and reflect on the most memorable part of the lesson. However, in order for them to do this, they must be provided with an opportunity to do so. According to the reflective assessment argument, it is vital for students to utilize metacognitive skills to reflect on their learning and comfortably share thoughts, questions and concerns with the instructor and with one another.

Empirical studies – reflective assessment. Cognitive theorists continue to be intrigued by metacognition as evidenced by the wealth of literature that promotes the vital role it plays in the learning environment. Bond and Ellis (2013) focused on fifth and sixth grade students and their ability to reflect meaningfully on concepts and skills in mathematics. The purpose was to examine the effects of metacognitive practice in the form of reflective self-assessment on the mathematics achievement of fifth- and sixth-grade students (p. 228).

The experimental posttest-only control group design consisted of 141 students who were randomly assigned to three groups (reflective assessment group, non-reflective review and control group) with each condition represented by subgroups or classes (Bond & Ellis, 2013). Each of the six teachers was randomly assigned to one of the subgroups. The random assignment of students strengthened the internal validity (Campbell & Stanley, 1963; Corrigan & Salzer, 2003) of the study and therefore, the investigators argued against the need for a pre-test as a covariate. Both experimental groups (reflective and non-reflective review) received identical instruction on statistics, a topic of a mathematics unit. At the closing of each class session, the reflective group practiced the reflective intervention. The students spent time completing "*I Learned*" statements and verbal "*Thinking Aloud*" protocols (Ellis, 2001). The non-reflective group spent the remaining five minutes reviewing the lesson activities and objectives. The control group, however, focused on geometry mainly in the form of area and perimeter lessons.

The results of the study supported the notion that student reflection enhances academic achievement. A one-way analysis of variance (ANOVA) was conducted to determine the effects of the reflective intervention on achievement mathematics test scores. The results indicated a statistically significant main effect ($p < .05$) and effect size of .273. This suggested about 27% of the variance in achievement was accounted for indicating a relationship exists between the reflective group experience and their resulting posttest scores. Students who were in the treatment group (reflective assessment) scored higher in their posttest (M = 29.40, SD = 4.33) than both the control group (M = 22.30, SD = 4.37) and the non-reflective group (M = 26.92, SD = 5.61). Additionally, a retention test was administered six weeks after the study to both experimental groups. Although the reflective group scored higher than the non-reflective group, there was no statistically significant difference between their post-test and retention test scores. The results of this study offer tentative support for reflective assessment strategies as embedded formative assessments in daily activities.

It would be challenging to generalize the results to a diverse population since the study took place in a suburban area. Further research is warranted to demonstrate effectiveness of reflective strategies in varied populations, for example, high poverty schools, schools in an urban area, places of cultural or ethnic diversity as well as at-risk populations that include special education and English Language Learners (ELL). Furthermore, the authors noted that the study was conducted at the time when a new curriculum was piloted. The post-test, however, though developed by the researchers and aligned with the piloted math curriculum, was found to be adequate and reliable with a Cronbach's alpha coefficient of .72 (Bond & Ellis, 2013).

Zan's (2000) Italy-based research focused on ways to improve the performance of university first year biology students who repeatedly failed the required mathematics examination. Twenty-seven such biology students were enrolled in an intervention course that lasted for four hours a week over six weeks starting October and ending in January. The intervention consisted of metacognitive strategies that would assist the students in passing the examination. A series of practice tests was administered to the students during the intervention. Prior to the tests, students were provided with self-reflection prompts to which they were to respond in writing. The students discussed how they prepared for the exam as well as their level of self-efficacy regarding the exam.

After the practice tests, students further reflected on whether or not their method of preparation for the tests worked (Zan, 2000). In addition, they outlined a plan for future preparations and discussed it with their teacher. At the end of the intervention, the researcher observed significant changes in the students' metacognitive behavior and attitudes. Specifically, she saw they were able to make connections between various topics; study in a critical way; identify their own doubts; and activate control strategies in their written tests. Furthermore, the students appeared to be more interested in the subject and felt more confident. While such anecdotal conclusions by the instructor are helpful, it should be noted that no formal pre- and post-measures of reflective growth were administered to the students and certainly, the absence of a control or comparison group weakens any inferential conclusions. In spite of the fact that this was not a cause and effect study, the results from the compulsory mathematics assessment appear to support the effectiveness of the intervention. All 27 students passed the assessment and 10 of them obtained scores of more than 25 out of 30 points.

Kramarski and Mevarech (2003) investigated the differential effects of four instructional strategies on students' mathematical reasoning in graph interpretation and transfer ability in graph construction, and metacognitive knowledge. The study included 384 students from 12 eighth-grade classrooms. The classes were randomly selected from four junior high schools with three classes from each school. The four schools were randomly chosen from a district of 15 junior high schools. As described by the Israel Ministry of Education, all four schools were similar in size and were of "average" socioeconomic status. The four instructional strategies in the study were cooperative learning with metacognitive training (COOP+META), individual learning with metacognitive training (IND+META), cooperative learning (COOP) and individual learning (IND). Each school was randomly assigned to one of the four conditions since it is established the teachers in the same school share materials and talk with each other about their teaching strategies.

A month after the start of the school year, all students from the 12 randomly chosen classes were administered three pretests: Graph Interpretation Test, Graph Construction Test, and Metacognition Questionnaire (Kramarski & Mevarech, 2003). The purpose of this was two-fold. First, the pretest was to ensure the heterogeneous composition of each cooperative group which would include one high-achieving student, two middle-achieving students, and one low-achieving student. Second, the pretest scores were used as a covariate to control for pretreatment differences.

Prior to the study, all 12 teachers who happened to be female underwent a two-day in-service training which focused on pedagogical issues related to the unit in the study (Kramarski & Mevarech, 2003). All 12 teachers were instructed that they would be teaching the linear graph unit using the same mathematical problems as examples. A set of learning materials that included metacognitive questions designed by the IMPROVE program (Mevarech & Kramarski, 1997) were provided to the teachers assigned to teaching COOP+META and IND+META strategies. The remaining six teachers from the other two schools were provided with general instructions. Additionally, the teachers in each instructional strategy were educated separately on the theoretical background of their learning methods and its practical implications (Kramarski & Mevarech, 2003).

Instruction for the groups was composed of three parts: introduction of content, cooperative or individualized seat work, and review with the whole class (Kramarski & Mevarech, 2003). For each lesson, the introduction made by the teachers was about 10 minutes long while the cooperative or individual work was 30 minutes long and the review was about five minutes long. The COOP+META classes were provided with metacognitive questions which included comprehensive questions, strategic questions and connection questions (Mevarech & Kramarski, 1997) to solve a problem or complete a task individually, in small groups and through class discussions. Additionally, the questions were used by the teacher as she introduced the concepts, reviewed the concepts, and provided additional support. For the IND+META classes, the metacognitive training was identical except it was implemented individually instead of in a collaborative setting.

At the end of the study, the same battery of tests was administered to all of the students in the 12 classes (Kramarski & Mevarech, 2003). Since there was a significant correlation between graph interpretation and graph construction scores (r =.48), a multivariate analysis of co-variance (MANCOVA) was conducted on the post-test scores controlling for the pre-test scores. The results indicated a statistically significant difference in posttest scores of graph interpretation and graph construction ($F(6, 744) = 6.17$, $p < .001$). Given the findings from the MANCOVA, a one-way analysis of covariance (ANCOVA) was conducted for graph interpretations, and statistically significant differences between the treatment groups were found in posttest scores ($F(3, 371) = 3.98$, $p < .05$). Post hoc analysis conducted based on pairwise comparison *t* tests suggested COOP+META students significantly outperformed the IND+META group. Furthermore, the IND+META significantly outperformed the COOP and IND groups. There were no significant achievement differences noted between the COOP and IND groups.

A second one-way ANCOVA for graph constructions was conducted and likewise, significant differences between the treatment groups were found in posttest scores ($F(3, 371) = 7.19$, $p < .05$) (Kramarski & Mevarech, 2003). Post hoc analysis conducted based on pairwise comparison *t* tests suggested statistically significant differences between the groups exposed to metacognitive training (COOP+META and IND+META) and the non-metacognitive groups (COOP and IND). Both metacognitive groups outperformed the two non-metacognitive groups. However, there was no significant difference found between the two metacognitive groups as well as between the two non-metacognitive groups. Reporting effect size for the parametric methods used in this study would further help to explain the amount of variance that was accounted for in the outcome.

Finally, to analyze metacognitive knowledge, the third assessment, a one-way MANCOVA was conducted with general strategy and specific strategy criteria as the two dependent variables (Kramarski & Mevarech, 2003). Significant differences were found between conditions on both general and domain-specific metacognitive ($F(6, 744) = 2.97$, $p < .01$). Post hoc analysis conducted based on pairwise comparison *t* tests suggested statistically significant differences between the groups exposed to metacognitive training (COOP+META and IND+META) and the non-metacognitive groups (COOP and IND). Both metacognitive groups outperformed the two non-metacognitive groups. However, there was no significant difference found between the two metacognitive groups as well as between the two non-metacognitive groups (Kramarski & Mevarech, 2003).

The three measurements used in the study were assessed for reliability. For the graph interpretation test, the Kuder Richardson reliability coefficient was .91, for graph construction test, the interjudge reliability coefficient was .92, and for the metacognitive questionnaire, the Cronbach alpha coefficient was .86. Though the researchers utilized sound measurements of high reliability coefficient, the outcome of the study should be tempered for a number of reasons. The study focused only on one instructional unit: linear graphs (Kramarski & Mevarech, 2003). The unit lasted only two weeks, a short time span. This suggests a need for more extended investigations and studies that can focus on additional instructional units. Furthermore, the findings from the study suggest junior high school students can productively think and reflect on their learning of mathematical concepts. This calls for additional studies to include school age students at both the primary and secondary level with focus on other disciplines including language arts, science, and social studies. Finally, although the authors disclosed the sample was made up of randomly chosen classrooms from four junior high schools of similar characteristics defined by the Israel Ministry of Education, replication studies are needed to further pursue these promising results.

A study conducted by Evans (2009) focused on the effectiveness of reflective assessment in the daily in-class learning for high school students in English literature (p. 37). Evans utilized a quasi-experimental posttest only control group design made up of a convenience sample of 235 ninth grade students. The sample was comprised of nine intact classrooms that were randomly assigned to one of the three treatment (control, comparison, and experimental) groups. The three classes identified as the control group were randomly assigned to one teacher. The experimental and comparison groups were split between two other teachers.

The treatment assigned to the experimental group was a scripted reflective assessment activity that was provided at the beginning and the end of daily class lessons (Evans, 2009). All three groups received the same curriculum aligned with the school district English course adoption; however, the control group was studied during second semester while the experimental and comparison groups were studied in the first semester.

The treatment applied to the experimental group was on a daily basis over the course of 22 lessons (Evans, 2009). It consisted of two strategies that were used in tandem with one another. *I Learned Statements,* and a personal statement written by the students about their learning (Ellis, 2001) were assigned as a closure activity of each lesson. Teachers collected and read them daily. They provided oral or written feedback to the students. *Variation of Prompted Think Aloud*, an extension and variation of the "*Talk About It*" strategy outlined by Ellis (2001), took place at the beginning of each day's lesson. This activity was based on the *I Learned Statements* from the previous lesson. Students verbalized their thinking about what they have learned through discussion prompts initiated by the teacher.

The comparison group which received the same lesson during the same time as the experimental group received vocabulary training instruction at the beginning and ending of each lesson while the experimental group received the treatment (Evans, 2009). The control group that received the traditional instruction in the second semester did not focus on the same unit of study as the experimental and comparison groups did. They neither practiced reflective strategies nor vocabulary training exercises (Evans, 2009). At the conclusion of the study, the control group was administered the same posttest as the other two groups.

Since there was only one independent variable and one dependent variable, an ANOVA was conducted (Evans, 2009). The results from the ANOVA were statistically significant ($F(2, 221) = 407.82$, $p = .000$, $\eta^2 = .779$), favoring the reflective assessment intervention group. This implies about 78% of the variance was accounted for indicating a very strong relationship between the treatment group and the posttest scores. Students in the experimental group had a significantly higher mean score ($M = 37.36$ $SD = 5.43$) on the posttest than did the students in the comparison group ($M = 32.75$, $SD = 7.73$). Furthermore, there was a significant difference between the experimental group and the control group. Students in the control groups had a significantly lower mean score ($M = 10.27$, $SD = 3.47$) than both the experimental group and the comparison group.

A common supporting instrument integrated in the classroom that has potential to promote student reflection includes reflective journals. This tool can enable students to think critically on their own learning as well as to understand their individual learning styles (Cisero, 2006). These journals can potentially reflect students' understanding and show how interactions are made between students and teachers. Such components will help both the students and teachers gain insights into the learning that is happening (Cisero, 2006). Lew and Schmidt (2011), evaluated whether reflective journal writing was effective in promoting self-reflection and learning (p. 532). They further investigated whether students become better at self-reflection if they engage continuously in reflective journal writing. The researchers hypothesized that students who reflect on how and what they learn will perform better in the classroom and acquisition tests.

The participants in the study were 690 first year applied science students of a three-year program at a polytechnic school in Singapore (Lew & Schmidt, 2011). The students were exposed to problem-based learning (PBL) where they worked collaboratively in teams of four or five. Their day consisted of initial discussion of the problem, individual study, and collaboration within the groups. Data for the study were collected based on four elements: 1) classroom performance assessed through the lens of the class tutor; 2) performance through a student activity; 3) performance based on peer evaluation; and 4) a reflective journal written by each student. Journal writings recorded students' reflection based on daily prompts provided by the tutor. In addition, every three to four weeks during the semester, students were assessed on four knowledge acquisition tests.

Data were analyzed during the 3rd week of the first semester as well as the 14th week of the second semester in the academic year (Lew & Schmidt, 2011). The journal writings and performance letter grades were recoded into numerical scaled measures. Weak correlations were reported for both weeks between the journal responses and classroom performance grades (.02 < *r* < .27). In addition, there were weak to moderately strong correlations between journal reflections and knowledge acquisition test grades (.20 < *r* < .34). The journal reflections were coded and categorized into two categories: how learning took place and what was learned. Both categories showed very little difference in helping students become more effective at learning or academic achievement. Interestingly, however, Lew and Schmidt argued that despite the weak correlations and no statistically significant differences, "it is impossible to conclude a relationship between students' ability to self-reflect and performances in the classroom as well as assessments on knowledge does not exist" (p. 537).

Lew and Schmidt (2011) reflected on some potential factors that could affect the results: a) student are generally poor at self-reflection; b) students in this study were identified as "inexperienced" meaning, they lacked experience of reflecting on what they have learned and how they learned it; and c) differences between responses in between the weeks could be due to a number of factors for example, the type of questions asked by the tutor each week. In addition, given the seeming importance of feedback (Hattie, 2012; Hattie & Timperley, 2007), it is concerning that the researchers did not indicate if the class tutors read the students' journals and conducted their own self-reflection. This would lead to adapting their instructional strategies and potentially positively impact student learning. For these reasons, it is premature to dismiss the idea that students can be competent reflective thinkers. One would also consider students' reflection as only one predictor that impacts academic achievement. Therefore, the researchers might well have considered other factors such as quality of instruction, family life and student interests.

This study and previous ones discussed offer emergent support for the need of metacognitive practice in the classroom. In 1983, the publication of *A Nation at Risk* led to a call for reform of the American education (Jones, Jones & Vermette, 2009; McCombs, 2010). About three decades later educational leaders, policy makers, teachers and parents continue to question the education system. They increasingly turn to various influences as variables for student achievement, and one of those components is the inclusion of student reflection and self-assessment (Jones et al., 2009). Teaching students how to practice metacognitive skills can positively impact academic achievement.

Effects of teacher feedback on student learning. Scholars have tried to cipher out from a large body of research on feedback that indicates there is evidence to support feedback as a powerful tool that positively influences learning outcomes (Hattie & Gan, 2011). Kluger and DeNisi (1996), in their meta-analysis review of feedback intervention, consisting of 607 effect sizes, suggested that feedback interventions improve performance on average showing a moderate overall effect size (d = .41). Additionally, Hattie (2012), placed feedback as one of the top 10 influences on student achievement (ES = .72).

Empirical studies – feedback. Butler and Nisan (1986) designed a study to test the effects of different feedback conditions on performance as well as motivation. This mixed design consisted of 261 sixth-grade children from nine classes dispersed in three city elementary schools of a predominantly middle-class population. Three classes were randomly assigned to one of three treatment groups. Group one, consisting of 88 students received task-related written feedback on their performance. Group two, 90 students, received numerical grades, and Group three of 83 students received no evaluation.

All three groups were given three assignments identified as interesting for sixth-graders determined by the pilot study. The three assignments were each administered as sessions one, two, and three respectively whereas, session one was done on one day and two days later, sessions two and three were completed with two hours in between. Each session consisted of two tasks. For sessions one and three, Task A instructed the students to construct as many words as they could from the letters of a longer word. Task B consisted of two examples from the divergent thinking "uses" test developed by Torrance and Templeton (Butler & Nisan, 1986, p. 211). For session two, in the first task, the children were asked to construct a word tree using the first and last letters of each preceding word while the other task was a test on "circles" developed by Torrance and Templeton (Butler & Nisan, 1986).

The experiment was conducted in the class during regular school hours and administered by one or two female graduate students. The instructions for session one were identical for each group and though printed in the booklet, it was read aloud. Students were given five minutes to complete each task. Two days later, at the start of session two, the booklet from session two was returned. The students from group one were told that each had received appropriate evaluation of his/her performance in the form of written comments. Students in group two were told they were given a numerical grade and group three were simply told the booklet was being returned to them. Then Booklet two was given to the students and procedures for Tasks A and B were repeated as described in Booklet one.

Two hours later, session three began but before they could begin the tasks in the third booklet, Booklet two was returned with the evaluation appropriate for each group as described earlier. Students in each group were given a few minutes to look through Booklet two and were then told that the experimenters had some tasks that had not yet been tried out (p. 211). The children were told they are to complete the tasks however; Booklet three would not be returned to them once it's submitted. The procedure for both tasks was identical to the previous ones.

The researchers tested their hypotheses using analyses of covariance (ANCOVA) with session one scores as the co-variants and scores from session three as the dependent variable. The results indicated a significant effect in scores ($F(2, 257) = 77.00$, $p < .001$) which supports the hypothesis that the performance on the quantitative task (Task A) would be higher in group one (comments group) than the other two groups (group one: $M = 55.49$, $SD = 19.26$; group two: $M = 52.59$, $SD = 25.32$; group 3: $M = 29.46$, $SD = 14.00$). Likewise, for Task B, the researchers hypothesized that performance on the qualitative task would be higher after receipt of comments (group one) than after receipt of grades (group two) and no feedback (group three). The results from the ANCOVA for the scores in session three suggested a significant effect ($F\ (2, 257) = 123.28$, $p < .001$) of manipulation for the final scores (group one: $M = 32.59$, $SD = 11.65$; group two: $M = 17.08$, $SD = 8.61$; group three: $M = 15.06$, $SD = 8.04$).

Though the results from this study support the research that task-related feedback positively affects performance, further research is warranted for several reasons. This study was conducted in 1986, close to three decades ago. Additionally, the participants in the study were sixth-graders with a median age of 12.3. One would argue, it would be unrealistic to generalize the results of the study due to simple nature that the mindset of sixth-graders are not comparable to high school aged students. Students in various grade levels perform and reflect differently to feedback received (Brookhart, 2008). Additionally, there was very little time allowed for students to reflect on their performances from both sessions one and two. Though the researchers explicitly indicated students had a few minutes to review the booklet from session two, they made no indication if students had time to review the booklet from session one prior to starting session two.

Fast forward to 2011 when Siewert investigated the types and necessity of teacher feedback for students with learning challenges. The researcher sought to determine whether fifth-graders with learning disabilities would be motivated to complete assignments when written feedback was provided within 24 hours (Siewert, 2011, p. 20). These students came from an urban city and Title 1 school located in the southeastern region of the United States. Furthermore, according to the No Child Left Behind Act of 2001 for three consecutive years, the school received a grade of C which suggests no substantial academic improvement was demonstrated school-wide.

The students who participated in the study represented a general education class with 11 boys and 11 girls. Of the 22 students, four required special education services, two were identified as gifted students and of the remaining 16 general education students, 10 were identified as at risk because of ethnicity and socio-economic status according to the school district's policy (Siewert, 2011). Prior to the study, the researcher acknowledged several concerns worth of reporting. Since this class consisted of a variety of learning needs, when students received instructional time in the area of writing, the special education students were pulled out for occupational therapy and various general education students were pulled for other instruction such as technology and violin. Additionally, 10 minutes of a 45-minute class were lost due to school wide announcements. Accompanying all of these distractions and delays which included late arrival of several students, 45 minutes of time scheduled for writing dwindled down to about 25 minutes.

The researcher observed several aspects that justified the need for an intervention. Siewert (2011) observed in students' early writing assignments punctuation and capitalization errors were prevalent and student work lacked any type of feedback whether it was written or verbal from the teacher. Students' writing assignments became incomplete and lacked effort. The goal of the intervention was to determine if provided with written feedback and correction in their punctuation and capitalization, students will feel more inclined to complete their assignments and be cognizant of their writing conventions.

The study was six weeks long with the intervention given to the student two to three times per week. Toward the last five minutes of instructional time, students were handed half sheets of paper with five sentences containing punctuation and capitalization errors. They were instructed to correctly copy the sentences below by inserting the correct punctuation and capitalization. Since the exercise was given two to three times per week, those days were strategically selected to include as many of the special education students. During times when these students were pulled for occupational therapy, they would take the paper with them and receive extra assistance. The papers were graded and written feedback from the teacher was given to the students within 24 hours.

The researcher used a number of modes for collecting data to examine the effects of the intervention. Anecdotal notations were collected at two separate points, at the beginning and at the midway point of the study. At the beginning, students were asked to write a letter to their county official about things they liked about county policies as well as address grievances about county policies that pertained to them. At the midway point, the students were asked to complete another sample writing. When both writings were collected and analyzed, the results confirmed that students need written teacher feedback to progress academically.

At the start of the study, in the area of capitalization, 31% of the total student body demonstrated correct capitalization in the first letter. By the midpoint check, the number increased to 47%. In the area of punctuation, 37% of the student body demonstrated ability to punctuate sentences and by the midpoint check, the number increased to 39%. Although this data results suggested there was in increase in the number of students demonstrating ability in punctuation and capitalization, it's important to take these percentages in perspective. The student body consisted of 22 students, so an increase by four students may seem significant, but an increase by only one student in punctuation is low. It is important to note, the students receiving exceptional special education services did not show significant improvement in capitalization however in punctuation the percent of students demonstrating ability increased from 50% to 66%. Again, one would argue that these students made up only six of the 22 students.

Students completed a five-question survey regarding their feelings toward receiving written feedback, the intervention. Of the 22 students surveyed, 78% rated the experience as positive, 63% expressed they would like to continue receiving feedback, and 72% believed the intervention made them better writers.

Of the students receiving special education services, all of them rated the best part of the intervention to be the feedback given to them by the teacher in the form of smiley faces. Additionally, all of them agreed that the intervention helped them in their writing and 75% of those students felt it was a positive experience for them. Finally, at the end of the intervention, two writing samples were collected and analyzed. These writings were considered free writes and they were completed three weeks and over one month after the study. In terms of errors in punctuation and capitalization, the percentage of errors made decreased from 61% to 26%, but more particular, students receiving special education services, the percentage of errors students made decreased from 80% to 33%.

No one would argue that effective feedback given to students in a timely manner can positively impact student learning as well as their confidence in developing the ability to understand knowledge. However, with a small sample size such as the one in this study, it is impractical to generalize the findings that written feedback benefits students who require special education services. Perhaps, many would agree with Siewert (2011), that the implications of this action research is the fact that students need feedback from teachers.

Nunez et al. (2015) examined the relationship between teachers' feedback on homework and academic achievement. The sample included 454 students from grades five to 12 from three schools in northern Spain. The aims were to determine how teacher feedback affects students completing their homework, the amount of time students spend on homework and homework management which leads to academic achievement.

Teacher's feedback on students' homework was measured using questionnaires by Walberg, Paschal, and Weinstein (1985) as well as Xu (2011): (a) the teacher emphasizes the importance of completing the homework; (b) the teacher checks whether students have done their homework; (c) the teacher takes homework into account when assigning final grades; (d) the homework was corrected in class to fix the errors students made; and (e) the teacher gives students positive reinforcement when their homework is done. During the instructions, students were asked to assess their teachers' homework feedback globally. This suggested students were not to focus on a particular teacher or class but assess on their overall perception. The students responded to the items using a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The Cronbach's alpha of $\alpha = .66$ was reported as the reliability for the instrument.

The three variables related to homework were measured using a homework survey developed by Nunez et al. (2013). Students responded to three questions pertaining to the amount of homework they complete, the perceived quality of homework time management, and the level of homework time optimization students spent when completing their homework. All questions required students to respond using a 5-point Likert-type scale. The Cronbach's alpha was reported for each question of $\alpha = .72$, $\alpha = .69$, and $\alpha = .78$ respectively.

A structural equation modeling (SEM) was analyzed using AMOS.18 (Arbuckle, 2009), to test the relationship between teachers' homework feedback as perceived by students, the three students' homework-related variables, and students' academic achievement. Based on the inter-correlations reported by Nunez et al. (2015), there were positively statistically significant correlations between students' academic achievement and the following variables: amount of homework students completed (r = .369, p < .001), perceived quality of homework time management (r = .330, p < .001), the teacher emphasizing the importance of completing the homework (r = .124, p < .001), the teacher checking whether students have done their homework (r = .177, p < .001), and the homework was corrected in class to fix the errors students made (r = .121, p < .001).

These results suggested that teachers' feedback on homework as perceived by students is positively and significantly related to the quality and amount of homework the students completed. Additionally, the quality and amount of homework completed positively and significantly predicted academic achievement. However, it is important to note that exploratory results indicated that perceived by the students, homework feedback from the teachers decreased significantly as grade levels increased. Perhaps, this warrants the need for future studies to investigate this causal relationship. Additionally, quality and type of feedback given by the teacher should be relevant to the needs of the students (Brookhart, 2008), therefore, additional studies are needed to examine how this invaluable component when linked with other formative assessments such as metacognitive practice potentially impacts academic achievement.

The idea of providing students with feedback on their work is not an innovative approach in K-12 schools, implying it can be a natural approach. However, one could argue, the type of feedback given to students should be examined and discussed. If integrated effectively, feedback promotes engagement, improved instruction, and deepens understanding (Guskey & Marzano, 2003). The studies discussed earlier support the theory that feedback motivates and encourages students to generate a desire to learn (Irons, 2008). Students will put effort into their homework (Nunez et al., 2015), generate a higher self-efficacy (Siewert, 2011) and feel motivated to improve learning (Butler & Nisan, 1986) when feedback is meaningful and provided in a timely manner.

Summary of Literature Review

Educators who integrate effective formative assessments such as metacognitive practice and feedback in the classroom as daily activities create an environment with potential to improve student learning which leads to improved measured student achievement. Reflective assessment involves students becoming metacognitive thinkers. How teachers respond to their students' reflection can be in the form of feedback. While public education is faced with the pressure of increased expectation and diminishing resources, these strategies should be considered as one of several avenues of student growth.

Hattie (2012) suggested that encouraging students to practice reflective thinking requires the teacher to promote an environment where students feel safe to be honest and open. As noted by Ellis et al. (2013), "where does the boundary lie when students are given a voice …?" (p. 8). If students are instructed to think reflectively about their learning, teachers have a responsibility to instill the value of trust, truth, openness, self-worth, and respect (Hattie, 2012). Additionally, to boost value to students' reflective practice, teacher feedback provided either in the form of written or oral dialogue (Evans, 2009) could improve student learning (Hattie & Timperley, 2007).

Kluger and DeNisi (1996) found that in 50 of 131 well designed studies, teacher feedback appeared to lessen academic achievement. They learned that the effects of feedback depended on the reactions of the recipients. On the other hand, Hattie (2012) reported an average effect size of 0.79 which puts feedback in the top 10 influences on achievement (p. 130). This would indicate that some types of feedback are more powerful than others, therefore, one has to take in consideration the differential effects of feedback on the learning as well as the learner.

Metacognitive practice should involve consideration of thought and action. There is evidence that it enhances the possibilities of learning through thoughtfully considered experience. As Donald Schön (1987), wrote, "we may reflect *on action,* thinking back on what we have done in order to discover how our knowing-in-action may have contributed to an unexpected outcome" (p. 26). Students, in turn, are taught to think using metacognitive strategies on their learning for several reasons. Reflection in the form of student voice allows the student to say for example, "I don't get it …… but I've seen this kind of problem before, therefore, I should…. ". Such thinking represents a transition from teacher to learner.

Dewey (1910) defined reflection as "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support and the further conclusion to which it tends" (p. 6). He further emphasized the idea that reflective thinking involves "an act of search or investigation directed toward bringing to light further facts which serve to corroborate or to nullify the suggested belief" (p. 9). Metacognition, thinking about one's thinking (Flavell, 1979), supports academic learning (Bandura, 1997) and, therefore, involves a number of components to its usefulness. Bandura (1997) placed the emphasis on students assuming responsibility for their own learning and therefore adopting self-regulatory and self-corrective strategies to generate successful academic outcomes.

Reflective teachers examine their instruction, their lesson plans, and students' academic achievements to revise their practice for improved outcomes. They use a mental modelling approach to determine how well students are identifying with the content. They seek to evaluate students' higher-level skills that are required for problem solving and decision making. Reflective thinking promotes critical thinking that leads to restructuring strategies of actions, understandings of phenomena, or ways of framing problems (Schön, 1987). As students benefit from practicing their metacognitive skills, teachers also benefit as it allows for more comprehensive reflection. The awareness of reflecting on teaching and learning pave the path for enriched instruction and enhanced learning. As appropriately explained by Ellis (2001), reflection is like a ship's compass. "we need to turn to it regularly in order to ensure that we are steering the true course" (p. 32).

Though numerous studies conducted suggest there is evidence to show a positive impact of reflective assessment strategies on student learning and achievement, further research is warranted to investigate the impact of metacognitive practice when linked to teacher feedback in the secondary schools more specifically in the mathematics curriculum, particularly in geometry and algebra, two required classes for most students. As stated by Schoenfeld (1987), "the relative amount of attention given to having students "think about their thinking" may just define another kind of cycle in school mathematics" (p. 269). The purpose of the present study is to further explore the efficacy of reflective self-assessment pooled with specific teacher feedback as means to improve academic achievement. This study postulates that meta-cognitive practice when linked with content-specific teacher feedback positively and significantly impact academic achievement in high school geometry students.

Chapter Three

Research Methodology

Chapter Overview

The purpose of this study was to bridge the research gap regarding the use of metacognitive practice and feedback as part of the daily teaching and learning routine for teachers and students. Though researchers have suggested there is a positive effect when students are able to reflect on their own learning and when they receive feedback, there are few studies that link reflective assessment with feedback as the independent variable specifically in secondary school mathematics. Therefore, the researcher in this study focused on the use of reflective assessment and content-specific feedback in the daily in-class learning for high school students in a required Geometry course.

In this chapter the methods, procedure and elements of statistical analysis that were utilized in the study are presented. The researcher adapted the specific intervention and procedure focused on integrating reflective assessment from several prior studies (Bianchi, 2007; Bond, 2003; Evans, 2009). The research hypotheses are stated in this chapter with an overall account of the research design. A description of the participants follows, including an explanation of how classes were randomly assigned to the comparison and experimental groups. The variables, specific interventions utilized, and procedure for the study are also presented. In the second part of the chapter, the statistical analysis of data from the pretest, posttest and retention test are discussed. The instrument utilized in the study and testing procedures are also discussed. Last, the context of the study and the research steps taken are explained in Chapter Four.

Research Hypotheses

The purpose of this study was to examine the effects of metacognitive strategies and content-specific feedback on the academic achievement of high school students in mathematics. Specifically, the researcher examined the effects of situated metacognition, in the form of reflective assessment, linked with feedback on high school students studying Honors Geometry. According to the Florida Department of Education, Geometry is a required course for all high school students. Honors Geometry is an advanced class designated for high achieving math students who want to learn at an accelerated pace and deeper level. The following null and research hypotheses have been generated based on the research questions that drove this inquiry:

Research Question 1: Is there a statistically significant difference on achievement of high school geometry students who practice metacognition or reflective assessment and receive teacher feedback, when compared to those who are provided with the same instruction but do not explicitly practice reflective techniques nor explicitly receive teacher feedback?

H_0 = There is a statistically non-significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

H_1 = There is a statistically significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

Research Question 2: Does the use of metacognitive strategies enhance student retention of Geometry concepts over time?

H_0 = There is a statistically non-significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

H_1 = There is a statistically significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

Research Design

A nonequivalent control-group design with repeated-measures, two-tailed test was employed in the study. The independent variable examined was the use of reflective assessment linked to feedback. The reflective prompts utilized in this study were based on those articulated by Ellis (2001) and Mevarech and Kramarski (1997). The dependent variable was the performance on a measure of geometry content covered over the course of the intervention. The criterion instrument was developed by the publisher of the geometry textbook utilized in the course and aligned with the geometry content knowledge (Hall, Kennedy, Bass, & Wiggins, 2012). The specific content of the instrument related to the *Pythagorean Theorem and Special Triangles* unit, a required topic taught in all geometry courses in Florida. The criterion instrument was administered prior to the intervention (pretest), at the completion of the intervention (posttest), and four weeks after the study (retention test). Table 1 presents an overview of the experimental design.

Table 1

Quasi-Experimental Design

		Geometry Sample		
Group	Pretest	Intervention	Posttest	Retention Test
N1	O	X	O	O
N2	O		O	O

The procedure lacked random selection of participants to treatments due to the constraint of preexisting classroom assignment for students. Despite this, the intact student groups were randomly assigned to either the comparison group or the experimental group. In total, this involved random assignment of five intact classes, with each class containing approximately 17 students. The characteristics of the participating student population are discussed later in this chapter.

A two-tailed test was selected for the following reasons: (1) the researcher wished to achieve a more rigorous test result, (2) a two-tailed test will provide the researcher with a “safeguard” against unexpected results, and (3) the researcher opted for a non-directional hypotheses in response to both research questions (Cho & Abe, 2013). A repeated-measures design was selected for the following reasons: (1) the researcher desired to control for the threat posed to internal validity by differential selection of the participants, (2) the researcher sought to measure student retention of the content at four weeks post intervention. Finally, a pre-test was administered because random selection at the level of participants was not feasible. The pre-test scores were utilized to elicit potential pre-existing differences between students in the experimental and comparison groups.

According to Field (2013), a repeated-measures design is used when there is a between-group comparison and data are collected from the participants at multiple time points. However, with this design, the testing poses threats to both internal validity and external validity. Regarding internal validity, the same instrument was administered to the participants on three different occasions. Because students were familiar with the assessment, there was potential for gains in the students' scores across tests (Campbell & Stanley, 1963). This phenomenon is described as students becoming "test-wise" (Gall, Gall, & Borg, 2003). To address this threat to internal validity, both the experimental and comparison groups received equivalent exposure to the instrument, which should therefore minimize the differential effects of testing between the two groups.

Regarding external validity, it is possible that the assessment would interact with the intervention in such a way that it could enhance the effect of the treatment, which is known as test sensitization (Gall et al., 2003). The researcher acknowledged the possibility that the administration of a pretest, post-test or retention test could activate the students' awareness of their attitudes toward the concept which could sensitize them to react to the content and intervention in a way that would affect the outcome.

Participants

A convenience sample was used to recruit participants at the classroom level. The sample consisted of students from five intact Honors Geometry classes taught by the same teacher in a private high school located in an urban city in Valousa County, Florida. According to the 2014 census, the city in which the school is located reported the following demographic data: 48.9% male and 51.1% female; ethnographic data include 8.7% Hispanic, 0.2% American Indian, 2.6% Asian, 31.7% African American, 54.2% Caucasian, and 2.6% Multi-ethnic. Furthermore, approximately 16.8% of the town's population is below the poverty line.

The school is a private, Catholic high school that serves students who come from both private and public middle schools in the area. At the time of the study, the school population was represented by over 50 different zip code areas within the county and consists of 474 students in grades 9 to 12. The school consisted of 48% male and 52 % female students. The ethnographic makeup of the students is as follows: 5% Hispanic, 10% Asian, 7% African American, 75% Caucasian, and 3% Multi-ethnic. Furthermore, 70% of the students received tuition assistance to help families bridge the gap between what they can afford to pay and the tuition cost. The breakdown of this assistance was as follows: 14% of the student are financially supported with one fourth of tuition assistance, 21% with one fourth to one half of tuition assistance, 7% with one half to three-fourths of tuition assistance and 28% with more than three-fourths of tuition assistance. Finally, 6% of the student body received some form of remedial accommodation based on their learning disabilities.

At the classroom level, the participating teacher reported the demographic data for students participating in this study: from a sample size of 75, 45.3% male and 54.7% female; the ethnographic data included 8.0% Hispanic, 10.7% Asian, 1.3% African American, 61.3% Caucasian, and 10.7% Multi-ethnic and 8% other. Additionally, 85.3% represent the grade 9 class and 14.7% are grade 10 students. These data are presented in Table 2.

Table 2

Demographic Information of Sample

		Frequency	Percentage
Ethnicity			
	Hispanic	6	8.0
	Asian	8	10.7
	African American	1	1.3
	Caucasian	46	61.3
	Multi-ethnic	8	10.7
	Other	6	8.0
Gender			
	Male	34	45.3
	Female	41	54.7
Grade Level			
	9^{th} Grade	64	85.3
	10^{th} Grade	11	14.7

A convenience sample was used because the students were assigned to the class period based on their schedule. The overall sample size for Honors Geometry students was 75 consisting of five classes of approximately 17 students in each class. These five intact classrooms were randomly assigned to one of the two groups by a “draw from the hat” process, which was observed by two individuals unaffiliated with the study. Of the five classes, two were randomly assigned to the experimental group and three to the comparison group. The total group sizes were 33 students in the experimental group and 42 students in the comparison group. Table 3 provides characteristics of the sample by gender and group.

Table 3

Sample by Gender

Grouping Variable	Male	Female	Total
Comparison Group	18	24	42
Experimental Group	16	17	33
Total	34	41	75

This study occurred over four weeks consisting of 16 instructional days. All five classes received instruction for ten class periods each 45 minutes long. The remaining six instructional days were identified as block periods in which the class periods were 90 minutes long and students attended half of the total number of classes each day. During the course of the study, each student attended ten 45 minute classes and three 90 minute classes.

Protection of Participants

There were no risks involved with the participants beyond the normal educational settings and practices, with only slight pedagogical differences between the experimental and comparison groups. Nevertheless, the researcher asked the students to provide assent (Appendix A) that granted the researcher permission to use their data in the study. Furthermore, the parents or guardians were asked to give consent (Appendix B) to allow the researcher to use their child's data in the study. Finally, to preserve confidentially, at no time was the researcher present in the class during the study and students' names were de-identified with random numbers.

Instrumentation

The criterion instrument used for data collection in the pretest, posttest and retention test was the unit test developed by publishers of the Geometry textbook utilized as a resource (Hall et al., 2012). The specific content of the instrument related to the *Pythagorean Theorem and Special Triangles* unit, the Geometry topic that students focused on during the study.

To examine the reliability of the instrument, the researcher conducted a test-retest analysis using the post-test and retention test. According to recommendations by Gall et al. (2003), a correlation coefficient is calculated to determine the reliability of the test scores, a direct measure of consistency, on the same measure between two different occasions. This is the most common type of reliability for tests when alternate forms are not available (Gall et al., 2003). The bivariate coefficient between the pretest and the posttest revealed a coefficient of stability of $r = .47$, which is statistically significant at the $p < .01$ level. More importantly, the bivariate coefficient between the post-test and the retention test revealed a coefficient of stability of $r = .53$, which is statistically significant at the $p < .01$ level. The measure of internal consistency was calculated using Cronbach's alpha. The Cronbach's alpha is .75, which indicates a high level of internal consistency (Field, 2013). Data related to the test-retest reliability of the instrument are presented in Table 4.

Table 4

Test-Retest Reliability of Instrument

		Pretest	Posttest	Retention test
Pretest	Pearson Correlation	1	.47**	.52**
	Sig. (2-tailed)		.000	.000
	N	75	75	
Posttest	Pearson Correlation	.47**	1	.53**
	Sig. (2-tailed)	.000		.000
	N	75	75	75
Retention test	Pearson Correlation	.52**	.53**	1
	Sig. (2-tailed)	.000	.000	
	N	75	75	75

** Correlation is significant at the 0.01 level (2-tailed).

Procedure

The intervention lasted four weeks, which aligned with the district-specified timeline for the *Pythagorean Theorem and Special Triangles* unit. Six of the instructional sessions were 90 minutes in length. The remaining 10 sessions were 45 minutes in length. Toward the end of the study, the participating teacher was absent for one day. Although the teacher provided an assignment that was aligned with the topic, the students were not exposed to the traditional instruction. This anomaly could present itself as a potential threat to internal validity so, to address this threat, all five classes were without a math instructor for a total of one day.

One comparison class and one experimental class met in the early morning while the remaining three classes met after lunch. This difference in the time of day potentially represents a confounding variable that constitutes a threat to internal validity (Gall et al., 2003). The time of day classes were held and the random assignment of the classes to the groups was beyond the control of the researcher.

Intervention. On the first day of the study, students in both groups were administered a pretest. During the remaining time, all of the students received equivalent instruction in the *Pythagorean Theorem and Special Triangles*, with one exception: The experimental group completed the metacognitive prompts and practice problems (Appendix C) and received content-specific feedback.

Over the course of the study, the teacher administered the metacognitive prompts and two problems related to the lesson taught. These prompts were administered 12 separate times to the experimental group, during the last five to 10 minutes of the instructional period. During this time, the comparison group either reviewed the learning target for that day or began their homework assignment. After receiving the metacognitive prompts and practice problems, the students in the experimental group recorded their responses on a note card, which was then collected by the teacher. This process constituted reflective assessment. The following metacognitive prompts provided to students in the experimental group included:

1. I Learned Statement (Ellis, 2001): *Today, I learned …*
2. Strategic Questioning (Mevarech & Kramarski, 1997): *I can now apply ... to solve …*

3. Clear and Unclear Windows (Ellis, 2001): *I understand ... but still don't understand …*

The participating teacher de-identified the reflective cards and made them accessible to another Geometry teacher (not affiliated with the participants) to provide the feedback. This procedure was done to avoid bias and to allow students to receive content-specific feedback. Brookhart (2008) proposed that immediate or slightly delayed feedback should be provided while students are still mindful of the learning goal, concept, or assignment. In this study, students received content-specific feedback within three to five days which could be considered slightly delayed feedback. Because the *Pythagorean Theorem and Special Triangles* unit included learning targets that cumulatively scaffold each other, the feedback provided by the teacher remained relevant and applicable throughout the study.

The teacher provided content-specific written feedback on the reflective assessment card in response to any specific questions or comments each student posed and the work shown by each student on the assigned problems. Additionally, common trends such as misconceptions in areas where most students showed they struggled were identified and communicated to the participating teacher. When the reflective cards with specific content feedback were returned to the experimental group at the beginning of the class, the teacher provided additional general feedback that would improve students' understanding of the concept. In contrast, the comparison group began class by practicing problems to review their prior knowledge. Quality feedback, in terms of content-specific and general can influence instructional revision in a positive sense when it is immediate and focused on student reflection and learning (Black & Wiliam, 1998; Guskey & Marzano, 2003; Hattie, 2012).

Attitudinal survey. Following the completion of the *Pythagorean Theorem and Special Triangles* unit, the criterion instrument was administered for the second time as the posttest to both groups. Additionally, a survey (Appendix D) developed by the researcher consisting of four questions on a 5 point, Likert scale (*1* = strongly disagree, *2* = disagree, *3* = neutral, *4* = agree, *5* = strongly agree) and one open ended question was administered after students submitted their completed assessment. The purpose was to measure the students' attitude on learning the concepts, how they felt in terms of being prepared for the end of unit assessment, and to inquire from the reflective group, their perception on reflecting and receiving feedback. The reliability of the survey according to *Cronbach's alpha*, was .51. This value suggests a medium level of internal consistency with this specific sample. Only students in the experimental group were required to respond to the open ended prompt since it pertained to the reflective assessment and feedback intervention. Exactly four weeks after the post-test, the same end of unit assessment was administered for the purpose of measuring longer-term retention of unit content in both groups.

Data Analysis

The researcher used SPSS version 23 general linear model to address the hypotheses. Descriptive data were analyzed to ensure parametric procedures would be appropriate (Field, 2013). Tables 5 and 6 provides the descriptive data for the variables used in the study. It is important to note that there was a non-normal distribution of the data. This statistic was confirmed by both the Kolmogorov-Smirnov and Shapiro-Wilk tests' of normality in Table 7. In applications with a moderate to large sample size, ANOVA with repeated measures only require approximately normal data because it is robust to violations of normality (Field, 2013). The researcher conducted an additional statistical analysis to determine if an adequate sample size was utilized in the study.

A priori power analysis using G*Power 3 (Faul, Erdfelder, Buchner, & Lang 2007) for a two-tailed, repeated measures mixed ANOVA with six cells and three measures was conducted to determine an adequate number of participants using an alpha of 0.05, a power of 0.80, and an effect size of f = 0.20. Output from these parameters indicates that a sample size of 78 participants will result in an 83% chance of detecting an effect if one actually exists. This study employed a sample size of 75, therefore, though the assumption was violated, the test can still produce valid results. Appendix E provides the data output from G*Power 3.

Table 5

Descriptive Statistics for Post-Test

Group	*N*	Mean	*SD*	Skewness	Kurtosis
Comparison	42	81.38	15.47	-1.14	.64
Experimental	33	89.06	11.06	-1.27	1.29
Total	75	84.76	14.15	-1.29	1.31

Table 6

Descriptive Statistics for Retention Test

Group	*N*	Mean	*SD*	Skewness	Kurtosis
Comparison	33	72.24	13.92	-.53	1.30
Experimental	42	79.42	9.98	-1.12	2.25
Total	75	75.40	12.78	-.83	1.52

Table 7

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre-Test	.10	75	.03	.96	75	.02
Post-Test	.15	75	.00	.87	75	.00
Retention Test	.08	75	.20*	.95	75	.00

* This is a lower bound of the true significance.

The researcher computed inferential analysis using an analysis of variance (ANOVA) with repeated measures two tailed test. The purpose for using the ANOVA with repeated measures was to examine the main effects of the independent variable: *group* with two levels (experimental and comparison) over time. When comparing mean scores, this approach, as opposed to other statistical procedures such as multiple ANOVAs or an analysis of covariance (ANCOVA), is considered powerful and reduces the likelihood of Type 1 error (Field, 2013; Tabachnick & Fidell, 2014). The analysis produces an *F* ratio of within-group differences and between-group differences. The *F*-statistics indicates if there is a significant difference between the mean scores.

In order to obtain valid results from using an ANOVA with repeated-measures, several statistical assumptions must be met: (a) the dependent variable should be measured at the continuous level; (b) the within-subject factor should consist of at least two categorical, “related groups”; (c) the between-subjects factor should consist of at least two categorical, “independent groups”; (d) there must be no significant outliers in each group; (e) the distribution of the dependent variable should be approximately normally distributed; (f) there needs to be homogeneity of variances for each combination of the groups; (g) the variances of the differences between all combinations of groups must be equal (Field, 2013).

The results of each statistical analysis are presented in Chapter Four. The assumptions underlying the statistical procedures utilized in this study are reviewed, followed by a discussion of suitability with respect to the obtained data. Inferential statistics are presented and summarized in terms of their significance for each of the research hypotheses. Finally, the results from the qualitative analysis are presented to expound on the students’ attitude toward the intervention.

Chapter Four

Results

Chapter Overview

In this chapter, the results of the study are presented in order of the research questions posed by the investigator. Additionally, descriptive statistics for all relevant variables are provided including measures of central tendency, variability, and characteristics pertaining to the normality of each distribution. The assumptions underlying the statistical procedures utilized in this study are also reviewed, followed by a discussion of suitability with respect to the obtained data. Inferential statistics are presented and summarized in terms of their significance for each of the research hypotheses. Finally, the results from the qualitative analysis is presented to expound on the students' attitude toward the intervention.

Research Questions

In the first research question, the researcher wanted to determine if there is a statistically significant difference in academic achievement of high school geometry students who practice reflective assessment and receive content specific feedback and those who do not practice reflective assessment. In the second research question, the researcher further attempted to determine if there is a statistically significant difference on post-test and retention test scores when the retention test is administered four weeks after the study. Both research questions were tested using an analysis of variance (ANOVA) with repeated measures two tailed test. The significance of effects was analyzed at an alpha level of .05 (Gall et al., 2003).

Descriptive Statistics

Prior to computing inferential statistics, the data were scanned for missing scores as well as any outliers. One case was identified in which the student was missing a pretest score. To address this, the missing case was replaced with the mean score of the pretest. Data were analyzed to ensure parametric procedures would be appropriate. Table 8 provides the descriptive for the pretest, post-test and retention test. Each of these variables represents a separate administration of the same instrument, *Pythagorean Theorem and Special Triangles* Test. The possible range of scores on the instrument was 0 to 100. Tables 9, 10 and 11 present the data disaggregated by group assignment for all three variables.

Table 8

Descriptive Statistics for Pre-Test, Post-Test, and Retention Test

	N	Mean	*SD*	Skewness	Kurtosis
Pre-Test	75	19.15	10.55	.68	.92
Post-Test	75	84.76	14.15	-1.29	1.31
Retention Test	75	75.40	12.78	-.83	1.52

Table 9

Descriptive Statistics for Pre-Test

Group	*N*	Mean	*SD*	Skewness	Kurtosis
Comparison	42	17.31	9.95	.15	-.78
Experimental	33	21.48	10.98	1.17	1.82

Table 10

Descriptive Statistics for Post-Test

Group	*N*	Mean	*SD*	Skewness	Kurtosis
Comparison	42	81.38	15.47	-1.14	.64
Experimental	33	89.06	11.06	-1.27	1.29

Table 11

Descriptive Statistics for Retention Test

Group	*N*	Mean	*SD*	Skewness	Kurtosis
Comparison	42	72.24	13.92	-.53	1.30
Experimental	33	79.42	9.988	-1.12	2.25

The pretest was administered prior to the study. According to the data presented in Table 8, pretest scores (M = 19.1, SD = 10.5) suggest that students knew very little of the unit content prior to the intervention. Additionally, both skewness and kurtosis statistics for the pre-test distribution fall within plus or minus one. This suggests the data for the pre-test produced a normal distribution.

The post-test was administered at the completion of the intervention. According to the data presented in Table 8, post-test scores (M = 84.7, SD = 14.1) suggest a ceiling effect occurred. The mode reported for the post-test was 99, which is one point below the maximum possible score. Both the skewness and kurtosis statistics for the post-test distribution fall outside the range of plus or minus 1, which suggest a non-normal distribution (Gall, et al., 2003). This was confirmed by both the Kolmogorov-Smirnov and Shapiro-Wilk tests' of normality. The skewness statistics of -1.29 (*SE* = .27) indicates a negative skew to the data and the kurtosis statistics of 1.31 (*SE* = .54) shows a peak in the data. The ceiling effect is a possible explanation for the negative skew and the mode in the post-test explains the kurtosis statistics. Figure 1 provides an illustration of the distribution of the post-test data.

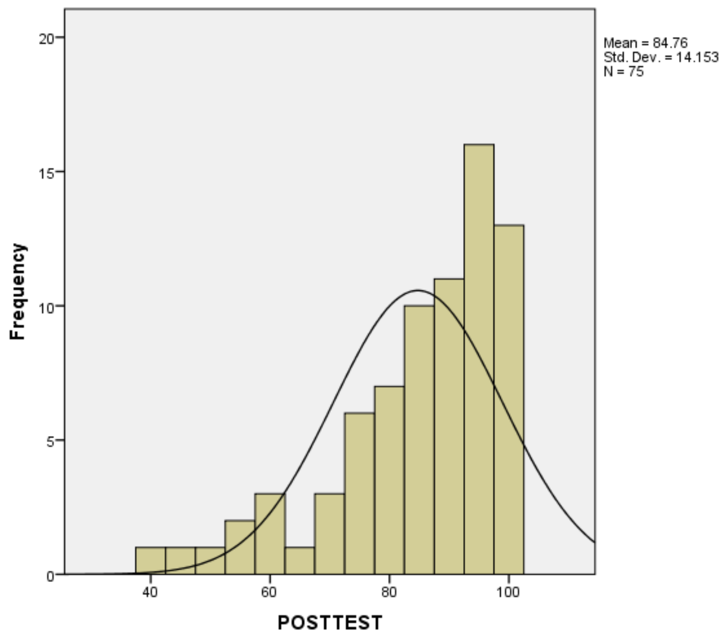


Figure 1. Post-Test Scores

The retention test was administered four weeks after the post-test. According to the data presented in Table 8, retention test scores (*M* = 75.4, *SD* = 12.7) suggest there was a slight regression from the post-test scores. Although the skewness statistics of -.83 (*SE* = .27) for the retention test distribution fell within plus or minus one, the kurtosis statistics of 1.52 (*SE* = .54) suggests a peak in the scores. Figure 2 shows a mode of 70 as a possible explanation for the kurtosis statistics falling outside the plus or minus one range.

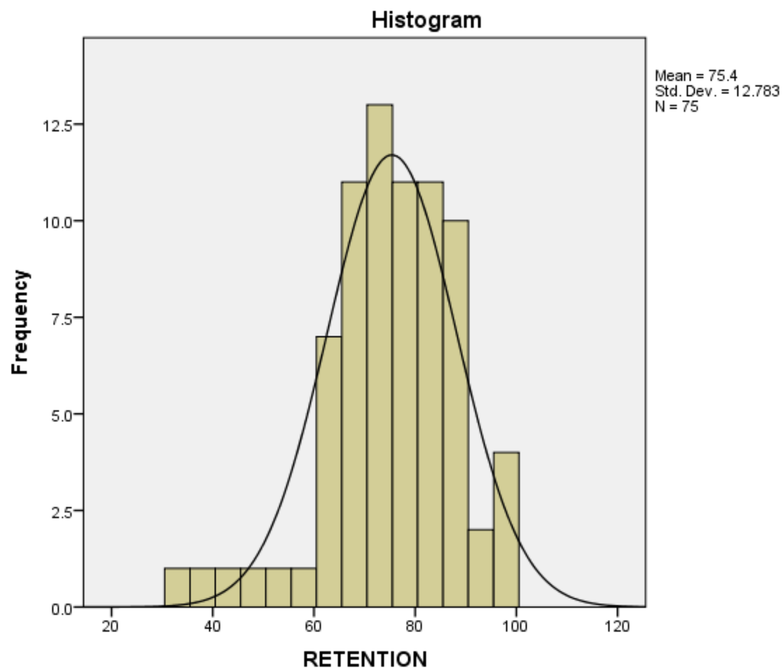


Figure 2. Retention Test Scores

Inferential Statistics

The researcher addressed the statistical assumption by first confirming that the dependent variable represented numerical test scores measured on a continuous scale ranging from 0 to 100 points. Both the comparison and experimental groups were measured at three separate times, which confirms that the within-subject factor consist of three "related groups." The between-subjects factor was organized into two independent groups: the experimental and comparison groups. After carefully scanning through the data, there were no obvious outliers; that is, any single data points that do not follow the usual pattern.

The researcher used SPSS version 23 to confirm the remaining assumptions were not violated. Normality of the data was tested using Shapiro-Wilk test of normality and results in Table 12 confirmed a non-normal distribution in the measures. The main threat to normality was the distributions of the skewness (-1.29 & -.83) and kurtosis (1.31 &

1.52) statistics in post-test and retention test scores. Those values that fall outside the plus or minus one range confirm the violation of normality. According to Field (2013), an ANOVA with repeated measures only require approximately normal data because it is robust to violations of normality. Furthermore, in applications with a moderate to large sample size, ANOVA with repeated measures may yield reasonably accurate *p* values even when the normality assumption is violated (Field, 2013). A priori power analysis using G*Power 3 (Faul et al., 2007) for a two-tailed, repeated measures ANOVA with six cells and three measures was conducted to confirm the adequate sample size of 75 was used for this study.

Table 12

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre-Test	.10	75	.03	.96	75	.02
Post-Test	.15	75	.00	.87	75	.00
Retention Test	.08	75	.20*	.95	75	.00

* This is a lower bound of the true significance.

Assumption of homogeneity of variance using Levene's Test of Equality of Error Variances was conducted. Table 13 shows non-significant values for all the variables ($p > .05$), suggesting that the variances are homogeneous for all levels of the repeated-measures variables. Finally, Table 14 shows Mauchly's test statistics is non-significance ($p > .05$), which implies the variances of the differences between groups are roughly equal.

Table 13

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Pre-Test	.008	1	73	.893
Post-Test	3.513	1	73	.065
Retention Test	1.610	1	73	.209

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + GROUP
Within Subjects Design: Time

Table 14

Mauchly's Test of Sphericity

Measure: Assessments

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
TESTS	.971	2.095	2	.351	.972	1.000	.500

Research Question One

In research question one, the researcher examined if there was a statistically significant difference between the comparison and experimental groups on the post-test scores administered to the students at the end of the study. When the ANOVA with repeated measures two tailed test was conducted, the *Tests of Between-Subjects Effects* indicated that there was significance in the main effect of the variable group (F(1, 73) = 7.27, p = .009, η_p^2 = .091). This suggests that at an alpha level of $p < .05$, there was a statistically significant difference between groups (comparison and experimental) on the end of unit assessment. At all three times the test was administered (pre-test, post-test and retention test), the experimental group outperformed the comparison group. Specifically, there was a statistically significant difference between groups in the post-test (p = .02, Cohen's d = .57). This effect size calculated using Cohen's d formula is considered medium in magnitude (Cohen, 1988, 1992).

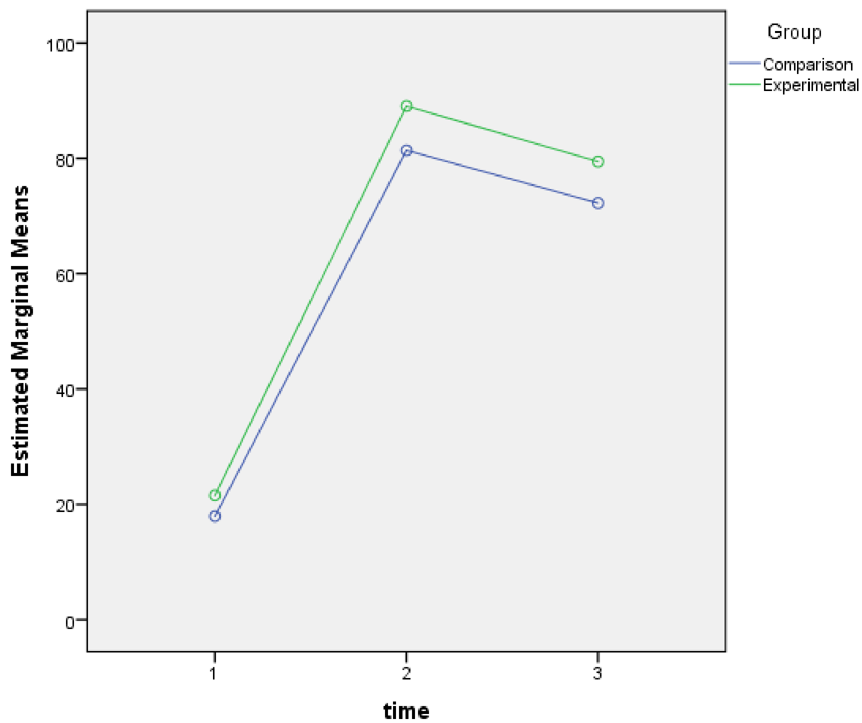


Figure 3. Estimated Marginal Means of Test

Research Question Two

Research question two explores if there is a statistically significant difference between the post-test scores and retention test scores. According to the ANOVA with repeated measures, *Tests of Within-Subjects Effects* showed a statistically significant time effect ($F(1, 73) = 1185$, $p = .000$, $\eta_p^2 = .942$). Furthermore, *Pairwise Comparisons* confirmed a mean difference of 9.390 between the post-test and retention test to be significant ($p = .00$, Cohen's $d = .69$). The retention test scores were lower than the post-test scores. This decrease in scores can be interpreted to mean that while reflective assessment techniques with feedback may enhance student performance, there could be other factors that contributed to the decline in scores. However, it is worth noting that similar to the post-test scores, the experimental group significantly outperformed the comparison group in the retention test scores ($p = .01$, *Cohen's* $d = .59$).

Qualitative Analysis

Qualitative analyses were carried out to determine if students' attitude about reflective assessment linked with content-specific feedback could be further differentiated. Following the post-test, both groups (comparison and experimental) were asked to complete a survey. Table 15 illustrates the bivariate correlation for the survey disaggregated by group.

Table 15

Correlations for Attitudinal Survey

Group				1	2	3	4
Comparison	1.	I enjoyed studying Ch. 8.	Pearson (*r*)	1	.320	.318	.416*
(*N* = 36)			Sig. (2-tailed)		.057	.059	.012
	2.	I was given an opportunity to reflect on my learning and express when I was struggling on the content.	Pearson (*r*)		1	.557**	.165
			Sig. (2-tailed)			.000	.336
	3.	I was provided with helpful feedback on my reflection on learning.	Pearson (*r*)			1	-.191
			Sig. (2-tailed)				.263
	4.	I felt prepared for this unit test.	Pearson (*r*)				1
			Sig. (2-tailed)				
Experimental	1		Pearson (*r*)	1	.058	.164	.332
(*N* = 27)			Sig. (2-tailed)		.773	.415	.091
	2		Pearson (*r*)		1	.525**	-.229
			Sig. (2-tailed)			.005	.250
	3		Pearson (*r*)			1	-.014
			Sig. (2-tailed)				.944
	4		Pearson (*r*)				1
			Sig. (2-tailed)				

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Of the 42 students in the comparison group, 36 students completed the survey and out of 33 students in the experimental group, 27 students completed the survey. Based on responses on a 5-point Likert Scale (*1* = strongly disagree, *2* = disagree, *3* = neutral, *4* = agree, *5* = strongly agree), interestingly, 29 out of the 36 (80.6%) students in the comparison group believed they were given an opportunity to reflect on their learning and express when they were struggling on the content whereas, 23 out the 27 (85.2%) students in the experimental group felt the same.

Similarly, 30 out of 36 (83.3%) students in the comparison group agreed or strongly agreed that they were provided with helpful feedback on their reflection on learning and 21 out of 27 (77.8%) students in the experimental group agreed or strongly agreed. Finally, when asked how prepared they felt for the post-test, in the comparison group, 31 out of 36 (86.1%) students agreed or strongly agreed and in the experimental group, 25 out of 27 (92.6%) students agreed or strongly agreed.

Students in the experimental were provided with an additional open-ended prompt. Students were asked “What did you like and did not like about the exit ticket you completed at the end of each class period?” Of the 27 who completed the prompt, 24 students provided positive perspectives on reflective assessment when linked with content-specific feedback and 16 students provided a negative perspective. Appendix F provides students narrative responses to the prompt.

Summary of Results

An ANOVA with repeated measures two-tailed test was computed in order to test the two null hypotheses. Accordingly, the following null and research hypotheses have been generated based on the research questions that drove this inquiry:

Research Question 1: Is there a statistically significant difference on achievement of high school geometry students who practice metacognition or reflective assessment and receive teacher feedback, when compared to those who are provided with the same instruction but do not explicitly practice reflective techniques nor explicitly receive teacher feedback?

H_0 = There is a statistically non-significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

H_1 = There is a statistically significant difference for Group (two levels: reflective/feedback and non-reflective/feedback) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

Research Question 2: Does the use of metacognitive strategies enhance student retention of geometry concepts over time?

H_0 = There is a statistically non-significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

H_1 = There is a statistically significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

Prior to performing the parametric procedure, the data were analyzed to check for major violations of parametric assumptions. Additionally, descriptive statistics were computed for all groups and reported. Parametric statistical data related to each of the two research questions that drove this study were reported. The results showed there was significance in the main effect of the variable group, which means that at an alpha level of $p < .05$, there was a statistically significant difference between the comparison and experimental groups on the end of unit assessment. Furthermore, there was a statistically significant difference between the post-test and retention test scores. Consequently, the researcher rejected both null hypotheses. Finally, the results of the qualitative analyses conducted were reported which included coding of the students' responses and comparison of students' survey questions.

The following chapter provides a summary of the purpose of this study and the methodology employed. The practical significance of the research findings is examined within the context of prior studies. A discussion of the limitations of this study is included, along with suggestions for future research.

Chapter Five

Discussion of Results and Conclusion

Chapter Overview

The purpose of this study was to determine the effects of metacognitive strategies and content-specific feedback on the academic achievement of high school students in mathematics. Specifically, the researcher examined the effects of situated metacognition, in the form of reflective assessment, and content-specific feedback on high school students studying geometry. Additionally, it was the intent of the researcher to apply the theories of metacognition and content-specific feedback that have been articulated by Bandura (1997), Dewey (1910), Flavell (1977), Hattie and Timperley (2007) and Vygotsky (1978) to the classroom setting.

Participants in the study consisted of a convenience sample of honors geometry students in grades 9 and 10 in a private high school located in Daytona Beach, Florida. Beyond answering the specific research questions, an additional aim in this study was to contribute to the growing body of knowledge pertaining to effective ways to use metacognitive instruction and provide effective content-specific feedback to improve student achievement and learning.

In the first part of this chapter, the researcher provides a rationale for this study and the methodology employed. The practical significance of the research findings is also examined within the context of prior studies. In the second part of this chapter, a discussion of the limitations is included, along with suggestions for future research.

Rationale for the Study

The research questions investigated in this study were developed for three main purposes. First, an extensive review of literature suggests metacognitive practice has potential to improve student achievement in mathematics (Bond, 2003; Kramarski & Mevarech, 2003). Similarly, other studies, for example, Butler and Nisan (1986) and Nunez et al. (2015) imply teacher feedback can be an effective predictor of academic achievement. While both strategies are highly regarded as best practices as suggested by Hattie (2012), there appear to be a limited number of empirical research studies that explicitly link reflective assessment with content-specific feedback specifically in secondary school mathematics.

Second, high-stakes standardized testing has heightened the pressure for teachers to help their students produce favorable outcomes on academic achievement (Guth et al., 1999). The current wave of these standardized tests takes into account how well students can perform on achievement tests designed by others, but seldom are students asked whether what they are being tested on is meaningful to them. Metacognitive practice provides an opportunity for students to determine how work done in class connects with their sense of meaningfulness of what is taught (Bandura, 1997). Advocates of reflective practice argue it is a skill that must be taught and used daily in order to produce effective outcomes (Borich, 2014; Costa, 2001; Ellis, 2001).

Third, quality feedback on students' reflective writing helps students learn (Ramsden, 2003). Irons (2008) and Brookhart (2008) posited that the formative activities involved when giving feedback should be relevant to the content and perceived as a worthy task for students to accomplish. Feedback is perceived as authentic and meaningful when the quality is well-thought-out and provided to students in a timely manner (Hattie, 2012). Similarly, noted by Brookhart (2008), feedback is a critical extension of formative assessment that should be used to help learners understand what they need to do to improve their learning as well as what was done well.

Research Methodology

A quasi-experimental, nonequivalent control-group design with repeated-measures was employed in the study. The independent variable examined was the use of reflective assessment linked to content-specific feedback. The dependent variable was the performance on the criterion instrument consisting of the geometry content covered over the course of the intervention. Descriptive and inferential statistics were computed to address the research questions. An ANOVA with repeated measures, two-tailed test was utilized for testing the hypotheses at a significance level of .05.

Discussion of Results

The results of this study offer tentative support for reflective strategies linked with content-specific feedback embedded as formative assessments in daily activities. Because of a lack of studies explicitly linking both strategies, it is premature to confirm any effects the intervention had in the learning environment. The findings of this study are reviewed and discussed in order of the research questions posed in Chapter One.

Research question one. Is there is a statistically significant difference on academic achievement of high school geometry students who practice metacognition and receive content-specific feedback, when compared to those who are provided with the same instruction but do not explicitly practice reflective techniques nor explicitly receive content-specific feedback? For the purpose of this study, metacognitive practice is defined as students reflecting on what they learned by responding to reflective prompts provided by the instructor and practicing two problems aligned with the content studied in class (Costa, 2001; Flavell, 1979; Schoenfeld, 1987). Feedback provided was content-specific and personalized based on each student's responses to the reflective prompts as well as the performance on the practice problems.

Research question one generated the following null and statistical hypotheses:

H_0 = There is a statistically non-significant difference for Group (two levels: experimental and comparison) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

H_1 = There is a statistically significant difference for Group (two levels: experimental and comparison) on academic achievement of high school geometry students as measured by their score in the end of unit assessment.

An ANOVA with repeated measures, two-tailed test showed that there was a statistically significant main effect for Group. The results along with the mean scores for both groups show that the experimental group outperformed the comparison group in the end of unit assessment. Previous studies focused either on reflective assessment or teacher feedback have shown that these approaches have positive effects on student achievement (Bond & Ellis, 2013; Butler & Nisan, 1986; Evans, 2009; Kramarski & Mevarech, 2003; Nunez et al., 2015). However, because the cited studies did not explicitly link reflective assessment with content-specific feedback, it is difficult to conclude the findings support the work of previous studies. For this reason, further research is warranted to examine the effectiveness of metacognitive practice linked with content-specific feedback on academic achievement.

In terms of setting and reflective assessment, the present study is similar to the work of Evans (2009) whose sample consisted of grade 9 high school English language students. In both Evans' (2009) and the present study, the researchers focused on the effectiveness of reflective assessment when used daily. In the present study, the researcher randomly assigned five intact classes to one of two groups, comparison and experimental, while Evans (2009) randomly assigned nine intact classes to one of three groups, control, comparison and experimental. The data gathered from both studies revealed a statistically significant difference in the students' achievement scores favoring the experimental group on both the post-test and retention test.

Three significant differences exist between the present study and the study conducted by Evans (2009). First, in the present study, the researcher used a nonequivalent control-group design with repeated measures. A pre-test was administered to the students to compensate for the non-equivalent group design and though the experimental group outperformed the comparison group, there was no statistically significant difference between both groups in terms of ability. The data from the pre-test showed that students in both groups knew very little of the unit content prior to the intervention. Evans (2009) used a post-test only control group design, which indicated that a pre-test was not administered to the students.

Second, in the present study, both the experimental and comparison groups consisting of five intact classes were instructed by the same teacher. In the study conducted by Evans (2009), three participating teachers each taught three of the nine intact classes. Although the classes were randomly assigned to the teachers, the three instructional styles could have affected the outcome of the study. This potential threat to internal validity is called selection bias (Campbell & Stanley, 1963).

Third, in the present study, the participating teacher collected the reflective cards at the end of each period and after de-identifying them, gave them to another mathematics teacher to provide the students with content-specific feedback, which were then returned to the students for further review. In the study conducted by Evans (2009), the participating teachers collected and maintained the written responses to the prompts from the students. The researcher then collected the student work during class visits and maintained them as part of the record-keeping for the study. It must be noted that in the study conducted by Evans (2009), no explicit feedback based on the written responses to the reflective prompts was provided to the students.

In terms of feedback, the present study is similar to the work of Nunez et al. (2015) who examined the relationship between teachers' feedback on homework and academic achievement. The present study was conducted in a high school class in the United States and focused on geometry students primarily from grade 9. The study conducted by Nunez et al. (2015) focused on students from grades 5 – 12 in three schools in northern Spain, and the researchers found teachers' feedback on homework was positively and significantly related to the quality and amount of homework the students completed. Additionally, the quality and amount of homework completed positively and significantly predicted academic achievement.

Two significant differences exist between the present study and the study conducted by Nunez et al. (2015). First, in the present study, written feedback was content-specific and provided based on the students' daily reflective assessment as opposed to Nunez et al. (2015) who provided feedback on students' homework, which was not identified as a form of reflective writing. Additionally, the type of feedback provided in the Nunez et al. (2015) study was a letter grade based on completion along with positive reinforcement. It is unclear if the positive reinforcement was written or oral. Though this quality of feedback positively and significantly predicted academic achievement, it was not content-specific.

Second, in the present study, the researcher utilized a quasi-experimental design and analyzed the data using an ANOVA with repeated measures, two-tailed test. The purpose of this type of test was to determine the impact the intervention had with the experimental group when compared with the comparison group over time. In the study conducted by Nunez et al. (2015), a structural equation model (SEM) was analyzed using AMOS 18 (Arbuckle, 2009) to test the relationship between the teachers' feedback as perceived by the students, the homework-related variables, and student achievement.

The qualitative data in the current study provided further insight about the efficacy of metacognitive practice and content-specific feedback. Following the post-test, a survey was administered to the students. The purpose was to determine if reflective assessment, when linked to content-specific feedback, could be differentiated. Students from both the comparison and experimental groups were asked to voluntarily complete a survey in which they responded using a 5 point Likert Scale (*1* = strongly disagree, *2* = disagree, *3* = neutral, *4* = agree, *5* = strongly agree). According to the results, there was no statistically significant difference between the experimental group and the comparison group. A high percentage of students in both groups believed they were given the opportunity to reflect on their learning and express when they were struggling with the content. It is unclear why students' responses were similar in both groups.

Interestingly, based on the bivariate correlation for the comparison group, there was no statistically significant difference between the students who enjoyed learning the content and their opportunity to reflect on learning (p = .06) and being provided with feedback (p = .60) at the $p < .05$ level. Therefore, there is no clear explanation for the findings in the comparison group. On the contrary, for the experimental group, it was no surprise that there was a statistically significant correlation between students reflecting and receiving helpful feedback. These results and the narrative responses from the open ended prompt administered to the experimental group support the anecdotal findings by Siewert (2011) and Zan (2000).

Research question two. Is there is a statistically significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study?

The second research question generated the following null and statistics hypotheses:

H_0 = There is a statistically non-significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

H_1 = There is a statistically significant difference on scores (two levels: post-test and retention test) when the retention test is administered four weeks after the study.

An ANOVA with repeated measures showed that there was a statistically significant time effect. The *Tests of Within-Subjects Effect* showed that scores changed over time. In this study, there was a decrease in scores from the post-test to the retention test. *Pairwise Comparisons* confirmed a mean difference of 9.39 between the post-test and retention test significant at the $p < .05$ level. However, students in the experimental group (M = 79.42, SD = 9.98) continued to outperform the students in the comparison group at a statistically significant level (M = 72.24, SD = 13.921).

Limitations of the Study

Apart from some specific limitations discussed earlier, there are other factors that limit the generalizability of this research. The limitations discussed in this section are categorized according to research design, participants, and methodological weaknesses.

Research design. The quasi-experimental, non-equivalent group design raises an immediate concern related to differential selection. Although the five intact classes were randomly assigned to one of the two groups, the design lacked random assignment at the level of the participants. This main threat to internal validity is the possibility that group differences in the post-test are attributed to the pre-existing group differences rather than the treatment effect (Gall et al., 2003). To mitigate this threat, a pretest was administered prior to the study. Despite this, it must be noted that statistical control of such differences is inferior to random assignment of subjects. However, given the difficulties of randomly assigning students who take a particular class, in this case honors geometry, at different times of the day, it would have been impossible to achieve random subject assignments. The utilization of intact classes represents a compromise, one which reflects the real world of secondary schools.

Participants. A limitation with regard to the participants, was that a convenience sample was employed in the present study. Because the participants did not consist of a scientifically selected probability sample, researchers argue that the derived inferential statistics cannot be interpreted meaningfully (Gall et al., 2003). A related matter is that the use of a convenience sample raises a threat to external validity, specifically in terms of population validity. When a sample such as the one made available for the present study is not necessarily reflective of a broad population, inferential statistics should be used with caution when certain conditions are not met (Gall et al., 2003). To address this issue, several characteristics of the sample were provided in Chapter Three including details pertaining to the participants in the study, the sample they were drawn from, and the defined population.

Another limitation, known as the Hawthorne effect, raises a threat to external validity. The Hawthorne effect occurs when individuals are aware that they are participating in an experiment (Gall et al., 2003). The nature of the current study required assent from the students, thus raising the possibility of the Hawthorne effect. Therefore, the external validity of the treatment was potentially compromised and encumbers the ability to generalize the findings.

Methodology. Three potential limitations with respect to the study's methodology surfaced. First, the same criterion instrument was used for the pre-test, post-test and retention test. With regard to internal validity, this is a possible concern associated with testing effect (Campbell & Stanley, 1963). Because students were familiar with the assessment, there was potential for gains in the students' scores across tests, which is known as becoming "test-wise" (Campbell & Stanley, 1963; Gall et al., 2003). To alleviate this threat to internal validity, both groups received equivalent exposure to the instrument, thus minimizing the differential effects. Regarding external validity, it is possible that pre-test and post-test sensitization occurred. Sensitization occurs when the pre-test serves as a learning experience on its own, which has meaningful impact on the treatment. This potential interaction of testing with the treatment hinders the ability to generalize from the study's findings (Gall et al., 2003).

Another potential limitation was the timeliness of the feedback provided to the students. Feedback was provided within three to five days from the time the students completed their reflection cards. Since researchers suggest feedback should be provided in a timely manner (Hattie, 2012, Hattie & Timperley, 2007), the delay of feedback could have negatively impacted the validity of the test scores. However, according to Brookhart (2008), slightly delayed feedback can be meaningful as long as it is provided while students are mindful of the learning goal and content.

Third, the researcher provided the feedback to the students. Although the purpose of this protocol was to strengthen the validity of the study, it raises the concern of experimenter bias. Experimenter bias occurs when the researcher unintentionally influences the results to produce a certain outcome (Gall et al., 2003). However, because the researcher was a former geometry teacher and not affiliated with the participants, the feedback provided was content-specific and therefore, avoided other potential biases related to the personal knowledge of the students.

Implications of the Findings and Suggestions for Future Research

The findings from this study offer a modest contribution to the body of empirical research on the impact of metacognitive practice and content-specific feedback on academic achievement at the high school level. Further studies are warranted to add to the body of literature and more specifically to provide greater clarity regarding the magnitude of the current investigation. Although the findings from the study show moderate effect sizes, based on the limitations referenced in this chapter and lack of studies that link student reflection with teacher feedback, the researcher recommends further study to support any broad-based conclusions. To date, the majority of studies have focused on either the impact of reflective assessment or the impact of teacher feedback (Bond, 2013; Butler & Nisan, 1986; Evans, 2009; Kramarski & Mevarech, 2003; Lew & Schmidt, 2011; Nunez et al., 2015; Siewert, 2011; Zan, 2000).

Future studies should be crafted to include larger carefully selected samples across diverse settings to examine the effects of reflective assessment linked with content-specific feedback on academic achievement and to probe its validity and usefulness for a broader population. This could include conducting studies across a variety of disciplines and grade levels from elementary to college level with the intent to more clearly develop a clear portrait of how the use of reflective assessment, when linked with teacher feedback, impacts learning and retention. Additionally, studies that employ the use of various designs and analyses are necessary to yield more generalizability. Although studies involving a convenience sample can provide valuable insights, "repeated replication of the findings is much stronger evidence of their validity and generalizability than is a statistically significant result in one study" (Gall et al., 2003).

Apart from conducting studies with the intent to examine broad-based effects of metacognitive practice and feedback, studies that examine discrete aspects of implementation are also recommended. For instance, advocates of reflective practice argue it is a skill that must be taught in order to be utilized effectively (Borich, 2014; Costa, 2001; Ellis, 2001). Perhaps, similar to the study conducted by Kramarski and Mevarech (2003), future studies should involve in-service teacher training that focuses on pedagogical practices involving metacognition. Furthermore, explicit metacognitive practice should be addressed in the classroom.

Additionally, it would benefit teachers to understand what qualifies as good feedback and decide how it should be given based on students' abilities, learning needs, and interests (Brookhart, 2008). Another topic of interest for future studies is the timeliness of feedback. When effective feedback is timely, it enables the students to process and implement the feedback (Brookhart, 2008). In turn, students become more receptive to the feedback while they are still mindful of the topic, assignment, or performance in question. In other words, feedback should be given when there is still time to correct errors. Otherwise, when it is no longer relevant to the current or future content, the feedback is pointless (Kulik & Kulik, 1988).

Implications for Classroom Practice

In Chapter Four, the effect sizes reported show that the results obtained carry practical significance for both teachers and students in classroom environments. This strategy that involved class closure in the form of reflective assessment may have positively affected what the students learned and the depth at which they learned it, when content-specific feedback was provided to each student.

Based on the findings of the current study, the researcher concludes that formative assessment, when linked with content-specific feedback, led to improved learning and higher academic achievement. Therefore, based on the results and the growing body of research that demonstrate their effectiveness in the classroom environment, it is recommended that educators become informed about the efficacious potential of metacognition and feedback in student learning.

For the students, reflective assessment provides an opportunity to take ownership of their learning on a regular basis. As suggested by Flavell (1979), by encouraging students to reflect and "think about their thinking," they foster a skill set that transcends the classroom experience and benefits their long-term learning process. For the teachers, students' reflection informs their instruction to promote improved learning and to better k now their students' thought processes.

There is an abundance of empirical evidence that supports the argument that reflective assessment positively impacts academic achievement (Bond, 2013; Bond & Ellis, 2013; Evans, 2009; Kramarski & Mevarech, 2003; Lew & Schmidt, 2011; Zan, 2000). Likewise, teacher feedback can positively and significantly impact student learning in terms of quality of homework, interest and motivation which lead to improved learning (Butler & Nisan, 1986; Nunez et al., 2015; Siewert, 2011). However, one could argue that many of these studies suffered from limitations in terms of the research design and data analysis. Additionally, further study is much needed to provide empirical evidence that links both approaches, reflective assessment and feedback, as an effective practice for improved learning. As such, it is vital that educational researchers continue to explore, research and refine the use of metacognitive practice linked with teacher feedback in the learning environment.

References

Anderson, R. C., Kulhavy, R. W., & Andres, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63*, 186-188. http://dx.doi.org/10.1037/h0032653

Arbuckle, J. L. (2009). *Amos 18.0 user's guide*. Crawfordville, FL: Amos Development Corporation.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: W.H. Freeman and Company.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research (61)*, 213-238. http://dx.doi.org/10.3102/00346543061002213

Bianchi, G. A. (2007). *Effects of metacognitive instruction on the academic achievement of students in the secondary sciences* (Doctoral dissertation). Seattle Pacific University, WA.

Black P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*, 7 – 74. http://dx.doi.org/10.1080/0969595980050102

Bond, J. (2003). *The effects of reflective assessment on student achievement* (Doctoral dissertation). Seattle Pacific University, WA.

Bond, J. B., & Ellis, A. K. (2013). The effects of metacognitive reflective assessment on fifth and sixth graders' mathematics achievement. *School Science and Mathematics, 113*, 227–234. http://dx.doi.org/10.1111/ssm.12021

Borich, G. D. (2014). *Effective teaching methods* (8th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Brookhart, S. (2008). *How to give effective feedback to your students.* Alexandria, VA: Association for Supervision and Curriculum Development.

Bruner, J. (1961). *The process of education*. Cambridge, MA: Harvard University Press.

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 78*, 210-216. http://dx.doi.org/10.1037/0022-0663.78.3.210

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin Company.

Cho, H., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research, 66,* 1261-1266. http://doi:10.1016/j.jbusres.2012.02.023

Cisero, C. A. (2006). Does reflective journal writing improve course performance? *College Teaching, 54*, 231–236. http://dx.doi.org/10.3200/CTCH.54.2.231-236

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hilsdale, NJ: Erlbaum.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science (1),* 98–101. http://doi:10.1111/1467-8721.ep10768783

Corrigan, P. W., & Salzer, M. S. (2003). The conflict between random assignment and treatment preference: implications for internal validity. *Evaluation and Program Planning*, *26*, 109-121. http://doi:10.1016/S0149-7189(03)00014-4

Costa, A. L. (2001). *Developing minds: A resource book for teaching thinking* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668. http://dx.doi.org/10.1037/0033-2909.125.6.627

Dewey, J. (1910). *How we think.* Boston, MA: D.C. Heath & Co. publishers

Ellis, A. (2001). *Teaching, learning, and assessment together: The reflective classroom*. Larchmont, NY: Eye on Education, Inc.

Ellis, A., Denton, D., & Bond, J. (2013). An analysis of research on metacognitive teaching strategies. *Social and Behavioral Sciences, 116*, 4015-4024. http://dx.doi.org/10.1016/j.sbspro.2014.01.883

Ellis, A. K., & Denton, D. (2010). *Teaching, learning & assessment together: Reflective assessments for middle and high school mathematics & science*. Larchmont, NY: Eye on Education.

Evans, L. (2009). *Reflective assessment and student achievement in high school English* (Doctoral dissertation). Seattle Pacific University, WA.

Faul, F., Erdfelder, E., Buchner, A., & Lang A. (2007). Statistical power analysis using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Thousand Oaks: CA: SAGE Publication Ltd.

Flavell, J. H. (1977). *Cognitive development.* Englewood Cliffs, NJ: Prentice-Hall.

Flavell, J. H. (1979). Metacognition and cognitive monitor: A new area of cognitive development inquiry. *American Psychologist, 34,* 906 – 911. http://dx.doi.org/10.1037/0003-066X.34.10.906

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: An introduction.* Boston, MA: A & B Publications.

Guskey, T. R., & Marzano, R. J. (2003). *Assessment as learning: Using classroom assessment to maximize student learning.* Thousand Oaks, CA: Corwin Press, Inc.

Guth, G. J. A., Holtzman, D. J., Schneider, S. A., Carlos, L., Smith, J. R., Hayward, G. C., & Calvo, N. (1999). Impact of standards-based accountability systems. In *Evaluation of California's standards based accountability system* (chapter 10). Retrieved from http://www.wested.org/online_pubs/accountability/SBA_chapter_10.pdf

Hall, C., Kennedy, B. J., Bass, D. L., & Wiggins, M. (2012). *Geometry Common Core.* Boston, MA: Pearson Prentice Hall.

Hamre, B. K., & Pianta, R. (2010). Classroom environments and developmental processes: Conceptualization, measurement, & improvement. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling and human development* (pp. 25-41). New York, NY: Routledge.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning.* London, UK: Routledge.

Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249-271). New York, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. http://dx.doi.org/10.3102/003465430298487

Irons, A. (2008). *Enhancing learning through formative assessment and feedback.* New York, NY: Routledge.

Jones, J. L., Jones, K. A., & Vermette, P. J. (2009). Using social and emotional learning to foster academic achievement in secondary mathematics. *American Secondary Education, 37*(3), 4–9.

Joyce, B., Weil, M., & Calhoun, E. (2009). *Models of teaching* (8th ed.). Boston, MA: Pearson Education, Inc.

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,* 254-284. http://dx.doi.org/10.1037/0033-2909.119.2.254

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal, 40*, 281-310. http://dx.doi.org/10.3102/00028312040001281

Kulhavy, R. W., & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. In J. Dempsey & G. Ales (Eds.), *Interactive instruction and feedback* (pp. 3–20). Englewood Cliffs, NJ: Educational Technology Publications.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, *58*, 79–97. http://dx.doi.org/10.3102/00346543058001079

Leung, D. P., & Kember, D. (2003). The relationship between approaches to learning and reflection upon practice. *Educational Psychology: An International Journal of Experimental Educational Psychology, 23(1),* 61-71.

Lew, M. D. N., & Schmidt, H. G. (2011). Self-reflection and academic performance: Is there a relationship? *Advance in Health Science Education, 16*, 529–545. http://dx.doi.org/10.1007/s10459-011-9298-z

Marzano, R. J., Boogren, T., Heflebower, T., Kanold-McIntyre, J., & Pickering, D. (2012). *Becoming a reflective teacher.* Bloomington, IN: Marzano Research Laboratory.

McCombs, B. (2010). Learner-centered practices: Providing the context for positive learner development, motivation, and achievement. In J. L. Meece & J. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 60–74). New York, NY: Routledge.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking.* New York, NY: Routledge.

Mevarech, Z. R., & Kramarski, B. (1997). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal, 34,* 365–394. http://dx.doi.org/10.3102/00028312034002365

Nunez, J. C., Suarez, N., Cerezo, R., Gonzalez-Pienda, J. A., Rosario, P., Mourao, R., & Valle, A. (2013). Homework and academic achievement across Spanish compulsory education. *Educational Psychology, 35,* 726-746. http://dx.doi.org/10.1080/01443410.2013.817537

Nunez, J. C., Suarez, N., Rosario, P., Vallejo, G., Cerezo, R., & Valle, A. (2015). Teacher's feedback on homework, homework-related behaviors, and academic achievement. *The Journal of Educational Research, 108,* 204-216. http://dx.doi.org/10.1080/00220671.2013.878298

Pellegrino, J. W. (2010). *The design of an assessment system for the race to the top: A learning sciences perspective on issues of growth and measurement*. Paper presented at the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, Chicago, IL.

Piaget, J. (1976). *The grasp of consciousness.* Cambridge, MA: Harvard University Press.

Plato. (1952). Theaetetus. In B. Jowett (Ed.), *Dialogues of Plato* (pp. 150– 151). Chicago, IL: University of Chicago Press

Popham, W. J. (2014). *Classroom assessment: What teachers need to know* (7th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). London, UK: Routledge.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice (5)*, 77 – 84. http://dx.doi.org/ 10.1080/0969595980050104

Schoenfeld, A. H. (1987). *Cognitive science and mathematics education.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Schön, D. A. (1987). *Educating the reflective practitioner.* San Francisco, CA: Jossey-Bass Inc.

Siewert, L. (2011). The effects of written teacher feedback on the academic achievement of fifth-grade students with learning challenges. *Preventing School Failure, 55,* 17-27. http://dx.doi.org/10.1080/10459880903286771

Stiggins, R. J. (1996). *Student-centered classroom assessment.* Upper Saddle River, NJ: Merrill Prentice Hall.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics.* Upper Saddle River, NJ: Pearson Education Inc.

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

Walberg, H. J., Paschal, R. A., & Weinstein, T. (1985). Homework's powerful effects on learning. *Educational Leadership, 42*(7), 76–79.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* San Francisco, CA: Jossey – Bass Publishers

Wiliam, D. (2012). Feedback: Part of a system. *Educational Leadership, 70*, 31-34.

Xu, J. (2011). Homework completion at the secondary school level: A multilevel analysis. *The Journal of Educational Research*, *104*, 171–182. http://dx.doi.org/10.1080/00220671003636752

Zan, R. (2000). A metacognitive intervention in mathematics at university level. *International Journal of Mathematical Education in Science and Technology, 31*(1), 143–150. http://dx.doi.org/10.1080/002073900287462

Appendix A

Student Assent

Engaging the culture, changing the world®

Seattle Pacific UNIVERSITY

INFORMED ASSENT

Title of the Study: Journals and Learning

Investigators: Nalline Baliram (561) 707 8823 baliramn@spu.edu
Arthur K. Ellis (206) 235 0816 aellis@spu.edu

DESCRIPTION OF THE RESEARCH
The purpose of the research is to examine the impact of reflective thinking practices and teacher feedback on learning.

This study will include everyone in your math class.

WHAT WILL MY PARTICIPATION INVOLVE?

Everyone will be participating in learning mathematics. Some classes will be doing problems at the start of class and some classes will be doing journaling at the end of class. Everyone will take a few tests on how well they are learning the content.

Once this unit is done, your teacher will remove everyone's names from all the tests, including yours. Then will send just the scores of the tests to the people whose names are on this form above (Nalline and Arthur). Nalline and Arthur will then use the scores on the tests to better understand how everyone learns math.

ARE THERE ANY BENEFITS TO ME?

Everyone will be learning mathematics, and we hope that by using your test scores teachers will know more about how students learn.

HOW WILL MY CONFIDENTIALITY BE PROTECTED?

Your name will not be attached to the test scores when your scores go to Nalline and Arthur.

WHOM SHOULD I CONTACT IF I HAVE QUESTIONS?

You may ask any questions about the research at any time. If you have questions about the research after you leave today you should contact Nalline Baliram, (561) 707 8823.

By signing your name here, you are allowing your teacher to send your anonymous test scores to Nalline. Thank you!

Participant's Name (please print):______________________________

Participant's Signature:________________________________ Date:_____________

Appendix B

Parent Consent

Engaging the culture, changing the world®
Seattle Pacific UNIVERSITY

INFORMED CONSENT

The Impact of Metacognition and Teacher Feedback on Academic Achievement

Investigators: Nalline Baliram (561) 707 8823 baliramn@spu.edu

Faculty Sponsor: Arthur K. Ellis (206) 235 0816 aellis@spu.edu

---

PURPOSE

Your child is invited to take part in a research study taking place during normal class time. The purpose of this study is to examine the impact of reflective thinking practices linked with feedback on academic achievement. All students in the mathematics class will be doing regular learning activities.

PROCEDURES

At the start of the new mathematics unit, students in all sections will take a typical pre-test on essential course content to find out what they already know about the material. While students in all sections will be learning the same unit of material, students in some sections will be engaged in reflective journaling on what they feel they've learned each day, and students in the other two sections will be engaged in doing practice problems. At the end of the unit (about 4 weeks), all students will be administered a typical post-test on the course material. Finally, 4 weeks after the unit has been completed, all students will be administered a test to assess how well they remember this first unit.

While all students will be participating in the learning, those students who want can have their de-identified (anonymous) test scores sent to the researchers (Nalline and Art) to compare learning between the different sections. At no time will Nalline or Art have access to your child's name or course work (like the journal or assignments).

BENEFITS

Since all students are doing work, there are no additional direct benefits to your family for participating.

However, there are several potential benefits to your child:

- S/he may gain a greater understanding of how s/he approaches learning
- S/he may learn how to self-regulate and adjust personal learning strategies

There are also several potential benefits for your child's math teacher, and teachers generally:

- They will learn which of the two strategies aids student learning the best
- They gain greater awareness of the role of reflection in student learning and can perhaps facilitate future professional development for other teachers.
- They gain a consistent, department-wide approach to formally utilize reflective assessment and improve on feedback in the classroom on an ongoing basis.

PARTICIPATION AND ALTERNATIVES TO PARTICIPATION

There are no alternatives to engaging in the classroom activities measured in the study, as the students will be engaged in the typical curriculum as prescribed by the School District and required per the guidelines for High School Graduation. If your student or you decide to be a 'NON-participant', your student will remain engaged in the regular classroom activities; however his/her data will not be collected for the study.

CONFIDENTIALITY

__The information in the study records will be anonymous__. All test scores will be compiled by the teacher into non-identifiable data. The coded, anonymous data will be stored securely and will be made available only to Nalline and Art (whose contact information is above). There will be no way to link any work or score to either you or your child.

SUBJECT RIGHTS

If you have questions at any time about the study or the procedures, you may contact the Principal Investigator, Nalline Baliram, at Seattle Pacific University, and (561) 707 8823. If you have questions about your rights as a participant, contact the SPU Institutional Review Board Chair at 206-281-2201 or IRB@SPU.edu .

CONSENT

Your signature on this form indicates that you have understood to your satisfaction the information regarding your child's participation in this research project and agree to have him/her participate in this study. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities.

I have read the above information and agree to have my child participate in this study. I have received a copy of this form.

Parent/Guardian's name (print) ________________________________	**Researcher's name (print)** ________________________________
Parent/Guardian's signature ________________________________	**Researcher's signature** ________________________________
Date _____________	**Date** _____________

Copies to: Participant Principal Investigator

Appendix C

Daily Reflection Notecard

Front of note card

Student #:	Period:	Date:
1. **Today I learned** *(Write down at least two things you learned in class today)*		
2. **I can now** …… *(Write down at least two concepts you feel comfortable with)*		
3. **I still don't understand**.. *(Reflect on areas you still need help with after today's lesson)*		

Back of note card

Practice Problem #1
Practice Problem #2

Appendix D

End of Study Survey

End of Study Reflective Questions:

Please complete the following prompt by circling the best choice that applies to you.

1. I enjoyed studying Chapter 8: Pythagorean Theorem and Special Triangles

1	2	3	4	5
strongly disagree	dis-agree	neutral	agree	strongly agree

2. I was given an opportunity to reflect on my learning and express when I was struggling on the content.

1	2	3	4	5
strongly disagree	dis-agree	neutral	agree	strongly agree

3. I was provided with helpful feedback on my reflection on learning.

1	2	3	4	5
strongly disagree	dis-agree	neutral	agree	strongly agree

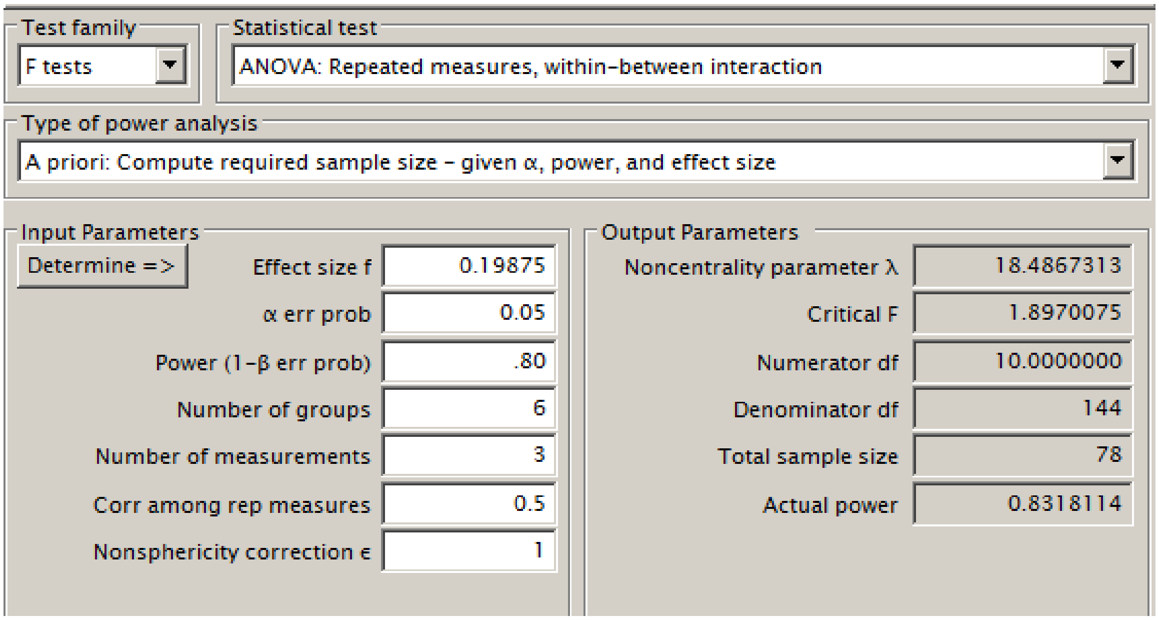
4. I felt prepared for this unit test.

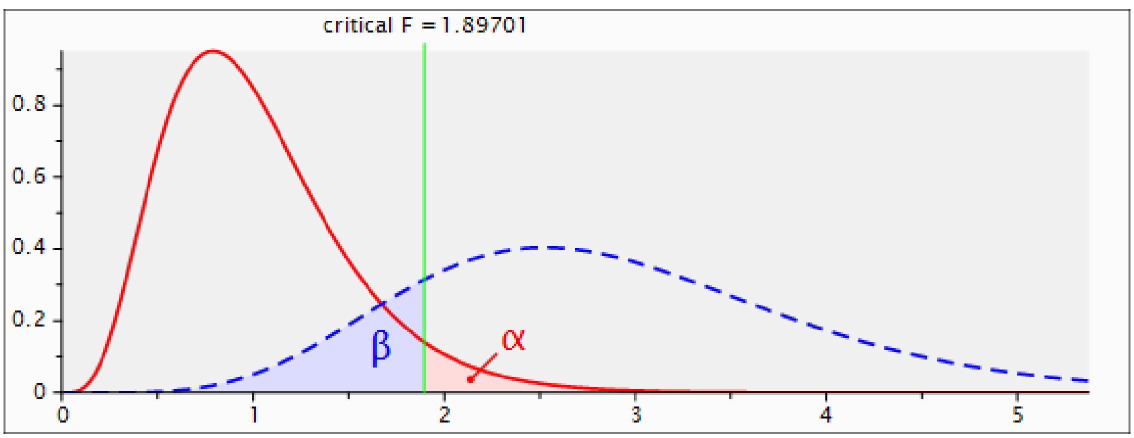
1	2	3	4	5
strongly disagree	dis-agree	neutral	agree	strongly agree

5. What did you like and did not like about the exit ticket you completed at the end of each class period?

Appendix E

G*Power 3 Output





Appendix F

Responses to the Open Ended Prompt Administered to the Experimental Group

	What did you like and did not like about the exit ticket you completed at the end of each class period?
1	I didn't like that it was more work to do, but I liked the feedback it gave me.
2	I didn't like it because it held me back from lunch and made us do extra work.
3	I liked doing the practice questions. I did not like saying two things I learned.
4	I liked how we were able to express our concerns and triumphs with the lesson. I did not like waiting for a response to my questions.
5	Help me see if I didn't understand anything, got too repetitive.
6	I like because I learn more new things.
7	I liked being able to reflect on what we learned at the end of class.
8	We didn't get them back soon enough so the notes were basically useless because I didn't have them to help me study for quizzes and the test.
9	I liked that it gave me a chance to practice what I learned in class. I did not like the fact that it was extra work I had to do at the end of class.
10	I like how it made you think about what you just learned.
11	I like how we received help for things we didn't understand. I don't think I didn't like anything.
12	I like how I was given the opportunity to reflect on my work after each lesson. I did not like how I felt rushed to complete my reflections at the end of each lesson.
13	I liked being able to apply what I learned. It helped me to realize what I fully understood and what I didn't. The feedback that I was given back was very helpful and I used it to study.
14	I liked the feedback but didn't like that it took the time out of class.
15	I liked the problems but it was difficult to put into words what I was having trouble with.
16	I liked how I felt I could ask questions more freely. I didn't mind the cards and the question. The only negative thing is that the time used for the card could have been used for class.
17	I think they help me because at the end of each period it summed it all up onto the card and gave me some examples.
18	Like review the things I've been learned.
19	So so, because sometimes we don't have enough time to finish it.
20	I did not like how sometimes I would not have enough time to finish the ticket but other than that I didn't mind it.
21	I really like getting a second way of teaching through the note cards. It was useful to see it on a notecard right next to our work. I can't say I disliked anything from this unit.

22	I liked that it accomplished two tasks at once, helping me gather my thoughts at the end of class while also wasting, I mean "using wisely' some time in class. I really don't have any negatives about the cards.
23	I like to do the last two problems.
24	I did not like how it took away from homework completion time and I liked the encouraging and useful feedback.
25	I like how it gave me an extra challenge, but it was tedious to try to get it done before the end of class.
26	I liked it because it actually made me understand the math we learned better.
27	I liked being able to evaluate how well I knew and learned the lesson. I did not like that the cards almost always came back after I learned what it was I was struggling with. Overall I really enjoyed them.