January 1st, 2003

The Effects of Reflective Assessment on Student Achievement

John B. Bond
Seattle Pacific University

The Effects of Reflective Assessment
on
Student Achievement

John B. Bond
Seattle Pacific University

The Effects of Reflective Assessment

on

Student Achievement

by

John B. Bond

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor in Education

Seattle Pacific University

2003

Approved by ____
(Dr. Arthur Ellis, Ed.D., Chairman of the Dissertation Committee)

____
(Dr. Rick Eigenbrood, Ph.D., Committee Member)

____
(Dr. Lisa Bjork, Ed.D., Committee Member)

Program Authorized to Offer Degree ________ School of Education ____________

Date ____

____
(Dr. William Rowley, Ed.D., Dean, School of Education)

Doctoral Dissertation

Signature ____

Date ____

Seattle Pacific University

Abstract

The Effects of Reflective Assessment

on

Student Achievement

by

John B. Bond

Chairperson of the Dissertation Committee: Dr. Arthur Ellis, School of Education

Over the last decade alternative assessment strategies have become an important part of the debate regarding the reform and restructuring of American education. The purpose of assessment should be to improve student learning, which means it should be integral to the teaching and learning process. For this to occur, a seamlessness needs to exist between teaching, learning, and assessment through which students are empowered to take increased responsibility for their learning. Reflective assessment grows out of strong theoretical roots including ancient Greek thought, the philosophy of John Dewey, and cognitive constructivist learning theories. Reflective assessment is a formative process through which students can experience assessment as a part of learning, rather than as a separate evaluative process.

This study was designed to investigate the effect of a metacognitive strategy on the mathematics achievement of fifth and sixth grade students. This was an experimental

study that employed a posttest-only control group design. Students were randomly assigned to three treatment groups, two of which received identical mathematics curriculum, except for a reflective intervention with one group. It was the reflective assessment intervention that served as the independent variable in the study. The third group served as the control group, and received instruction in an alternate mathematics curriculum.

An instrument was developed that was aligned with the mathematics lessons that were taught to the two experimental groups. This measure served as the dependent variable in the study. Descriptive statistics were used to describe the demographic nature of the sample. Analysis of variance and nonparametric procedures were used to analyze the data and to make inferences from it. These inferential statistics were also used to analyze the results of a retention test that was administered six weeks following the end of the study. A performance test was also administered at the end of the study to measure practical application of the concepts taught in the mathematics lessons.

The data gathered revealed a statistically significant difference between the achievement of students who practiced the reflective strategy and the achievement of students in both the comparison and control groups. Students who practiced the reflective strategy also scored significantly higher on the retention test, which was a re-administration of the posttest. On the performance assessment, which was a subjective measure, students in the reflective strategy group were more successful applying the concepts taught in the mathematics lessons than were students in the other two groups.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

This dissertation is dedicated to my wife, Ginny, who has provided unwavering patience, support, encouragement, and love throughout the entire course of this project. I am thankful for my beloved children, Jessica, Michael, and Bridget, for their love, understanding, and support as I progressed on this scholarly journey.

I would like to thank my committee chair, Dr. Arthur Ellis, for his wise counsel, consistent guidance, and timely encouragement during this project and throughout my doctoral studies. In addition, I would like to thank Dr. Rick Eigenbrood and Dr. Lisa Bjork for their excellent support as members of my dissertation committee.

Finally, I would like to acknowledge and thank Dr. David Steele, Dr. Richard Smith, and Dr. Chris Sink for their support, encouragement, and inspiration over the course of my doctoral program.

Seattle Pacific University

Abstract

The Effects of Reflective Assessment

on

Student Achievement

by

John B. Bond

Chairperson of the Dissertation Committee: Dr. Arthur Ellis, School of Education

Over the last decade alternative assessment strategies have become an important part of the debate regarding the reform and restructuring of American education. The purpose of assessment should be to improve student learning, which means it should be integral to the teaching and learning process. For this to occur, a seamlessness needs to exist between teaching, learning, and assessment through which students are empowered to take increased responsibility for their learning. Reflective assessment grows out of strong theoretical roots including ancient Greek thought, the philosophy of John Dewey, and cognitive constructivist learning theories. Reflective assessment is a formative process through which students can experience assessment as a part of learning, rather than as a separate evaluative process.

This study was designed to investigate the effect of a metacognitive strategy on the mathematics achievement of fifth and sixth grade students. This was an experimental

study that employed a posttest-only control group design. Students were randomly assigned to three treatment groups, two of which received identical mathematics curriculum, except for a reflective intervention with one group. It was the reflective assessment intervention that served as the independent variable in the study. The third group served as the control group, and received instruction in an alternate mathematics curriculum.

An instrument was developed that was aligned with the mathematics lessons that were taught to the two experimental groups. This measure served as the dependent variable in the study. Descriptive statistics were used to describe the demographic nature of the sample. Analysis of variance and nonparametric procedures were used to analyze the data and to make inferences from it. These inferential statistics were also used to analyze the results of a retention test that was administered six weeks following the end of the study. A performance test was also administered at the end of the study to measure practical application of the concepts taught in the mathematics lessons.

The data gathered revealed a statistically significant difference between the achievement of students who practiced the reflective strategy and the achievement of students in both the comparison and control groups. Students who practiced the reflective strategy also scored significantly higher on the retention test, which was a re-administration of the posttest. On the performance assessment, which was a subjective measure, students in the reflective strategy group were more successful applying the concepts taught in the mathematics lessons than were students in the other two groups.

Chapter One

Introduction

Purpose of the Study

The central purpose of this study is to determine the effects of reflective assessment strategies on the mathematics achievement of fifth and sixth grade students. The reflective assessment strategies will be based on those articulated by Ellis (2001) and Wilson (1986), which are influenced by cognitive-constructivist theories of learning.

In addition, a practical aim of this study is to provide classroom teachers with a rationale in support of integrating metacognitive and reflective components into lessons. The review of the literature and the present investigation are designed to provide insights concerning the effects of reflective practice by elementary age students in daily learning activities.

Background

Over the last decade alternative assessment strategies have become an important part of the debate regarding the reform and restructuring of American education. In contrast to the more familiar focus upon standardized testing, this discussion and inquiry has explored the informal, ongoing, formative assessments that might occur in classrooms. Particularly, it is how such formative assessment is seamlessly integrated into the teaching and learning experience that has sparked increased attention to this area (Ellis, 2001; Wragg, 1997). Unfortunately, the role of student reflection in a seamless teaching, learning, and assessment process has received only slight empirical attention in this ongoing discussion.

The purpose of assessment has become confused with the overemphasis that is placed on standardized tests (Stiggins, 1997). Rather than integral to teaching and learning, assessment is often viewed as a separate process (Herman, Aschbacher, & Winters, 1992). Such significant matters as problem solving and complex thinking are often neglected in favor of specific and routine skills that are learned out of context (Wiggins, 1993). As a result, the seamlessness between teaching, learning, and assessment that should occur naturally in the classroom (Ellis, 2001; Perrone, 1994; Simmons, 1994; Wragg, 1997) is not at the center of the learning process.

Michael Scriven (1967) is credited with first using the term *formative* to describe evaluation that is intended to assess the effectiveness of new curricula (Bloom, Hastings, & Maudaus, 1971; Brookhart, 2001). Still writing in the field twenty-five years later, Scriven (1991) expanded his definition of formative assessment to include evaluation designed, done, and delivered to someone who can make improvements. It is the "someone who can make improvements" Scriven (1991) refers to that is at the crux of the assessment dilemma. When the student is included, along with the teacher, as someone who can make such improvements, the potential exists to transform the classroom experience. Students need to become the ultimate users of classroom assessment information that is elicited to improve learning (Black & Wiliam, 1998; Stiggins, 1997). Of the three reasons offered by Ellis (2001) for assessing teaching and learning—to classify students, to diagnose students, and to encourage and support student learning—it is the last one that captures the rationale for involving students in the assessment process. Effective assessment practices offer students an opportunity to reflect upon what they are learning (Earl & LeMahieu, 1997). It is through this reflection that students experience assessment as a part of learning, rather than as a separate evaluative process (Earl & LeMahieu, 1997; Wiggins, 1993). Assessment should begin to answer two fundamental questions: How are we doing? and How can we do it better? (Herman et al., 1992). The "we" that Herman and colleagues refer to should be a partnership that includes the learner.

Metacognition, or reflective practice, is a concept of cognitive psychology which "focuses on the active participation of the individual in his or her thinking process" (Stewart & Landine, 1995). It involves thinking about thinking and emphasizes the individual's self-awareness of his or her thinking patterns, learning characteristics and techniques which aid both memory and comprehension (Schneider, Borkowski, Kurtz, & Kerwin, 1986). Costa (2001) uses the terms "metacognition" and "thinking about thinking" interchangeably to mean the ability to know what we know and what we don't know. Metacognition refers to self-knowledge about cognitive states and processes (Flavell, 1976). Others define reflection in similar terms as "an important human activity in which people recapture their experience, think about it, mull it over, and evaluate it" (Boud, Keogh, & Walker, 1985). Costa & Kallick (2000, p. 27) provide further definition of terms in the following statement:

> Intelligent people plan for, reflect upon, and evaluate the quality of their own thinking skills and strategies. Metacognition means becoming increasingly aware of one's actions and the effect of those actions on others and on the environment.

Research Questions

The following research questions drive this inquiry:

1. Is there a difference in the mathematics achievement of elementary students who receive statistics instruction and practice reflective assessment techniques, as compared to students who are taught the same curriculum but do not practice reflective assessment techniques?
2. Is there a difference between the mathematics achievement of elementary students who receive statistics instruction and practice reflective assessment techniques, as compared to students who receive instruction in a geometry curriculum?
3. Is there difference between the mathematics achievement of students who receive statistics instruction without reflective assessment techniques, as compared to students who receive instruction in a geometry curriculum?

Significance of the Study

This study will add to the limited body of research in the area of metacognitive practice by elementary-age students and its effects on academic achievement. While there is an increasing amount of attention to reflective practice in the research literature, few studies have focused on measuring the effects of student reflective strategies when integrated into lesson activities. In order to build upon the existing body of research, this study will examine the effects of reflective assessment strategies as defined by Ellis (2001) and Wilson (1986) on the mathematics achievement of fifth and sixth grade students.

Although a number of research studies (Naglieri & Johnson, 2000; Scheider et al., 1986; Schunk, 1983; Wang, Haertel, & Walberg, 1993) have reported positive effects of metacognitive activities on student achievement, these are typically embedded as only a component of such studies. While the literature is rich with philosophy and opinion regarding the value of reflective practice for students, there has been very little research designed specifically to measure such effects. This study will help to fill this research gap by using an experimental design to isolate the effects of reflection in the elementary classroom setting.

Chapter Two

Review of Literature

For more that one hundred years the study of how people know what they know has been of interest to cognitive psychologists and educational researchers. While focus on notions akin to metacognition far precede John Flavell's introduction of the term *metacognition* three decades ago, since that time scholarly work on the topic greatly increased. An indication of this growing interest is the more than 500 journal articles and 169 book chapters that were written on topics concerning metacognition between 1979 and 1995 (Hacker, 1998). In addition to continued dialogue regarding the theoretical aspects of metacognition, publications over the last decade have shifted greater attention to both empirical investigations and applications to educational practice.

Metacognition has accumulated a wide range of definitions and interpretations (Manning & Payne, 1996). The term metacognition came into vogue about 1975, according to Ann Brown (1978), who has been a prolific writer in the area over the last twenty years. It has been referred to as mysterious, fuzzy, faddish, and a "many-headed monster" (Brown, 1978, p. 105). The focus of this study, however, is on the application end of the spectrum, rather than near the theoretical pole. While the following review of theories, definitions, and issues that surround metacognition is essential in order to place reflective assessment appropriately within the metacognitive construct, the actual study described in chapters three, four, and five is focused upon the extent to which the application of a metacognitive strategy effects student learning. For this reason, the research that is reviewed in this chapter is limited to studies that deal with metacognitive interventions and their relation to achievement.

With awareness of the broad construct that includes metacognition, reflection, and related areas of cognitive study, the purpose of this literature review is to accomplish four goals: (a) explicate the theoretical bases underlying metacognition; (b) describe the composition of metacognition as a construct; (c) identify issues related to application of reflective assessment concepts in the classroom; and (d) summarize specific research related to reflective interventions and their relationship to student learning.

Theoretical Underpinnings

The roots of metacognition, and for the purposes of this paper, reflective assessment, are found in the constructivist theories of John Dewey (1933) and Jean Piaget (Piaget, 1976; Gredler, 2001). Thinking about thinking is firmly based in the cognitive theories of learning that have been developed over the last fifty years, as opposed to a behaviorist view that is often seen in the school setting (Shepard, 1991). Metacognition is also rooted in ancient Greek thought, as seen in Plato's *Theatetus*, where Socrates described thinking as "a discourse the mind carries on with itself" and judgement as "a statement pronounced…silently to oneself" (Plato, trans. 1956).

Contemporary cognitive psychology holds that learning is an ongoing process during which learners are continually receiving information, interpreting it, and connecting it to what they already know (Herman et al., 1992). Within a constructivist view of learning, knowledge is assumed to be personally and socially mediated (Vygotsky, 1978), rather than accumulated in a linear, bit-by-bit fashion (Soodak & Martin-Kniep, 1994). As Piaget (1976) theorized, learners' interactions lead to structural changes in how they think about something as they assimilate and accommodate new information. Knowledge is organized in mental models, knowledge structures, or schema

(Herman, et al., 1992). Consistent with this cognitive, constructivist perspective, Tittle (1994) supported the view of teachers and students both using assessment information:

> A cognitive constructivist perspective … suggests that teachers and learners construct schemas or integrate representations from assessments into existing views of the self, of teaching and learning, and of the curriculum, broadly construed. These interpretations include knowledge and beliefs and may also result in intents to use and actual use of assessments. (p. 151)

Cognitive learning theory and its constructivist approach to knowledge acquisition support the need to integrate assessment methodologies with instructional outcomes and curriculum content (Herman, et al., 1992). The notion that assessment is part of learning is deeply rooted in a constructivist theory that learning is a process of taking in information, interpreting it, connecting it to existing knowledge or beliefs, and, if necessary, reorganizing understanding to accommodate that information (Shepard, 1991).

Albert Bandura (1997) theorized that reflective processes mediate the information about perceived efficacy that learners acquire from different sources of their learning experience. Self-efficacy refers to beliefs about one's capabilities and these beliefs also motivate learners in particular ways (Gredler, 2001). While an in-depth analysis of self-efficacy theory is beyond the scope of this paper, the relationship between reflection, self-efficacy, and achievement are related to the topic of reflective assessment and need to be acknowledged. A subsection on metacognitive beliefs is included in this chapter.

Bernard Weiner's (1980) attribution theory also has relevance to the concept of reflective assessment. Weiner (1980, 1986) theorized that academic performance is influenced by an individual's perceptions regarding the reasons for their achievement. "Attribution theorists assume that individuals utilize a number of ascriptions both to postdict (interpret) and to predict the outcome of an achievement-related event" (Weiner, 1980, p. 328). It is the postdicting and predicting referred to by Weiner that applies to this focus on metacognition, for to do so a student must engage in reflection. A thorough review of the work of Weiner is far beyond the scope of this dissertation, however, it would be a noticeable omission to not acknowledge its place in the theories underlying metacognition. As with self-efficacy theory, attribution theory has relevance to the metacognitive beliefs held by the learner, which will be reviewed later in this chapter.

John Dewey (1933) considered reflection central to all learning experiences, enabling "us to act in a deliberate and intentional fashion" (p. 212). In discussing why reflective thinking must be an educational aim, Dewey (1933) stated that the act of reflecting upon the consequences of actions ahead of time "enables us to know what we are about when we act. It converts action that is merely appetitive, blind, and impulsive into intelligent action" (p. 17). Dewey (1933) outlined five phases of pre-reflective to post-reflective thought that closely resemble the scientific method:

1. Suggestion, in which the mind leaps forward to a possible solution;
2. Intellectualization, in which a problem is converted from an emotional quality to an intellectual process;
3. Hypothesis, in which ideas that "pop into the mind" are used to guide observation and other operations in collection of factual material;
4. Reasoning, in which ideas are mentally elaborated upon, pondered, and synthesized;

5. Testing the hypothesis, in which an idea is tested and the consequences of action are evaluated (pp. 106-115).

Dewey emphasized that in practice the five phases of reflection might happen in a different order and that some might be passed over, while others might be expanded upon.

The importance of reflective thinking has also been written about by philosopher Jürgen Habermas (1996; Young, 1990). Habermas, who is a critical theorist, explained that the goal of education should be the emancipation of students' thinking in order to empower them to take control of their lives (Ornstein & Hunkins, 1998). According to Habermas (1996), "Reflection coincides with a step forward in the progress toward the autonomy of the individual, with the elimination of suffering and the furnishing of concrete happiness" (pp. 77-78). Ornstein and Hunkins (1998), placed Habermas in the same broad learner-centered camp that includes Dewey. Habermas, however, is included in the "romantic or radical" subgroup of learner-centered designs, while Dewey is included in the "child-centered" subgroup (Ornstein & Hunkins, 1998).

Ernst von Glasersfeld (2000) included student reflection as being of high importance in an effective classroom. Von Glasersfeld (1991), a radical constructivist (Moshman, 1982; Phillips, 1995), argued that reflection in the classroom is essential for students to construct new learning. According to von Glasersfeld (1991), "Leading students to discuss their view of a problem and their own tentative approaches, raises their self-confidence and provides opportunities for them to reflect and to devise new and perhaps more viable conceptual strategies" (p. xix). Von Glasersfeld (2000) placed strong emphasis on the responsibility of teachers to create opportunities for reflective thinking, to monitor the appropriate development of student learning, and to intervene to avoid inappropriate constructions of knowledge. Von Glasersfeld's views regarding the role of the classroom teacher in closely monitoring and responding to student reflective activities support the ongoing, informal classroom assessment that is at the center of current assessment reform conversation.

"Assessment reform, for those who hold this view, is not connected to compliance with mandates but is rooted in the constructivist view that learning depends on self-monitoring and reflection" (Earl & LeMahieu, 1997, p. 158-159). It is ironic that the word *assessment* is derived from the Latin word *assidere*, which means to sit with (Wiggins, 1993). In contrast to images of bubble-in standardized tests and percentile ranks, the phrase "to sit with" conjures up images of teachers observing, discussing, and working with students. Earl & LeMahieu (1997) stated that assessment is not only part of learning, but that it is the critical component that allows learners and their teachers to check their understanding against the views of others and the collective wisdom of the culture. John Dewey advocated placing the learner at the center of experiences, and defined the teacher as the learner's "co-partner and guide in a common enterprise—the child's education as an independent learner and thinker" (Dewey, 1964, p. 10).

What is Metacognition?

John Flavell (1976) introduced the concept of *metacognition* to the literature and defined it as "one's knowledge concerning one's own cognitive processes and products or anything related to them" (p. 232). According to Flavell (1976), metacognition refers to the active monitoring of these processes in relation to a concrete goal or objective. Stewart and Landine (1995) stated that metacognition is a branch of cognitive psychology, which focuses on the active participation of the learner in his or her thinking process. Others have referred to metacognition as knowledge about how we perceive, remember, think, and act (Jans & Leclercq, 1997), awareness and regulation of cognitive activity (Palincsar & Brown, 1987), thinking about your own thinking (Schoenfeld, 1987; Stewart & Landine, 1995), reflections on cognition (Schoenfeld, 1987), reflective intelligence (Garofalo & Lester, 1985), and self awareness of your own thinking patterns, learning characteristics, and techniques which aid both memory and comprehension (Schneider, Borkowski, Kurtz, & Kerwin, 1986). Metacognition is, simply put, what we know about what we know (Jans & Leclercq, 1997).

While there has been a great deal written about metacognition over the last thirty years, there is not agreement among scholars regarding the components of the construct (Allen & Armour-Thomas, 1992; Desoete, Roeyers, & Buysse, 2001). Schoenfeld (1987) stated that there are three related, but distinct, categories of intellectual behavior. These are one's knowledge about one's own thought processes, control or self-regulation, and beliefs and intuitions. Schoenfeld (1987) described the first component as *metamemory* or people's ability to describe how good they are at remembering things. The second component, control or self-regulation, has to do with how well you keep track of what you are doing when you are solving problems, and how well you use the input from those observations to guide your problem-solving actions (Schoenfeld, 1987). This component has received much attention in the literature and will be reviewed in greater detail later in this paper. Beliefs and intuitions, the third component, has to do with the ideas, preconceptions, and feelings that a person brings to a learning experience.

In a competing view, Garofalo and Lester (1985) proposed that there are two separate, but related aspects to metacognition, rather than three. These are knowledge and beliefs about cognitive phenomena and the regulation and control of cognitive actions, both of which overlap with the three-component models components. Since it is not always easy to distinguish between what is metacognitive and what is cognitive, confusion often results when communicating about metacognition (Garofalo & Lester, 1985)). In an effort to provide greater clarity to the metacognition conversation, Garofalo and Lester define *cognitive* as being involved in doing and *metacognitive* as being involved in choosing and planning what to do and monitoring what is being done. Other scholars (Biggs, 1988; Nelson & Narens, 1990; Schoenfeld, 1987) further defined the terms making communication about metacognition more clear and efficient.

Nelson and Narens (1990) described cognitive processes as being divided between two interrelated levels—meta-level and object-level. Meta-level contains a dynamic model or mental simulation of the object-level, which is the *doing* that Garofalo and Lester (1985) ascribed to the term cognition. According to Nelson and Narens (1990), there are two dominance dynamics involved in metacognitive thought, “control” and “monitoring,” which are defined in terms of the flow of information between the meta-level and the object-level. *Control* refers to the meta-level modifying the object-level. For example, when a person is driving too fast he or she thinks about the act of speeding and makes a decision to either slow down or to continue at the same rate. Control takes place when the person decides to slow down or change his or her behavior, which is meta-level thinking modifying object-level behavior. On the other hand, if the person chose to continue driving too fast, it would be an example of *monitoring*.

Monitoring takes place when the meta-level is informed by the object-level, but does not modify the behavior (Nelson & Narens). Similarly, Schoenfeld (1987) explained that this aspect of metacognition is a management issue, which includes understanding the problem, planning how to act, keeping track during a solution, and allocating resources such as strategies and time.

A metalearning model composed of three clusters of factors was developed by John Biggs (1988). While much of Biggs' work restated that of previous scholars, his model brought greater clarity to the construct of metacognition by placing it as a process that is preceded by some factors and followed by others. First, in Biggs' metalearning model is the *presage* cluster, which includes personal factors, situational factors, and tactics. Personal factors refer to such traits, qualities, or experiences that an individual brings to a metacognitive process. Cognitive ability, gender, age, and upbringing are examples of personal factors that could have influence on a process. Situational factors might include the nature of the task, its difficulty, and the time required. Tactics, in Biggs' model, refer to the study skills or techniques that a learner might use to solve a problem.

The presage cluster in Biggs' metalearning model is followed by the *process* cluster, which deals with actual metacognition. In this cluster of factors, process combines with a strategy to produce an approach to learning. According to Biggs (1988), a strategy is a self-regulatory system that is based on self-awareness. Biggs theorized that strategy interacts with motive, which is how a student views the learning context, resulting in the depth of information processing. Biggs described a continuum from Deep Approach, in which students seek to actualize interest and competence, to Surface

Approach, in which students meet the minimum requirements. In between the two extremes is Achieving Approach, in which a student might be motivated by grades or other external rewards.

The third cluster in Biggs' metalearning model is the *product* cluster (Biggs, 1988). This cluster includes the ratio of structure and detail in a learning task, which has impact upon difficulty level and motivation. Affective involvement is also a factor in the product cluster, which has to do with a learner's feelings about the activity or task. As do other writers in the field of metacognition (Bandura, 1986; Borkowski, 1992; Cornoldi, 1998; Larkin & Chabay, 1989; Palincsar & Brown, 1989; Pintrich & DeGroot, 1990; Wiens, 2001; Zimmerman, 1990), Biggs acknowledged in his model the importance of motivation to the learning process.

In addition to the theoretical proposals regarding the components of metacognition as a construct, efforts have been made to define the construct of metacognition using quantitative research methodology. Desoete and colleagues (2001) examined the relationship between metacognition and mathematical problem solving in a study aimed at contributing some data to the debate on whether there are two or three components within metacognition. Typically, the debate has been whether there are two (knowledge and skills) or three (knowledge, skills, and beliefs) components within the metacognition construct. Using principal components analysis, these researchers identified three metacognitive components—global metacognition (metacognitive knowledge), off-line metacognition (prediction, evaluation, and monitoring), and attribution (beliefs) to effort—which explained 66.86% of the common variance. While this investigation identified three components, they were different from the components, knowledge, skills, and beliefs, commonly used in the field. As the authors stated, they failed to validate the traditionally used components of metacognition (knowledge, skills, and beliefs) related to successful mathematical problem solving (Desoete, et al., 2001).

Construct validation of metacognition was also the focus of research by Allen and Armour-Thomas (1992). In a study of 126 high school students the researchers used a self-report measure based on Sternberg's conceptualization of metacognition or metacomponents. These metacomponents, which further define the broad components of knowledge and control that people have over their own cognition, are as follow: (a) Deciding upon the nature of the problem; (b) selecting components or steps needed to solve the problem; (c) selecting a strategy for ordering the components of problem solving; (d) selecting a mental representation for information; (e) allocating resources; and (f) solution monitoring (Allen & Armour-Thomas; Sternberg, 1986). While this study also included research questions that delved into the interdependence of metacognitive factors and the function of context in the selection of metacognitive components, it is the researchers' focus on the validity of metacognitive components that is relevant to this paper. The results of this study supported Sternberg's conceptualization of "metacomponents," and in doing so contributed to the validation of the construct of metacognition.

In summary, over the last three decades the study of metacognition has become a focus of study and debate among researchers in the field. Beginning with Flavell's definition of the term metacognition in the 1970s, the educational and psychological literature has become rich with theory, opinion, and argument regarding what comprises metacognition. While there is typically a great deal of overlap in the conceptual substance of the constructs proposed by different authors, a wide range of difference exists among their theoretical frameworks, number of components included, and vocabulary used to explain metacognition. Empirical studies aimed at identifying the factors that comprise the metacognitive construct have also been conducted (Allen & Armour-Thomas, 1992; Desoete et al., 2001). However, as with metacognitive theorists, researchers to this point are unable to agree upon common definitions and use of terms.

As far as identifying the broad parameters of the construct, though, it appears that there is much general agreement among experts in the field of metacognition. For the purposes of this paper, I have chosen a three-component model of metacognition, which includes *metacognitive knowledge, metacognitive skills,* and *metacognitive beliefs.* In addition, I have included an overview of *self-regulated learning*, which is seen by some authors as within metacognitive skills, and by others as a related construct on its own.

Metacognitive Knowledge

In describing metacognitive knowledge, Palincsar & Brown (1987) used the example of a student who sees the need to prepare differently for an essay exam than for a multiple-choice test. By reflecting upon his or her knowledge about tests, personal experience with such exams, and specific study skills appropriate for the task, the student accesses his or her metacognitive knowledge. According to Flavell's (1979) model of metacognition and cognitive monitoring, "four classes of phenomena affect the ability to control a wide variety of cognitive enterprises" (p.906). Flavell included metacognitive knowledge, metacognitive experiences, goals (or tasks), and actions (or strategies) in a conceptual description of metacognition. It is the first component of this model, metacognitive knowledge that will be summarized in this section.

Metacognitive knowledge refers to one's "stored world knowledge" (Hacker, 1998). Through both conscious memory search and automatic cognitive processes, metacognitive knowledge may influence the cognitive choices a person makes (Cornoldi, 1998; Hacker). Flavell (1979) stated that there is interaction between knowledge, experiences, goals, and actions, which results in a wide range of thought and behavior. Cornoldi, in describing metacognitive knowledge, distinguished between pre-existing metacognitive knowledge and the metacognitive conceptualization that is activated when a person is faced with a cognitive task. Cornoldi argued that metacognitive knowledge cannot be reduced to a "corpus of declarative knowledge" (p. 139).

Metacognitive knowledge consists of "what we have learned through experience about cognitive activities" (White, 1999, p. 38). Garofalo and Lester, (1985) stated that metacognitive knowledge "is concerned with what a person knows about cognitive abilities, processes, and resources in relation to the performance of specific cognitive tasks" (p. 164). Personal, task, and strategy factors are often used to categorize metacognitive knowledge (Flavell & Wellman, 1977; Garofalo & Lester; White). A summary of each of these factors follows:

1. Personal Knowledge— Personal or self knowledge consists of what learners believe about themselves and others (Garofalo & Lester, 1985; White, 1999). This includes how learners view their knowledge of themselves in comparison to others (White). Included in this category is the study of memory prediction, in which learners predict their performance based, in theory, upon their metacognitive knowledge (Garofalo & Lester).

2. Task Knowledge— Task knowledge includes what individuals know about the tasks they are about to undertake (White, 1999). It includes knowledge about "…the scope and requirements of tasks as well as knowledge about the factors and conditions that make some tasks more difficult than others" (Garofalo & Lester, 1985, p. 164). Examples of task knowledge include a student's belief that familiarity with story content enhances reading comprehension and young or poor readers view of reading as decoding activity, rather than a search for meaning (Garofalo & Lester).

3. Strategic Knowledge— Strategic knowledge has to do with the awareness of the strategies available to achieve a cognitive goal. It also includes a knowledge of the usefulness of specific strategies (White, 1999). "The metacognitive aspect of such knowledge lies in knowing where it can be used and in knowing when and how to apply it" (Garofalo & Lester, 1985, p. 165). Decisions about the use of such strategies as mnemonics and retrieval strategies are examples of strategic knowledge. According to Garofalo and Lester, purely rote memory strategies are cognitive strategies, not metacognitive.

Other researchers define metacognitive knowledge as including different components. In studies of mathematical problem solving, metacognitive knowledge has been described to include three components that differ from those described above. According to several researchers (Desoete et al. 2001; Jacobs & Paris, 1987; Montague, 1992) metacognitive knowledge is comprised of *declarative, procedural,* and *conditional* or *strategic* metacognitive knowledge. Declarative metacognitive knowledge is defined as "what is known in a prepositional manner" (Jacobs & Paris, p. 259) and the knowledge of the influencing factors of human thinking (Desoete et al.). Montague, in relation to mathematical problem solving, stated that declarative knowledge includes "quantitative concepts, mathematical operations and algorithms, and specific problem-solving strategies" (p. 230). The second component, procedural metacognitive knowledge, is comprised of the methods of achieving goals and the application of skills (Desoete et al.). Jacobs and Paris referred to procedural metacognitive knowledge as the "awareness of processes of thinking" (p. 259). The final component in this model, *conditional* or *strategic* metacognitive knowledge, has to do with the conditions that influence learning (Jacobs & Paris). According to Montague, conditional knowledge "enables a learner to select appropriate strategies and to adjust behavior to changing task demands" (p. 230). Desoete et al. stated that these metacognitive components may "help children to know how to study a new timetable (procedural knowledge), to make use of the awareness of previously studied number facts (declarative knowledge), and to select appropriate study behavior (conditional knowledge)" (p. 435).

While there is much overlap in how researchers and theorists define the term metacognitive knowledge, there continues to be a lack of agreement regarding what it comprises. This appears to be due to the lack of communication and collaboration by researchers studying metacognition as it relates to their area of specialty. Montague (1992), Desoete et al. (2001) and Garofalo and Lester, Jr. (1985), for example, have studied metacognition as it applies to mathematics problem solving. White (1999), on the other hand, has focused on distance learning. As with metacognition as a construct, there is a need for a common language that can serve researchers and theorists working in different content areas.

Metacognitive Skills

Research in metacognition is concerned with two basic problems (Palincsar & Brown, 1987; Weinert, 1987). The first has to do with the availability of factual or declarative knowledge about cognition. The second problem has to do with the "availability of strategies (procedure knowledge) that may be applied in order to control (monitor) and regulate cognitive activity (thinking about thinking)" (Weinert, p. 17). A learning strategy is an individual's way of organizing and using a particular set of skills in order to learn content or accomplish other tasks. Learning strategies have to do with how to learn, rather than with specific curricular content or skills. Strategy instruction focuses upon teaching students "how to use their existing knowledge and skills, coordinate their thinking, and make decisions and take actions during the learning process" (Van Reusen & Head, 1994). Main (1985) used the term "reflective activities" to describe strategies that "involve learners in a monitoring and evaluating of their own learning capacities and style" (p. 91).

The study of metacognition has significant implications for the teaching of study skills (Stewart & Landine, 1995). Several researchers in the area of metacognition believe that study skills must be presented from a metacognitive perspective in order to achieve long-term learning (Biggs, 1985; Brown, 1978; Desoete et al., 2001; Elen & Lowyck, 1998; Hartlep & Forsyth, 2000; Palincsar & Brown, 1987; Pressley, Borkowski, & O'Sullivan, 1984; Resnick & Klopfer, 1989; Schneider et al., 1986; Stewart & Landine). According to Stewart and Landine, students who are taught study skills in a metacognitive perspective "will have increased self-confidence, will have skills to accomplish learning tasks, will be able to apply these skills in appropriate contexts and generally will be better able to achieve the academic goals they establish for themselves" (p. 5).

Pressley et al. (1984) emphasized the distinction between specific strategies, such as mnemonics, and strategy knowledge, which is what a learner knows about the strategy. The researchers hypothesized that there is a "dynamic bidirectionality" between strategy use and strategy knowledge (Pressley et al.), which is consistent with the interrelatedness of metacognitive components (Flavell, 1979). Pressley and colleagues stated that metamemory instruction (strategy knowledge) did not always have to be explicit, especially with older students and adults. However, metamemory interventions have a "pronounced effect upon youngsters" and should be incorporated into study skills instruction (Pressley et al., p. 104).

In a review of metacognitive instruction, Palincsar & Brown (1987) identified five features of successful metacognitive instruction. These are:

1. Careful analysis of the task at hand.
2. Identification of strategies which will promote successful task completion.
3. Explicit instruction of these strategies accompanied by metacognitive information regarding their application.
4. Provision of feedback regarding the usefulness of the strategies and the success with which they are being acquired.
5. Instruction regarding the generalized use of the strategies. (p. 73).

The authors concluded that metacognitive instruction should be conceptualized as an integral part of teaching activity.

"Research in cognitive and metacognitive strategies provides a basis for understanding the processes involved in higher level cognitive activities such as mathematical problem solving, and also provides a foundation for process-based instruction" (Montague, 1992, p. 231). There have been, however, mixed results in the research on metacognitive strategies. Scruggs, Mastropieri, Monson, and Jorgenson (1985), for example, reported in a research review that studies show instruction in metacognitive strategies have resulted in increases in learning. On the other hand, Resnick & Klopfer (1989) writing four years later, stated that research had not yet examined the effect of combining thinking skills with content, only on the effects of one or the other alone. Interestingly, Resnick and Klopfer went on to argue, however, that practitioners should not wait for such research to conducted. As will be reviewed later in this chapter, my own review of research on metacognition is consistent with that of Resnick and Klopfer, however, there appears to be an accelerated interest in this research topic in the literature.

Brown (1978) suggested a general description of the kinds of skills needed for efficient predicting, planning, checking, and monitoring. These include:

1. Recognizing problem difficulty
2. Using inferential reasoning to check validity of information
3. Predicting outcome of strategy use
4. Predicting task difficulty
5. Planning study-time needs
6. Monitoring success of strategy
7. Checking outcomes for internal consistency

8. Checking outcomes against common sense criteria (reality testing) (p. 96).

Brown stated that the goal is to guide the student toward “self-interrogation” where the student thinks dialectically, as in the Socratic teaching method. “The desired end-product is that the student will come to perform the teacher’s functions himself via self-interrogation” (Brown, p. 97). Hartlep and Forsyth (2000) found that self-referencing, a similar concept to that of Brown’s self-interrogation, to be an effective strategy in enhancing learning and retention.

Meichenbaum (1977) theorized that a person can be taught to change what he or she says internally, using reflection, explanation, interpretation, information-giving, and cognitive modeling. It is the internal conversation that a person has with himself or herself, that influences his or her behavior, according to Meichenbaum. He suggested that as a person learns to use new tools, that the person’s behavior will change. Similarly, Montague (1992) stated that “successful problem solvers, consciously or unconsciously (depending on task demands) use self-instruction, self-questioning, and self-monitoring to gain access to strategic knowledge, guide execution of strategies, and regulate use of strategies and problem-solving performance” (p. 231). Alex Main (1985) referred to such internal problem solving as “reflective activities” in which the learner monitors and evaluates his or her own learning capacities and styles (p. 91). Internal reflection by students was also supported by Gibbs and Northedge (1979), who stated that encouraging students to have a questioning and self-analytical attitude toward their own learning strategies results in independent learners, while training in specific techniques and the giving of advice can result in a dependent learner.

The list of specific learning strategies and skills extends far beyond the scope of this paper. However, a brief review of commonly used skills and strategies is included. Among a multitude of strategies and skills that are commonly used in education today are mnemonics (Brown & Barclay, 1976), debriefing (Pearson & Smith, 1985), coaching (Candy, Harri-Augstein, & Thomas, 1985), graphic organizers (Robinson & Kiewa, 1995), modeling (Meichenbaum, 1977; Schoenfeld, 1987; Williams, 2000), thinking aloud (Williams), questioning (King, 1991; King 1994; Osman & Hannafin, 1994), monitoring (Schraw, Dunkle, Bendixen & Roedel, 1995), and writing (Ellis, 2001; Powell, 1985; Walker, 1985; Wilson, 1986). Whether these skills are used to guide students toward internal reflection and metacognitive skill or toward dependence upon a concrete skill, depends upon how they are implemented by the teacher.

In summary, study skills that are presented in a metacognitive perspective are more powerful than such skills taught in isolation. By reflecting on their own learning strategies—and those of others—students can be become more purposeful in their learning (Bruner, 1986). Pressley & El-Dinary (1993) observed that the very best strategies instruction emerges from collaborations between educational scientists and teachers, with each making critical contributions that cannot be made by the other. Metacognitive skills can be seen as the voluntary control that individuals have over their own cognitive processes, which is dealt with in more detail in the next section on self-regulated learning. The goal of strategy instruction is to “demystify the classroom and increase the opportunities for students to be intrigued and challenged, rather than baffled, by classroom experience” (Palincsar & Brown, 1989, p. 37). In the next section reviewing self-regulated learning, the examination of how the use of specific learning strategies to support independent learning will be expanded.

Self-regulated Learning

It would be a serious oversight in a literature review on the topic of metacognition to omit a discussion and review of self-regulated learning. After all, most definitions of metacognition include reference to regulatory aspects as an integral component. However, self-regulated learning is a huge theoretical construct in its own right, and its exhaustive examination is beyond the scope of this paper. With this in mind, an overview follows of the linkages between self-regulated learning and metacognition. Included is an overview of the definitions, components, models for teaching, example strategies, and theoretical roots of self-regulated learning.

"Psychologists interested in mechanisms of growth and change have traditionally been concerned with self-regulatory processes, because a great deal of learning takes place in the absence of external agents" (Brown, 1987, p. 76). Self-regulated learning emerged as a construct that encompassed the research conducted over the last 30 years on cognitive strategies, metacognition, motivation, task engagement, and social supports in the classroom (Paris & Paris, 2001). According to Paris and Paris, self-regulated learning is a synthesis of many constructs in learning and motivation.

It is the evaluating of one's own thinking, according to Manning & Glasner (1996), that links self-regulated learning strategies and metacognition. Writing in 1987, Brown stated that "recently, the term metacognition has been extended to encompass regulatory functions" (p. 77). The view that self-regulation is an integral part of metacognition is also held by Borkowski (1992), who referred to self-regulation as "the heart of metacognition." Borkowski explained that its function is to analyze or "size up" tasks in order to select appropriate problem solving strategies (p. 253). It is the interaction that occurs between knowledge about tasks and of strategies that lead individuals to "monitor one's task understanding and regulate one's strategy use" (Garofalo & Lester, 1985).

Since self-regulated learning is relevant to many aspects of learning and control, diverse theoretical views have been proposed as useful in its examination (Paris & Paris, 2001). Among these are the constructivist theories of Piaget, Vygotsky's sociocultural theory, social learning theories, and information-processing theories (Paris & Paris). The gradual shift of responsibility from teacher to student, for example, is an application of Vygotky's theory that children progress from external regulation by adults to internal regulation (Vygotsky, 1978). Brown (1987), however, proposed that self-regulation grows from different roots than metacognition.

While knowledge about cognition and regulation of cognition are closely related, they are readily distinguishable and are derived from different historical roots (Brown, 1987). Brown cautioned, however, that since the two concepts are interrelated, attempts to separate them lead to oversimplification. Knowledge about cognition, which I have labeled as metacognitive knowledge in this paper, is rooted in the developmental theories of Piaget (Brown). This type of knowledge is "stable, statable, often fallible, and often late developing" usually requiring learners to reflectively consider their own cognitive processes (Brown, p. 68). Piaget (1976) theorized that such reflective abstraction requires a person to have reached a formal operational developmental stage. The roots of self-regulated learning, however, can be found in the works of early educational psychologists, including Dewey, Thorndike, and Binet (Brown, 1987). Self-regulated learning activities are often “relatively unstable, not necessarily statable, and relatively age independent” (Brown, p. 69). Since knowledge and regulation of cognition are linked to different types of activity, the use of the term metacognition to describe them both can be confusing.

Zimmerman (1989) proposed a social cognitive foundation for self-regulated learning based on Bandura’s (1986) triadic theory. In this model “students’ self-regulated learning is not an absolute state of functioning, but rather varies on the basis of the academic context, personal efforts to self-regulate, and outcomes of behavioral performance” (Zimmerman & Martinez-Pons, 1990, p. 51). The intrinsic motive to learn will be dealt with in the following section of this paper on metacognitive beliefs.

Self-regulated learning is defined as the active process through which individuals direct and sustain their metacognition, motivation, and strategies to optimize their learning (Zimmerman, 1990). Self-regulated learners are students who plan and check their work, are aware of their thought processes, use cognitive strategies to achieve their goals, and exert effort” (Malpass, O’Neil, & Hocevar, 1999, p. 283). Successful self-regulated learners are good at “arguing with themselves” (Schoenfeld, 1987).

Using a classroom example to operationally define self-regulated learning, Palincsar & Brown (1989) stated that “students’ responses to the mysteries of classroom activity reflect, in part, their awareness of the variables that are important to learning and their ability to take control of their learning environment” (p. 19). It is the autonomy and control by the individual, who “monitors, directs, and regulates actions toward goals” that is emphasized in self-regulated learning (Paris & Paris, 2001, p. 89). This is typically seen in the classroom in the contrast between students who are inconsistent with assignment completion and classroom participation, and those who are conscientious about classroom responsibilities and engagement. Expert learners have an awareness of their personal strengths and weaknesses, the knowledge about whether or not they have succeeded, and the ability to create reasonable goals (Weinstein & Van Mater Stone, 1993).

The components that make up the construct of self-regulated learning are subject to a wide range of definition by researchers. In their definition of self-regulated learning, Pintrich and DeGroot (1990) included metacognitive strategies, management of effort, and actual strategies. Similarly, Zimmerman (1990) included strategies, which involve purpose of agency, a self-oriented feedback loop, and recognition of the need for preparation as components of self-regulated learning. Schoenfeld (1987) included the same three components in his definition, although he divided metacognitive strategies into two separate components, which were problem understanding and planning.

Zimmerman (1989) described three major determinants of self-regulated learning. The first, *personal influences*, involve a student's self-efficacy perceptions in relation to his or her knowledge, metacognitive processes, goals, and affect. *Behavioral influences*, the second determinant, includes self-observation, self-judgement, and self-reaction. Self-observation refers to monitoring of one's own performance. Self-judgement refers to comparing one's own performance to a standard or goal. Self-reaction refers to "such personal processes as goal setting, self-efficacy perceptions, and metacognitive planning, as well as behavioral outcomes" (p. 334). The third determinant, *environmental influences*, includes modeling, verbal persuasion, direct assistance, literary or symbolic information, and structure of the learning context. Modeling effective self-regulatory strategies can improve the self-efficacy of learners, which in turn has a positive impact upon the student's self-regulatory ability (Bandura, 1986; Zimmerman, 1989). Verbal persuasion, according to Zimmerman, is less effective in conveying self-regulatory strategies, but is effective in combination with modeling. Direct assistance by a teacher or peer and accessing information in written or symbolic form are both sources of social support widely used by self-regulated learners. Finally, structure of the learning context, the last environmental element, affects self-regulated learning in that any change in a learning setting can affect the difficulty of tasks. In turn, the difficulty of tasks influences self-efficacy (Bandura, 1986), a key component of self-regulation.

Over the last two decades there has been considerable interest in teaching strategies to students designed to assist their self-regulated learning. There is wide acceptance that individuals can be taught to regulate their behaviors, and that "these regulatory activities enable self-monitoring, and executive control of one's performance" (Bransford, Brown, & Cocking, 1999, p. xii). Palincsar and Brown (1989) stated that a model of instruction for self-regulated learning was "emerging from the research literature" that supported strategy instruction as an integrated part of the curriculum (p. 37). Such instruction includes assessment of current strategy use, explanation regarding the nature and use of the strategies, modeling and guiding practice in the use of strategies, and opportunities to use strategies across the contexts in which they are useful (Palincsar & Brown, 1989).

While Paris and Paris (2001) embraced a constructivist philosophy in believing that every student constructs his or her own theory of self-regulated learning, the authors offered three ways that understanding of self-regulated learning can be enhanced. These are indirectly through experience, directly through instruction, and elicited through practice (Paris & Paris). All three of the opportunities—indirect induction, direct instruction, and elicited actions—"probably operate together in the classroom as children create their theories about learning in school and their own abilities as they work with teachers, parents, and peers" (Paris & Paris, p. 99). There has been debate among experts in the field, however, about generality and specificity of self-regulatory skills (Campione, 1987).

Campione (1987) pointed out a weakness in teaching the use of specific self-regulated learning skills to students. While Feuerstein (1980) believed that skills should be taught separate of content, rather than in context, including contextual skill instruction along with content-free skill instruction is now commonly seen. Often skills are specific to a very small class of situations, which requires the learner "to discriminate the situations in which the routine would be appropriate from those in which it would not" (Campione, p. 125). Campione advocated the instruction of "truly general skills" in which "the skill or routine could simply be used in a whole battery of problem solving situations, eliminating the need for a detailed analysis of the task being attempted" (p. 125). An outcome of this argument may be the balanced approaches for skill instruction recently seen in the literature on self-regulated learning (Bransford, Brown, & Cocking, 1999; Paris & Paris, 2001; Paris & Winograd, 1999).

Paris and Winograd (1999) synthesized research on self-regulated learning as it applied to classroom instruction. Among 12 principles that the authors presented, are three that focus on diverse ways that self-regulated learning can be taught. First, self-regulation can be taught through explicit instruction, directed reflection, metacognitive discussions, and participation in practices with experts. Second, self-regulation can be promoted indirectly by modeling and by activities that entail reflective analyses of learning. Third, self-regulation can be promoted by assessing, charting, and discussing evidence of personal growth (Paris & Winograd). Paris and Paris (2001) articulated a need for expanded research into how self-regulated learning can be applied in the classroom and offered three broad questions to guide future research. These are: "What does it mean for students to be self-regulated? How do students become self-regulated? Are there individual differences in self-regulated learning?" (pp. 98-99). "The synergy between practices in classrooms and research on self-regulated learning should be useful for many years" (Paris & Paris, p. 99).

In summary, self-regulated learning is a vast and expanding area of study that is integral to the construct of metacognition. It is also a separate body of research and literature that focuses upon both skill development and internal mechanisms of regulation and control. Self-regulated learning is certainly too broad a topic on its own for a dissertation, let alone as a component of one focusing on the nature and effects of personal reflection. However, it would be negligent to not include a brief overview of self-regulated learning, for it offers a different lens through which metacognition may be viewed. Self-regulated learning is an important area of study that offers both theoretical and practical support for the classroom teacher and individual student. Self-regulated learning is "more that a developmental milestone tied to grade levels or an educational achievement tied to specific learning. Both experience and context contribute to self-regulated learning" (Paris & Paris, 2001, p. 99).

Metacognitive Beliefs

Metacognitive beliefs are the general ideas and theories people have about their own and other people's cognition (Cornoldi, 1998; Desoete, et al., 2001). As a third component of metacognition, metacognitive beliefs includes such aspects as attribution, motivation, self-esteem, and self-efficacy. As with the previously reviewed components of metacognition, metacognitive beliefs is a huge and multi-dimensional construct that can easily lead to study far beyond the scope of this paper. It is important, however, to include an overview of metacognitive beliefs in order to recognize and acknowledge the broad construct in which reflective assessment is enmeshed.

There is ongoing debate among researchers about whether metacognitive beliefs should be considered a separate component of metacognition, or included within metacognitive knowledge (Dickson, Collins, Simmons, & Kameenui, 1998). However, these unresolved arguments are typically based upon theoretical concepts, rather than upon empirical evidence (Desoete, et al., 2001). Desoete and colleagues pointed out that "even authors who are in favor of a two-component approach to metacognition have found it important to study attribution" (p. 436).

Bandura (1986) defined self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance" (p. 391). According to Malpass, O'Neil, & Hocevar (1999), self-efficacy refers to people's "specific beliefs" about their capability to perform tasks or accomplish outcomes (p. 282). How students view their own capabilities, in turn, influences their effort on tasks (Bandura, 1993; Schunk, 1984). Bandura (1993) theorized that beliefs about self-efficacy contribute to motivation by influencing the goals people set, the degree of effort that is exerted, their perseverance when challenged, and their resilience when not successful. According to Bandura, "self-directed learning requires motivation as well as cognitive and metacognitive strategies" (1993, p. 136). Weiner (1980, 1986) theorized that students develop causal inferences, or attributions, regarding their achievement that influences their ongoing academic motivation. Similarly, Zimmerman (1990) proposed that a cycle exists in which students' use of cognitive and metacognitive strategies enhance perceptions of self-efficacy, which in turn provide a motivational basis for continued self-regulation during learning.

Bortkowski (1992) asserted that all cognitive acts have motivational consequences and that as students come to recognize the importance of being strategic their feelings of self-efficacy emerge. This view was supported by Weinert (1987), who advocated an integrated study of the effects of metacognition and motivation upon learning activities. According to Weinert, "There is sufficient overlap in the tasks and concerns of the fields of cognition, metacognition, and motivation to urge the development of an integrated research program" (p. 14). Understanding the interaction between cognitive and affective development is essential if the reasons why "a child will or will not use active and efficient strategies for efficient information processing" (Wiens, 2001, p. 145). In support of this view, Biggs (1985) suggested that a reason prescriptive strategy approaches did not sustain long-term results was that students' motives had not been adequately addressed. Stewart & Landine (1995) suggested that intrinsic motivation provides the energy to sustain effort on tasks longer than will extrinsic motivation. Internal locus of control is an important antecedent to high academic achievement (Stewart & Landine).

Paris and Paris (2001) articulated the role motivation plays in the learning equation in a discussion of self-regulated learning. " Self-regulated learning is informed by metacognition from self and others and is fueled by affect and desire" (Paris & Paris, p. 98). Borkowski (1992) referred to the relationship between motivation and learning as "the final link in the metacognitive chain" (p. 254). This positive relationship between motivational variables and learning has been supported by research (Kurtz & Borkowski, 1984; Malpass et al., 1999; Pintrich & DeGroot, 1990). It is also a dynamic that can be enhanced by teacher guidance in ongoing, informal classroom activities (von Glasersfeld, 1991).

In addition to the focus upon teaching students specific learning strategies, some models of strategy instruction also emphasize learner self-perception, self-efficacy, and motivation (Van Reusen & Head, 1994). Students often feel that academic success is due to factors out of their control, such as ability or luck, rather than due to factors within their control, such as effort and efficiency (Van Reusen & Head). Jans & Leclercq (1997) used the term "metacognitive realism" to describe how learners view themselves as learners within the context of an experience (p. 101). Through experience, both successful and unsuccessful, learners develop views regarding their own abilities to achieve in different contexts, which may or may not be accurate (Jans & Leclercq). However, it is students' perception of their academic efficacy that determines their choice of strategies when approaching a task (Zimmerman & Martinez-Pons, 1990).

In summary, while a body of research clarifying the impact of motivation and other elements of metacognitive beliefs is growing (Zimmerman & Martinez-Pons, 1986), further research is needed. As with self-regulated learning, the study of metacognitive beliefs reaches into vast related constructs that while essential to include, also challenge efforts to provide greater clarity to the study of metacognition. It is commonly accepted among researchers and theorists of metacognition, though, that personal beliefs influence performance. Therefore, it will be important for continuing research to be conducted on the effects of such factors as attitude, effort, motivation, and self-efficacy as they relate to metacognition and academic performance.

Content vs. Thinking

Conversation and debate continue among curriculum experts regarding the respective roles of content and reflection, but there appears to be agreement that both are essential ingredients in a student's education (Stiggins, 1995). Some educators feel that teaching content is overemphasized to the detriment of considering thinking as the core of the curriculum (Costa, 2001). Costa (2001) stated that "content, when judiciously selected for its rich contributions to thinking and learning, becomes the vehicle to carry the learning processes. The focus is on the learning from the objectives instead of learning of the objectives" (p. 246).

Student Reflection vs. Teacher Feedback

Sadler (1989) distinguished between "feedback" information that is provided by the teacher to the student, and "self monitoring" information about performance from a student's own appraisal of their work. These contrasting uses of information about learning unveil the critical issue of who owns the information, the teacher or the student. Wilson (1986) stated that the purpose of informal reflective activity is to both encourage student reflection and to provide feedback to the teacher about their teaching (p. 199). Diagnostic information can inform both students' studying and teachers' teaching

(Brookhart, 2001). This view is also held by Orsmond, Merry, and Reiling (2000), who stated that tutor feedback and student learning should be inseparable and “if they become uncoupled, the formative aspect of assessment is lost” (p. 24). According to Taras (2001), “Assessment is not owned by tutors or anyone else” (p. 608). While it is unlikely that students will become competent judges of quality on their own, effective assessment practices empower students to become expert assessors (Earl & LeMahieu, 1997).

When students are involved with the assessment of their learning they are empowered to take ownership of their learning. According to Stiggins:

> To tap the full potential of the bond between classroom assessment and the effectiveness of schools, we must expand our collective vision to include students as assessment users, too. Yet we never think of students in this way. We see them as the examinees. We assess them and then we use the assessment results to decide how to treat them. Our definitions of the roles of assessment in school effectiveness virtually always cast students in a passive role (1996, p. 2).

It is this passive role that Stiggins referred to that is avoided when informal assessment opportunities are integrated into the classroom experience. “It is the goal of empowerment of all learners for which educators should strive” (Ellis & Worthington, 1994). Ellis and Worthington outlined the following motivational characteristics for the empowered student:

1. Empowered students have an internal locus of control.
2. Empowered students expect to be successful.
3. Empowered students are goal oriented.
4. Empowered students are invested in the learning process (p. 9-10).

The crucial role that intrinsic and extrinsic motivation play in the empowerment of students is further discussed by Yancey (1998), who argued that in our culture we are obsessed with external evaluation and because of this teachers need to consciously encourage students to reflect upon their learning. Yancey further stated that if teachers do not ask students to assess their own work—a process based on internal factors and criteria—students are likely to be dependent on external rewards, not knowing where to begin to consider their own performances. In other words, if teachers do not ask students to reflect and self-assess, those processes are not likely to occur due to long established practices of external assessment and extrinsic rewards.

Research on Metacognition

Over the last decade there has been increased interest in research on alternative assessment in the classroom. The lack of consistent structure and language used in this research, however, makes challenging any easy gleaning of findings. To alleviate this problem, Weston, McAlpine, and Bordonaro (1995) suggested that research on classroom assessment should to provide clear definitions regarding the level of participation for teachers and learners, their roles in the learning activity, and the knowledge domain of the learning activity. While there is not yet such clarity in either definitions or methodology, the body of research on alternative classroom assessment is growing and is beginning to offer guidance for classroom application. The ability of students to reflect upon their own learning, while acknowledged in the research literature, is usually found as a component of studies focusing on self-assessment, feedback strategies, or transfer of learning. Few studies focus specifically on the effects of reflection practice on achievement.

Among the most prominent studies identifying what improves student learning is a meta-analysis conducted by the Mid-Atlantic Regional Educational Laboratory of more than 11,000 statistical findings correlating school factors with achievement (Wang, Haertel, & Walberg, 1993). Wang et al. combined results from 179 handbook chapters and reviews, compiled 91 research syntheses, and surveyed 61 educational researchers to estimate the effect of a given method or condition on student achievement. According to this study, students' metacognitive processes had an influence on learning second only to teachers' ability to maintain active student participation.

Students' ability to reflect on their learning and make adjustments accordingly has been identified as one of the most significant determinants of student success (Conzemius & O'Neill, 2001). Schunk (1983) found that systematic observation of one's own learning progress resulted in significantly higher measures of self-efficacy, content skill, and task persistence during posttesting, than did a control group that did not self observe. In this experimental design, eight and nine year olds were randomly assigned to one of three treatment conditions— self-monitoring, external monitoring, or no monitoring. Preliminary analyses of variance revealed no differences due to tester, school, or gender on any pre- or posttest measure nor any significant interactions. Schunk found that the two monitoring conditions did not differ significantly from one another on measures of math subtraction skill and self-efficacy, but that both were found to be significantly different than the no-monitoring condition. Schunk concluded that as children observe their progress during training a heightened sense of efficacy is developed, which is validated as they continue to monitor their own work. He states that self-monitoring allows students to gain capability information on their own (Schunk, 1983, p. 92).

In a cross-cultural study on metamemory and motivation Schneider et al. (1986) found significant difference on post-treatment tasks between the experimental and control groups. In this study of 102 German and 91 American fourth graders, students were randomly assigned to treatment and control groups. While the primary purpose of this research was cross-cultural comparison, embedded within it were findings that have practical significance regarding the value of teaching reflective strategies to students. Students from both cultures who received training on reflective strategies performed significantly higher than did their non-trained counterparts on tests of metamemory and recall. Metamemory was defined by Schneider et al. (1986) as knowledge about memory states and processes.

In another experimental study of sixty first and third grade students, Kurtz and Borkowski (1984) found that, contrary to expectations, children who received both strategy training and metamemory training were not at an advantage in terms of strategy use when compared with students who received only strategy training. In this study students were randomly assigned into three treatment groups—metamemory and strategy training, metamemory-only training, or strategy-only training. General metacognitive instructions did not appear to alter reflective knowledge about memory processes. In their discussion the authors suggest that changes in metamemory take place over long periods of time and that "metacognitive training in their study was insufficient in focus and/or duration to produce these permanent changes" (Kurtz & Borkowski, p. 350). Similarly, Andrade (1999) found no significant difference between the treatment and control groups in terms of metacognitive processing in a study of the effects of self-assessment on metacognitive engagement. In this study of 47 seventh graders from a rural Massachusetts public school, think aloud protocols were collected and coded to provide insight into spontaneous self-assessment, the classification of self-assessment, and the influence of self-assessment on metacognitive engagement and learning. While the author found that treatment students tended to outperform the control group on periodic posttest measures, the results were not statistically significant. Andrade found that approximately three-quarters of the students across both experimental and control groups assessed themselves spontaneously as measured by think aloud protocols. Control group students, however, were not specifically asked to assess their own progress, as were students in the treatment group.

Naglieri and Johnson (2000), in a quantitative study of students with learning disabilities, reported that students with cognitive weakness in planning improve significantly as the result of group instruction on self-reflection and verbalization of arithmetic strategies. This study involved a sample of 19 students who received special education instruction for math from two public middle schools in rural and suburban southern California. The authors divided the sample into one experimental and four comparison groups for purpose of data analyses. While the results of this study suggest that cognitive strategy instruction that teaches students to better use planning processes is useful for those who need it the most, the external validity is limited due to the small sample size and its inclusion of only special education students. The authors recommend that further research needs to be conducted with larger samples of children.

Qualitative Research on Metacognition

Several qualitative studies have been conducted related to metacognition and student learning. It is important to acknowledge that a growing body of qualitative literature exists on the topic of metacognition and student learning, even though the emphasis in this paper is on quantitative, empirical results. This points to what Snow, Burns, and Griffin (1998, p. 35) referred to as a "convergence of evidence" in which qualitative methods can play a complementary role in describing in greater detail the contextual and interpersonal variables not easily captured in tightly controlled empirical investigations. A brief review of qualitative studies related to reflective assessment follow.

In a study on self-assessment, Brookhart (2001) found through interview data that successful high school students constructed what their learning ought to mean to them in terms of both mastery and emergent understandings. The researcher observed one class of 10^{th} grade English students, two sections of 11^{th} grade honors English students, and three sections of 12^{th} grade anatomy elective class. A total of 52 interviews were conducted, involving 50 different students, including 28 in English classes and 24 in anatomy classes. The interview questions were designed to elicit information about the perceived task characteristics, perceived self-efficacy to meet the challenge posed by the task, amount of effort expended and the reasons for that effort, and expected grade and how students felt about the grade. The findings showed that this sample of high achieving students talked in terms of using assessment information both formatively and summatively. Brookhart referred to these subjects as having learned the art of "studenting," of figuring out what their teachers expect of them and doing it well (p. 165). Buehl (1996) found, in another qualitative study, that incorporating self-reflectiveness into a reading curriculum led to increased self-evaluative thinking among high school students. Using a convenience sample of 23 students in a semester-long college-preparatory class, the effectiveness of learning logs as a reflective strategy was focused upon. Students were required to make reflective entries once or twice each week, on which the teacher made written comment every two weeks. Buehl employed two researcher-designed measures to determine the results of this study. First, a rubric was used to evaluate the reflective quality and thoroughness of the learning log entries. Second, a survey was administered at the end of the course to measure student perception regarding their reflective thinking. Since the survey was comprised of open-ended questions that required narrative responses, a statistical analysis of the results was not conducted.

In another study Powell and Makin (1994) investigated the effect of teacher intervention in the process of reflection upon gains in pupils' abilities as learners. The authors designed their study around the three phases of reflection described by Boud et al. (1985). In this investigation of a small group of 12 and 13 year-old students with moderate learning disabilities, the researchers found that students were able to take increasing control of their own learning through reflection on their thinking and reading. The researchers also found that the effectiveness of particular teaching and learning episodes in promoting reflection in children was affected by the social and emotional context of the classroom. In their conclusion, however, the authors state that the increase in performance might have been due to students learning the language of reflective thinking, rather than an actual increase in reflective ability. Powell and Makin called for further research in this area, including evaluative methodology that accommodates the assessment of problem-solving abilities.

Teaching Metacognition

Can metacognitive skills be taught? While there is strong support in the literature for emphasizing metacognitive or reflective behavior among students (Black & Wiliam, 1998; Buehl, 1996; Dewey, 1933; Ellis, 2001; Ellis & Worthington, 1994; Ellison, 2001; Loughran, 1996; O'Neill, 1998; Starnes & Paris, 2000; Yancey, 1998), there is also strong support for teacher guidance of such reflection (Andrade, 1999; Burchell & Westmoreland, 1999; Earl & LeMahieu, 1997; Kirkwood, 2000; Loughran, 1996; Powell & Makin, 1994; Taras, 2001; Wilson, 1986). If learning strategies are to transfer, they must first be consciously articulated and then practiced, so that eventually they become part of a habitual and unconscious approach to learning (Kirkwood). Students should not be left to discover reflective practices on their own. It is this portion of the teaching and learning partnership that teachers need to own. The actual reflective assessment, however, must belong solely to the learner.

If we want students to take responsibility, then we have to allow them to do so (Taras, 2001). If we exclude them from assessment, we are excluding them from any real responsibility and we need to examine our own position and our own motives (Boud, 1995, p. 180). While historically the primary purpose of assessment has been evaluative (Earl & LeMahieu, 1997), assessment is not a single entity, nor does it have a singular purpose or audience (Haney, 1991). "No longer is learning thought to be a one-way transmission from teacher to students with the teacher as lecturer and students as passive receptacles" (Herman et al., 1992, p. 12). As Conzemius & O'Neill summarized, "Reflection is as much a mindset as it is a process, or a set of tools or methods. Reflection is a way of thinking about the world and one's relationship to it" (2001, p. 15).

Conclusion

In American education we are currently involved in a reform effort that has placed assessment practices at the center of the conversation. The emphasis, however, has been often misplaced on standardized testing and other summative measures, rather than appropriately focused upon formative assessments that occur within the school classroom (Stiggins, 1997). The purpose of assessment should be to improve student learning, which means it should be integral to the teaching and learning process (Herman et al., 1992). A seamlessness needs to exist between teaching, learning, and assessment through which students are empowered to take ownership and responsibility of their learning. Reflective assessment, as I have argued in this paper, is a form of metacognition using a formative approach that places students at the center of assessment practice.

Reflective assessment grows out of strong theoretical roots including ancient Greek thought, the philosophy of John Dewey, and cognitive constructivist learning theories. If assessment is accepted as a part of the learning experience that is deeply rooted in constructivist theory, it can be viewed appropriately as an essential component as learners construct new schema and integrate them into their thinking. Learning occurs when students reflect upon what they have experienced, which means that the locus of control for learning rests with the learner.

If one accepts a constructivist view of knowledge acquisition, it follows logically that learners have always owned their learning, for it is a natural cognitive experience. If this is the case, and I have argued that it is, then full acknowledgement that students are empowered learners is long overdue. Not only does accepting this belief enhance the role of the learner, but also it transforms the role of the teacher into the "co-partner and guide" about whom Dewey wrote (1964, p. 10).

While there has been a great amount written on such topics as alternative and formative assessments, reflective practice, metacognition, learning strategies, and constructivism, there are surprisingly few empirical studies dealing specifically with the effects of reflective practice on achievement. I found no studies evaluating program implementations. Typically, I have found and reported on empirical studies that relate tangentially to the topic of reflective assessment. Similarly, there is also a rapidly growing body of qualitative research related to reflective assessment, which as stated earlier, points to a convergence of evidence from different approaches. It is a research area, however, that lacks clear focus, and thus, is integrated into a wide range of seemingly unrelated topics in the fields of education, business administration, medicine, and philosophy.

For full implementation of reflective assessment practices to occur in American classrooms, it will be essential that a specific body of empirical research be conducted. The theoretical construct is defined and a broad base of related research exists. It is time, though, for the research spotlight to focus tightly on the effects of reflective assessment on student achievement. The continued pressures on students and teachers to improve test scores, while usually well intended, are misguided for they fail to acknowledge the central role of the student. It is time to fully embrace student participation and ownership in the assessment process, for in reality they have always owned it. It is a working hypothesis of this study that reflective assessment practice by students is a crucial factor contributing to achievement.

"The future belongs to individuals who can identify their own learning needs and who have the resources to orchestrate and manage their own learning activities" (Weinstein &Van Mater Stone, 1993), p. 32).

Chapter Three

Methodology

It is common practice during lesson closure activities for teachers to facilitate student reflection upon what has been taught. However, in a large body of metacognitive literature relatively few experimental studies have narrowed the focus of study to a single reflective intervention as the independent variable. It is contributing to the filling of this research gap that this study is aimed.

This chapter describes the methods and procedures that were used in the present study. It begins with a description of the subjects who participated in the study, including how subjects were randomly assigned to treatment groups and how a coinciding school district program evaluation led to the choice of upper elementary school students as subjects. A description of the experimental posttest-only control group research design follows, which includes an explanation of the power analysis that was used to arrive at the sample size. In following sections of this chapter the setting, independent variable, instrumentation, procedures and apparatus, data analysis, and limitations and delimitations are described in an effort to clearly spell out how the mechanics of the study were implemented. This chapter provides both the context and the actual research steps that were taken, which will serve as a reference point for the report of results that will follow in Chapter Four.

Subjects

A sample of 141 students from a suburban elementary school in the Northshore School District of Washington State participated in the study. The school had an enrollment of 466 students in grades kindergarten through sixth. The school where the study was conducted serves a neighborhood population of middle to upper-middle class socio-economic status (SES). The percentage of student participation in the federally funded free-and-reduced lunch program, a statistic commonly used to indicate poverty level, was 10%. The ethnic mix of the enrollment was 85% Caucasian, 5 % Hispanic, 5% Asian, and 5% African-American. The school has scored above Washington State averages on both the Iowa Test of Basic Skills (ITBS) and Washington Assessment of Student Learning (WASL), with 2001 total fourth grade percentile scores of 65 and 50, respectively. Special programs available at the school included Learning Assistance Program (LAP), Special Education, and English as a Second Language (ESL). The percentages of the student population enrolled in special programs during the 2001-2002 school years were LAP, 10%, special education, 9%, and ESL, 5%.

Random Assignment

The 141 student participants were randomly assigned to three treatment groups, which each included two separate classroom settings. Each group was comprised of two sub-groups of approximately 24 subjects each for total group sizes of 47, 48, and 46. Random assignment of subjects ensured that homogeneity existed among the three groups regarding gender, ability level, socioeconomic status (SES), and ethnicity. Of the 141 subjects who participated in this study 61 were male and 80 were female, 61 were fifth graders and 80 were sixth graders, 7 received special education instruction for both reading and mathematics, 3 were English Language Learners (ELL), and 5 qualified for free or reduced lunch. Excluding gender and grade level, a total of 15 students fit in subgroups of ELL, special education, and low SES, commonly accepted in educational literature as at-risk factors to academic performance. Of the four ELL students, Spanish was the first language of three subjects and Russian was the first language of one subject.

Fifth and sixth grade students from six different classrooms were randomly assigned to the three treatment conditions using an internet-based research tool (Urbaniak, Plous, & Lestick, 2002). The random assignment procedures resulted in a representative mix of the overall sample in each of the three treatment groups. Experimental Group I included 4 students with at-risk factors, Experimental Group II included 6 students with at-risk factors, and the Control Group included 5 students with at-risk factors.

Program Evaluation Component

Fifth and sixth grade students were selected for this study because of an ongoing program evaluation in the Northshore School District regarding the effectiveness and appropriateness of the mathematics curriculum at the elementary and junior high school levels. Since the two levels currently have different mathematics textbook adoptions—*Everyday Mathematics* (2000) at the elementary level and *Connected Mathematics Program* (Lappan, Fey, Fitzgerald, Friel, & Phillips,1998) at the junior high level—concern existed regarding mathematics performance as students transition to from 6th to 7th grades. The school district was interested in finding out if components of the current junior high school mathematics curriculum would be appropriate for grades five and six. The interest of the school district curriculum department in piloting junior high mathematics curriculum at the elementary level created an opportunity to conduct this experimental study. Since needs of the school district were met through the study, the concerns about the disruption to the educational process that often accompany random assignment were minimized.

Since the study was incorporated within a school district approved pilot of mathematics curriculum, acquiring parent permission proved to be a minimal challenge. A letter announcing the curriculum pilot and inviting inquiries was sent to parents early in the school year (see Appendix B, p. 115). As a result, four special education students whose Individual Education Program (IEP) would have been disrupted by the random assignment were excluded prior to the random assignment to treatment groups.

Research Design

This investigation of student reflective assessment and its effects upon mathematics achievement was an experimental posttest-only control group design (Campbell & Stanley, 1963). The independent variable, referred to in Table 1 as the intervention, was reflective assessment strategies, which were practiced only by Experimental Group I. The dependent variable was the mathematics scores for Experimental Group I, Experimental Group II, and the Control Group, as measured by a researcher-designed instrument (see Appendix F, pp. 124-132).

In addition to the random of assignment of students, the six teacher participants were randomly assigned to the one of the three treatments. Since sixth grade students at the school were already grouped for mathematics instruction, approximately half of the teacher and student participants in the study were accustomed to regrouping for mathematics lessons.

Table 1

Posttest-Only Control Group Design

Fifth/sixth Grade Sample		
Group	Intervention	Posttest
R_1	X	O
R_2		O
R_3		O

R_1 = Randomly assigned experimental group I
R_2 = Randomly assigned experimental group II
R_3 = Randomly assigned control group

Power Analysis

A *power analysis* was conducted to determine the number of subjects necessary to detect any effects that might result from the independent variable. Statistical power "refers to the probability that a particular test of statistical significance will lead to rejection of a false null hypothesis" (Gall, Borg, & Gall, 1996, p. 187). As will be explained in greater detail in Chapter Four, a one-way analysis of variance (ANOVA) was selected for the statistical analysis of the differences in performance of the three groups of subjects. Gall et al. (1996) recommend a minimum of 126 subjects for a three-group analysis of variance at a significance level of 0.05. In combination, sample size, level of significance, directionality of a test, and effect size were considered to estimate the statistical power of the posttest. After consideration of these four factors, it was decided to select a sample greater than 126, which with a *p* value of .05 would result in statistical power at the .7 level (Gall et al., 1996).

Setting

The elementary school in which the study was conducted is located in suburban north King County, northeast of Seattle, Washington. It is one of twenty elementary schools in the Northshore School District, which encompasses the cities of Bothell, Woodinville, and Kenmore, as well as large areas of suburban north King County and south Snohomish County. The school in which the study was conducted is located within the city limits of Bothell, and serves a suburban neighborhood that includes middle class housing developments, single-family homes on acreage, and a mobile home park.

The school, which has been in existence for ten years, is located in what had been a high-growth area of the school district, although enrollment growth had leveled over the three years preceding the study. Three portable classrooms were used at the school, which was an indication of the enrollment growth in the school's service area. Of the six teacher participants in the study, one was assigned to a portable classroom.

Parent involvement is strongly encouraged and embraced at the school. The PTA had 350 registered members during the 2002-2003 school year, an indication of high parental involvement in a school with an enrollment of 466. Typical of a suburban school community, approximately 80% of the student population at the school was represented by a PTA member.

Independent Variable

Intervention

The Experimental Group I intervention involved 16 sessions of statistics instruction spread over four weeks during November 2002. The statistics lessons were taken from the *Data About Us* unit of the *Connected Mathematics Program* (*CMP)* sixth grade curriculum. Included in each lesson was a reflective assessment activity, which was the independent variable in the study. The reflective assessment activity was provided to accompany each of the 16 mathematics lessons. Two separate reflection strategies were combined to form the independent variable, with both being used each day (see Appendix C, pp. 117-119). A description follows of the two reflective assessment strategies that served as the independent variable in the study:

1. *I Learned Statement*— I Learned Statements are statements of personal learning that are written by learners during the closure of a lesson (Ellis, 2001; Simon, Howe, & Kirschenbaum, 1972). As the title implies, an I Learned Statement is a student's individual written reflection upon the key concepts or understandings that he or she has learned in a lesson. In this study, each lesson for Experimental Group I concluded with an I Learned Statement, which students were allowed five minutes to complete.
2. "*Think Aloud*" Strategy— The Think Aloud strategy is a technique in which students verbalize their thinking as they solve problems. It is an extension of a reflective strategy that Ellis (2001) referred to as *Talk About It* (pp. 89-91). In this study the Think Aloud strategy was practiced by students in Experimental Group I as they shared their written I Learned Statements with a partner.

For the purposes of this study, the Think Aloud strategy was used in conjunction with I Learned Statements. Together, these reflective strategies served as the independent variable in this experimental study. As closure to lessons for Experimental Group I, students first wrote an I Learned Statement, then practiced the Think Aloud strategy by discussing their written statement with a classmate, and finally, students edited their first I Learned Statement or wrote a second one. The I Learned Statements were turned in to the teacher at the end of each class period and were collected by the researcher during regular site visits.

Experimental Group II received 16 sessions of statistics instruction using the *Data About Us* component of the *CMP* curriculum. Students were exposed to the same materials, lesson activities, instructional approach, and 50-minute instructional periods as were students in Experimental Group I. However, Experimental Group II did not receive the reflective assessment strategy treatment, which was the independent variable in the study. Experimental Group I teachers were instructed to limit the reflective assessment strategy to five minutes at the close of each lesson. Experimental Group II and Control Group teachers, however, were instructed to follow lesson scripts that did not include a closure activity or summary at the end of each class period.

The Control Group, received 16 sessions of geometry instruction taken from the *Covering and Surrounding* unit of the *CMP* curriculum. The Control Group did not practice reflective strategies at the end of each lesson. At the conclusion of the study, the control group was administered the statistics posttest (the dependent variable) along with the experimental groups.

Instrumentation

The posttest instrument for the study was a researcher-designed measure that was aligned with learning objectives in the statistics unit of the *CMP* sixth grade curriculum. The posttest included 36 multiple choice items that measure specific learning objectives stated in the sixteen *CMP* lesson outlines that were presented to Experimental Groups I and II. Both face validity and content validity of the instrument appeared to be high in consultations with university and curriculum development experts.

A pilot test of the instrument was conducted at a neighboring elementary school not participating in the study and involved 23 sixth grade students. As a result of the pilot, one of 37 original items on the posttest was found to be unreliable, and was dropped from the instrument resulting in the final 36-item test that served as the dependent variable in the study. Two internal consistency estimates of reliability were computed for the Statistics Test. These were Cronbach's alpha coefficient and a split-half coefficient, referred to as a Spearman-Brown corrected correlation (Green, Salkind, & Akey, 2000). The Cronbach's alpha and split-half coefficients were .72 and .71, respectively, each indicating satisfactory reliability (see Appendix G, pp. 133-144). According to Vogt (1999), Cronbach's alpha scores above .70 "suggest that the items in an index are measuring the same thing" (p. 64).

On each of the 36 items included in the instrument, subjects were directed to select one of four possible answers. Raw scores were used in the data analysis to calculate descriptive and inferential data. Since test items were equally weighted, the possible raw scores on the instrument ranged from 0 to 36.

Procedure and Apparatus

To ensure that the intervention was conducted properly and consistently over the course of the study, scripted plans for each of the 16 lessons were provided for the teachers of the experimental and control groups. The six teacher participants were trained prior to the study, during which emphasis was placed on their need to precisely follow the lesson scripts. Daily Lesson Logs were used for teachers to provide written feedback to the researcher (see Appendix D, p. 120). The experimenter also made frequent informal visits to the classroom and interacted regularly with the teacher participants through the course of the study. Informal and regular observation of each of the three groups over the four weeks of the study verified that the prescribed lesson scripts and reflective interventions were implemented as planned (see Appendix E, p. 122).

While the focus of the study was on the statistics lessons, which were taught to Experimental Groups I and II, it is important to emphasize that the Control Group was also involved in a worthwhile learning experience. As were the statistics lessons taught to the experimental groups, the geometry lessons that were presented to the Control Group were aligned with school district and Washington State mathematics standards. This alignment of lesson objectives was reviewed during the teacher training that preceded the study and diffused their potential of concern about the loss of four weeks of instructional time.

Materials

The curriculum used for this study of the effects of individual reflective assessment was purposefully selected for compatibility with the current school district mathematics adoption. From the initial meetings with the Northshore School District curriculum department staff, the school district testing and research committee, the six teacher participants, and the school principal, it was emphasized that the mathematics lessons taught during the study would be aligned with their current goals and objectives. Furthermore, it was pointed out in a review of the instructional benchmarks for grades five and six, that the content presented during the study would both reinforce previous learnings and introduce that which would required in the future. Since *CMP* was the current mathematics curriculum used in junior high schools (grades 7 through 9) in the Northshore School District, its use at grade six was seen as a strategy to enhance mathematics performance as students transition from elementary to junior high school.

Lessons were taken from two instructional units in the 1998 edition of *CMP* sixth grade curriculum— *Data About Us* (statistics) and *Covering and Surrounding* (geometry). Experimental Groups I and II were taught lessons from the *Data About Us* unit, and the Control Group was taught lessons from the *Covering and Surrounding* unit. Student participation in this study introduced them to mathematics concepts that would be part of their seventh grade curriculum and on which they would be assessed on the seventh grade Washington Assessment of Student Learning (WASL). Therefore, in addition to the research purposes of this study, there was practical value for all student participants.

A second practical use for this study was field testing of the *CMP* curriculum at the elementary level. It is this program evaluation component of the study that was emphasized with the six teacher participants. The experimental nature of the study was not discussed with any of the study participants until the study was completed in order to protect against threats to internal validity. Following the study, the Northshore School District Teaching and Learning Department consulted with the teacher participants regarding the appropriateness of adopting components of the *CMP* curriculum for the sixth grade level. Since the existing elementary mathematics adoption was seen as not meeting all of the needs at the sixth grade level, this study provided valuable information to the school district regarding the appropriateness of incorporating components of *CMP*.

Posttest, Performance Test, and Retention Test

Three measures were administered following the four-week study. First, the posttest was administered to both experimental groups and to the control group immediately following completion of the instructional units as part of the experimental posttest-only control group design. In addition, the posttest was re-administered six weeks later to determine the retention levels of the three groups of students, which is referred to as the "retention test" in this dissertation. Finally, an end-of-unit performance test from the *Data About Us, CMP* (1998) curriculum (see Appendix H, pp. 145-148) was administered to determine if there was similarity of results when compared to the researcher-designed posttest.

Internal Validity

According to Campbell & Stanley (1963), a posttest-only control group design controls for threats to internal validity caused by history, maturation, testing, instrumentation, regression, selection, mortality, and interaction of selection and maturation. Subject mortality, however, proved to be a potential issue in the analysis of the retention test data, which will be reported and discussed in Chapters Four and Five.

In addition to the initial eight threats to internal validity, " Campbell and Thomas Cook expanded this list to include four more extraneous variables" (Gall et al. 1996, p. 468). These four additional threats to internal validity, experimental treatment diffusion, compensatory rivalry by the control group, compensatory equalization of treatments, and resentful demoralization of the control group, were the most serious risks to the internal validity of this study. In order to control for these possible threats, careful attention was paid to keeping the experimental nature of the study confidential until after its completion. While school district curriculum administrators, the school district testing and research committee, and the school principal were fully aware of the experimental design of the study, the teacher participants, the student subjects, and parents did not. Since the study was conducted as part of a genuine school district curriculum pilot, something that occurs on a regular and cyclical basis, it was possible to keep confidential the experimental nature of the study without violating ethical considerations. It was the program evaluation strand of the overall project that was reviewed with parents prior to the study (see Appendix B, p. 115) and emphasized with teachers in training and as the study progressed.

It is important to emphasize that prior to communicating with staff members and parents, both the school district curriculum department and the joint district-teachers' union Testing and Research Committee had approved the communication plan about the study (see Appendix A, pp. 110-114). By keeping the visible focus on the pilot of new mathematics curriculum, rather than the experimental nature of the study, the remaining four threats to internal validity were controlled.

Data Analysis

This investigation was experimental; therefore, both descriptive and inferential statistics were computed in analysis of the research questions. Tests of significance were analyzed ($p < .05$), and appropriate post hoc analyses were completed where warranted. Descriptive statistics included frequencies, means, medians, standard deviations, and skewness and kurtosis statistics for all dependent variables.

In order to determine the nature of the relationships between the independent variable and dependent variable, inferential statistical procedures were conducted. Research hypotheses were each tested using a one-way analysis of variance (ANOVA). The significance of effects was analyzed at an alpha level of .05. An ANOVA was an appropriate statistical procedure due to the sample size, single dependent variable, single independent variable, and the use of posttest-only design (Gall et al., 1996).
The Tukey Honestly Significant Difference (*HSD*) Test, the Scheffe Test, and Dunnett's *C* Test were used as post hoc procedures to follow-up statistical significance found by the ANOVA procedures. In addition, in response to descriptive statistics regarding kurtosis of the group distributions, the nonparametric Kruskall-Wallis and Mann Whitney *U* procedures were conducted.

Limitations and Delimitations

The limitations of this research were related to the generalizability of the study. Generalizability is determined by the adequacy of the sample size and the degree to which the sample is representative of the target population (Gall et al., 1996). While the sample size and accompanying randomization procedures were not a problem with this study, the demographic data indicated that the sample represented a narrow segment of the population of students in Washington State. The socio-economic status (SES), ethnic mix, and standardized test scores all indicated that the school in which the study occurred was representative of a suburban and relatively affluent population. Any generalizing of results to other settings will need to be done with this in mind.

A major limitation to doing research in a school setting is the lack of control over intervening or confounding variables. Random assignment of both students and teachers provided a high level of control during the study, however, such variables as teaching style and personality, student attendance patterns, and other varying conditions within a particular classroom could have confounded the results of this study. Since the study was completed in the sixteen instructional days in November 2002 prior to a Thanksgiving holiday, a potential distraction of student attention due to the December holiday season was avoided. During the final five days of study parent-teacher conferences were conducted, for which students were dismissed from school early. The mathematics lessons, however, were not disrupted during this final week of the study. In fact, the coinciding of parent-teacher conference week with the study, helped to minimize the perception of disruption often ascribed to the random assignment of students with the school setting.

Delimitations, which are deliberately imposed limitations of a study imposed by a researcher (Rudestam & Newton, 2001), included the focus upon the fifth and sixth grade level, the use of mathematics curriculum as a content to examine the effects of reflective assessment, and the selection of two specific metacognitive strategies to serve as the independent variable. In addition, the choice to use a researcher-designed instrument to measure the dependent variable is a delimiting factor in this study.

Summary

The study aimed at measuring the effects of reflective assessment upon student achievement. A sample of fifth and sixth grade subjects from a suburban elementary school were randomly assigned to three treatment groups. In addition, six teacher participants were randomly assigned to one of the three treatment groups, resulting in two mathematics classes for each of the treatments.

The study compared the performance of students who practiced a reflective assessment strategy with students who did not. Scripted lesson materials were provided to the teachers to control threats to internal validity, with the only difference between Experimental Groups I and II treatments being the reflective assessment intervention with Group I. The Control Group was taught a geometry curriculum without the reflective assessment intervention.

A researcher-designed instrument was used that was tightly aligned with the lesson objectives stated in the *Connected Mathematics Program* (Lappan et al., 1998), from which lesson materials were drawn. The instrument reliability and its face and content validities were found to be appropriate for this study. In addition to the posttest, a retention test, and a performance test were also administered to subjects as a part of this study.

A posttest-only control group research design was used for this experimental study. The posttest was administered to all subjects immediately following the end of the study. The posttest was re-administered six weeks later to measure student retention. In addition, a subjective performance assessment taken from the *CMP* curriculum was administered at the end of the study.

The results were analyzed by conducting one-way analysis of variance (ANOVA) procedures. In addition, a two-way mixed-measures ANOVA and two nonparametric analyses, Kruskall-Wallis and Mann Whitney *U* procedures, were conducted. Finally, a subjective analysis of the results of the performance assessment was conducted. In Chapter Four the results of the present study will be reported.

Chapter Four

Results

The purpose of this study was to examine the effects of a metacognitive intervention on the mathematics performance of fifth and sixth grade students. This chapter describes the analysis and interpretation of data generated by testing students at the end of the four-week experimental study. First, the sample of student participants will be described. Second, statistical data related to each of the three research questions that drove this study will be presented. Finally, the results of secondary analyses that were conducted will be reported.

Description of the Sample

The unit of analysis in this study was the elementary school student. A description of the characteristics of the fifth and sixth grade student subjects serves to provide a context for the results associated with the three research questions. General demographic data regarding the sample is displayed on Tables 2 through 4.

As the data in Table 2 indicates, 61 subjects were fifth grade students and 80 were sixth grade students. Of the 141 subjects, 61 were male and 80 were female. The ethnicities of student participants were 121 Caucasian, 10 Asian, 7 Hispanic, and 3 African American. Three students were English Language Learners (ELL), 7 received special education instruction for a specific learning disability, and 5 qualified for free or reduced lunch. Descriptions of the gender and grade level, ethnicity, and academic risk factor distributions are displayed on Tables 2 through 4, respectively. Experimental Group 1 was composed of 16 male and 31 female students, of which 21 were in grade 5 and 26 were in grade 6. The ethnicities of subjects in Experimental Group 1 were:

Caucasian, 39; Asian, 5; Hispanic, 1; African American, 2. One ELL student, 1 special education student, and 2 students receiving lunch subsidy were in Experimental Group 1.

Table 2

Gender and Grade Level by Treatment Group

	Gender		
	Male	Female	Total
Experimental Group I	16	31	47
Experimental Group II	21	27	48
Control Group	24	22	46
Total	61	80	141

	Grade Level		
	5th Grade	6th Grade	Total
Experimental Group I	21	26	47
Experimental Group II	25	23	48
Control Group	16	30	46
Total	61	80	141

Experimental Group II was composed of 21 male and 27 female students, of which 25 were in grade 5 and 23 were in grade 6. The ethnicities of subjects in Experimental Group II were: Caucasian, 40; Asian, 5; Hispanic, 2; African American, 1. One ELL student, 3 special education students, and 1 student receiving lunch subsidy were in Experimental Group II. The Control Group was composed of 24 male and 22 female students, of which 16 were in grade 5 and 30 were in grade 6. The ethnicities of subjects in the Control Group were: Caucasian, 42; Asian, 0; Hispanic, 4; African American, 0. One ELL student, 3 special education students, and 5 students receiving lunch subsidy were in the Control Group.

Table 3

Ethnicity by Treatment Group

	Caucasian	Asian	Hispanic	African American
Statistics with reflection	39	5	1	2
Statistics without reflection	40	5	2	1
Control group	42	0	4	0
Total	121	10	7	3

Results

A one-way analysis of variance (ANOVA) was conducted to evaluate the relationship between the performance of students who received statistics instruction and practiced reflective assessment techniques (Experimental Group I), students who received only the statistics instruction (Experimental Group II) and students who received geometry instruction (Control Group). The independent variable, the reflective assessment techniques, included one level. The dependent variable was the raw score on the statistics test, which served as the posttest in this study (see Appendix F, pp. 124-132).

Table 4

Academic Risk Factor by Treatment

	ELL	Special Education	Free/Reduced Lunch	Total
Experimental Group I	1	1	2	4
Experimental Group II	1	3	1	5
Control Group	1	3	2	6
Total	3	7	5	15

Descriptive statistics, which are displayed on Table 5, were calculated from posttest scores for Experimental Group I ($N = 47$; $M = 29.40$; $Mdn = 31.00$; $SD = 4.33$), Experimental Group II ($N = 48$; $M = 26.92$; $Mdn = 29.00$; $SD = 5.61$) and the Control Group ($N = 46$; $M = 22.30$; $Mdn = 22.00$; $SD = 4.37$). Levene's Test for Equality of Variances showed equal variances among the three groups ($p < .087$). The ANOVA was significant (F (2, 139) = 25.962, p = .000, η^2 = .273; see Table 6). The strength of the relationship between the treatments and the posttest scores, as assessed by η^2, was strong with two statistics treatments accounting for 27% of the variance of the dependent variable. Since the posttest measured performance related to statistics content, it was expected that Experimental Groups I and II would score higher on the posttest, thus the strong effect size. Pairwise comparisons are reported below in relation to each of the three research questions, including further calculations of effect size.

Table 5

Descriptive Statistics: Posttest

Treatment	*Mean*	*Median*	*SD*	*N*
Statistics with Reflection	29.40	31.00	4.33	47
Statistics without Reflection	26.92	29.00	5.61	48
Control Group	22.30	22.00	4.37	46
Total	26.24	27.00	5.61	141

In addition to the one-way ANOVA, a Kruskall-Wallis test was conducted to evaluate the differences between the three treatment groups on the median change on the posttest results. This nonparametric statistical procedure was conducted as a response to the Levene's Test results that approached a significant finding, which would indicate non-homogeneity of groups. The use of nonparametric procedures was also supported by description statistics that revealed leptokurtosis (Ku = 2.45) reflected in the Experimental Group I posttest scores. The Kruskall-Wallis test was significant at the .05 level of confidence, χ^2 (2, N = 141) = 41.95, p = .000, η^2 = .30. As with the effect size of the one-way ANOVA, the η^2 of .30 is potentially misleading, since it does not discriminate between the statistics with and without reflection treatments which were the focus of this study. Pairwise comparisons among the three groups were conducted and are reported related to each of the following research questions.

Table 6

Analysis of Variance: Posttest

Source	*df*	Mean Square	*F*	*p*	η^2
Corrected Model	2	602.538	25.962	.000	.273
Error	138	23.208			

Research Question 1

Research question 1 was " Is there a difference in the mathematics achievement of elementary students who receive statistics instruction and practice reflective assessment techniques, as compared to students who are taught the same curriculum but do not practice reflective assessment techniques?"

Follow-up tests were conducted to evaluate pairwise differences among the means. Post hoc analyses found significant difference at the .05 level of confidence between the scores of Experimental Groups I and II (Tukey *HSD* = $p < .035$; Scheffe = $p < .045$). Since Levene's Test of Equality of Error Variances approached significance, Dunnett's *C* test was conducted and found significant difference at the .05 level of confidence. Effect size calculations were conducted indicating that the magnitude of the difference was of medium size ($d = .496$; $\eta^2 = .058$).

A Mann Whitney *U* test was conducted as a pairwise follow-up to the significant Kruskall-Wallis test. This test found significant difference between the scores of Experimental Group I (statistics with reflection) and Experimental Group II (statistics without reflection), $z = -2.37$, $p = .018$. The statistics with reflection group had a mean rank of 54.74, while the statistics without reflection group had a mean rank of 41.40. Difference in mean rank between groups can serve as an effect size index, according to Green et al. (2000).

Research Question 2

Research question 2 was " Is there a difference between the mathematics achievement of elementary students who receive statistics instruction and practice reflective assessment techniques, as compared to students who receive instruction in a geometry curriculum?"

Follow-up tests were conducted to evaluate pairwise differences among means. Post hoc analyses found significant difference at the .005 level of confidence between the scores of Experimental Groups I and the Control Group (Tukey *HSD* = $p < .000$; Scheffe = $p < .000$). Dunnett's *C* test was also conducted and found significant difference at the .05 level of confidence. Effect size calculations were conducted indicating that the magnitude of the difference was strong, as expected ($d = 1.63$; $\eta^2 = .405$).

A Mann Whitney *U* test was conducted as a pairwise follow-up to the significant Kruskall-Wallis test. This test found significant difference between the scores of Experimental Group I (statistics with reflection) and the Control Group (geometry), $z = -6.279$, $p = .000$. The statistics with reflection group had a mean rank of 64.39, while the control group had a mean rank of 29.27, which is indicative of effect size.

Research Question 3

Research question 3 was " Is there a difference between the mathematics achievement of students who receive statistics instruction without reflective assessment techniques, as compared to students who receive instruction in a geometry curriculum?"

Follow-up tests were conducted to evaluate pairwise differences among the means. Post hoc analyses found significant difference at the .005 level of confidence between the scores of Experimental Groups II and the Control Group (Tukey $HSD = p < .000$; Scheffe $= p < .000$). Dunnett's *C* test was also conducted and found significant difference at the .05 level of confidence. Effect size calculations were conducted indicating that the magnitude of difference was large ($d = .915$; $\eta^2 = .176$).

A Mann Whitney *U* test was conducted as a pairwise follow-up to the significant Kruskall-Wallis test. This test found significant difference between the scores of Experimental Group II (statistics without reflection) and the Control Group (geometry), $z = -4.247$, $p = .000$. As expected, the statistics without reflection group had a mean rank of 41.40, while the control group had a mean rank of 29.27, which is indicative of effect size.

Secondary Analyses

In addition to the analyses related to the three research questions, a retention test and a performance assessment were administered. The retention test was a re-administration of the posttest six weeks following the completion of the study. The performance assessment was administered immediately following the completion of the study. Finally, a two-factor mixed-measures ANOVA was conducted to analyze the posttest and retention test scores.

Retention Test

Two statistical procedures were conducted to analyze data collected on the retention test. First, one-way ANOVA procedures were conducted to compare the mean differences between the three groups. The statistical procedures that were conducted on the posttest were duplicated on the results from the retention test. Second, the nonparametric Kruskall-Wallis test was conducted in order to provide comparison data with the posttest results. While the results of Levene's Test of Equality of Error Variance approached a level of significance for the posttest, those for the retention test found non-homogeneity among the treatment groups (F [2, 134] = 3.277; p = .041). Finally, a two-way mixed-measures ANOVA was conducted to compare the results of the statistics tests, which was administered as both the posttest and the retention test.

One-way Analysis of Variance: Retention Test

To analyze the results of the retention test, a one-way analyses of variance (ANOVA) was conducted to evaluate the relationship between the performance of students who received statistics instruction and practiced reflective assessment techniques (Experimental Group I), students who received only the statistics instruction

(Experimental Group II) and students who received geometry instruction (Control Group The retention test data included the scores of the 137 subjects, all of whom took the both the posttest and the retention test.

Descriptive statistics, which are displayed on Table 7, were calculated from the retention test scores for Experimental Group I (*N* = 44; *M* = 29.18; *Mdn* = 30.00; *SD* = 3.54), Experimental Group II (*N* = 48; *M* = 26.77; *Mdn* = 27.00; *SD* = 5.54) and the Control Group (*N* = 45; *M* = 22.42; *Mdn* = 22.50; *SD* = 4.45).

Table 7

Descriptive Statistics: Retention Test

Treatment	*Mean*	*Median*	*SD*	*N*
Statistics with Reflection	29.18	30.00	3.54	44
Statistics without Reflection	26.77	27.00	5.54	48
Control Group	22.42	22.50	4.45	45
Total	26.12	27.00	5.36	137

The ANOVA was significant, F (2, 134) = 24.606, p = .000, η^2 = .269 (see Table 8). As was reported in relation to the posttest results, the strength of the relationship between Experimental Groups I and II treatments and dependent variable, as assessed by η^2 was found to be strong. However, this is potentially misleading since the focus of this study was on the difference between the scores of the statistics with and without reflection groups. The effect sizes reported in relation to Research Question 1 are of more relevance to the purposes of this study.

A Kruskall-Wallis test was conducted to evaluate the differences between the three treatment groups on the median change on the retention test results. This nonparametric statistical procedure was conducted as a response to the significant Levene's Test results (p = .041) that indicated non-homogeneity of groups. The Kruskall-Wallis test was significant at the .05 level of confidence, χ^2 (2, N = 137) = 41.252, p = .000. Pairwise comparisons among the three groups were conducted and are reported related to each of the following research questions.

Table 8

Analysis of Variance: Retention Test

Source	*df*	Mean Square	*F*	*p*	η^2
Corrected Model	2	524.064	24.606	.000	.269
Error	134	21,299			

Follow-up tests to the significant ANOVA were conducted to evaluate pairwise differences among the means. Related to Research Question 1, post hoc analyses indicated significant difference between the scores of Experimental Groups I and II (Tukey *HSD* = $p < .036$; Scheffe = $p < .047$). In addition, Dunnett's *C* test was conducted and found significant difference at the .05 level of confidence. Effect size calculations were conducted indicating that the magnitude of the difference was of medium size ($d = .514$; $\eta^2 = .05$). As follow-up to the nonparametric Kruskall-Wallis test, a Mann Whitney *U* test was conducted, which found significant difference between the scores of the statistics with reflection group and the statistics without reflection group ($z = -2.292$; $p = .022$). The statistics with reflection group had a mean rank of 92.92 and the statistics without reflection group had a mean rank of 74.36, which are indications of effect size (Green et al., 2000).

Related to Research Question 2, post hoc analyses found significant difference between the scores of Experimental Group I and the Control Group (Tukey $HSD = p < .000$; Scheffe = $p < .000$; Dunnett's $C = p < .05$). Effect size calculations were conducted indicating that the magnitude of the difference was of large size ($d = 1.679$; $\eta^2 = .418$). As follow-up to the significant Kruskall-Wallis test, a Mann Whitney *U* test was also conducted, which found significant difference between the scores of the statistics with reflection group and the control group ($z = -6.257$; $p = .000$). Mean ranks were 92.92 for the statistics with reflection group and 39.89 for the Control Group.

Related to Research Question 3, post hoc analyses found significant difference between the scores of Experimental Group II and the Control Group (Tukey $HSD = p < .000$; Scheffe = $p < .000$; Dunnett's $C = p$ .05). Effect size calculations were conducted indicating that the magnitude of the difference was of large size ($d = .866$; $\eta^2 = .16$). As a follow-up to the significant Kruskall-Wallis test, a Mann Whitney *U* test was conducted, which found a significant difference between the scores of the statistics without reflection group and the control group (*z* = -4.235; *p* = .000). Mean ranks were 74.36 for the statistics without reflection group and 39.89 for the Control Group.

Two-factor Mixed-measures ANOVA

A two-factor mixed-measures ANOVA was conducted to further analyze the posttest and retention test data. This design has the advantage of being able to examine between-subject variables or factors, as well as, within-subject variables or factors (Tabachnick & Fidell, 1996; Turner & Thayer, (2001). The between-subject variable was the experimental treatment with three levels—statistics lessons with reflection, statistics lessons without reflection, and the control group. The within-subject factor was the statistics assessment with two levels—posttest and retention test. This procedure is often referred to in the literature as a "mixed measures" design because the between-subject factor contains independent groups and the within-subject factor contains repeated measures (Turner & Thayer, 2001).

The between-subject results confirmed those of the previously reported one-way ANOVAs. As was the case with the previous analysis involving the retention test data, equality of the treatment groups was an issue with this analysis as reflected in a significant Box's *M* test result (Box's $M = 14.553$; F [6, 432089.1] = 2.373; $p = .027$). As displayed on Table 9, an analysis of estimated marginal means found significant difference between the statistics with reflection treatment and the statistics without reflection treatment ($M_{difference} = 2.697$; *Std. Error* = .909; $p < .004$). Significant difference was also found between the estimated marginal means of the statistics with reflection treatment and the control group ($M_{difference} = 7.178$; *Std. Error* = .924; $p = .000$)

and between the statistics without reflection treatment and the control group ($M_{difference}$ = 4.499; *Std. Error* = .904; *p* = .000).

Table 9

Between Subjects Pairwise Comparisons

(I) Treatment	(J) Treatment	Mean Difference (I-J)	*Std. Error*	*p*
Statistics with reflection	Statistics without reflection	2.679*	.909	.004
	Control group	7.178*	.924	.000
Statistics without reflection	Statistics with reflection	-2.679*	.909	.004
	Control group	4.499*	.904	.000
Control group	Statistics with reflection	-7.178*	.924	.000
	Statistics without reflection	-4.499*	.904	.000

Based on estimated marginal means
* The mean difference is significant at the .05 level.

Post hoc procedures were also conducted on mean differences between the three treatments. Tukey *HSD* procedures found significant differences between the statistics with reflection treatment and the statistics without reflection treatment ($p < .011$), between the statistics with reflection treatment and the control group ($p < .000$), and between the statistics without reflection treatment and the control group ($p < .000$). Scheffe Test procedures found significant differences between the statistics with reflection treatment and the statistics without reflection treatment ($p < .015$), between the statistics with reflection treatment and the control group ($p < .000$), and between the statistics without reflection treatment and the control group ($p < .000$). Dunnett's *C* test for unequal homogeneity of groups found all three pairwise comparisons to be significant at the .05 level of confidence.

Tests of within-subject contrasts found no significant difference between the posttest and retention test results ($N = 137$; Wilks' $\Lambda = .996$; $F = .605$; $p = .438$). There was also no significant interaction found between the test factor and the treatment factor ($N = 137$; Wilks' $\Lambda = .990$; $F = .704$; $p = .496$).

Performance Assessment

A performance assessment was administered to students immediately following the study. The format of these results included short answers, creation of graphs, and required students to show their work (see Appendix H, pp. 145-148). A qualitative analysis of these results showed that students in Experimental Group I provided slightly more accurate information and more detailed responses than did students in Experimental Group II. Consistent with the quantitative analyses, students in Experimental Groups I and II provided much more accurate and detailed information that did students in the Control Group.

Summary

In this chapter, a description of the analysis and interpretation of data collected during the present study has been reported. Included were descriptions of the grade levels, genders, ethnicities, and at-risk factors of the sample and how these were represented in each of the three treatment groups. Both parametric and nonparametric statistical data related to each of the three research questions that drove this study were reported, showing that the reflective assessment intervention resulted in statistically significant posttest achievement results when compared to the other groups. Finally, the results of secondary analyses conducted were reported, which included a retention test, a performance assessment, and a one-way repeated-measures analysis of variance. In Chapter Five, a discussion of these findings, their relevance to the ongoing conversation on school reform, and their practical significance to the classroom will be focused upon.

Chapter Five

Discussion

The purpose of this study was to examine the effects of metacognitive strategies on the academic performance of elementary school students. An additional intent of the study was to apply the theories of reflection or metacognition that have been articulated by Dewey (1933), Bandura (1997), and Flavell (1976) to the classroom setting. Over the last thirty years, there has been a growing accumulation of literature on the topic of metacognition. The emphasis in the literature, however, has been largely upon theory, rather than research. As was stated in the literature review, actual empirical research on the impact of metacognitive interventions on student performance has been scattered across a wide range of related subtopics. It has been the goal of this study to make a contribution to the developing body of research on the effects of reflective or metacognitive strategies and to integrate new knowledge into existing theory and research.

Overview and Discussion of Findings

In the following section the results of the study are discussed in relation to each of the research questions. In addition, the results of secondary analyses that were conducted are considered in relation to their confirmation or refutation of the primary study results.

Research Question 1: Reflection vs. No Reflection

Statistical analyses conducted to determine if the inclusion of a reflective strategy would improve student achievement were statistically significant. Since the study included an experimental design with random assignment, both of the student groups that were compared received the same content, instructional approach, and amount of instructional time. The findings of both parametric (one-way ANOVA) and non-parametric (Kruskall-Wallis) procedures clearly revealed that the inclusion of the metacognitive strategy (I Learned Statement) caused an increase in student performance. Since the independent variable (metacognitive strategy) was directly manipulated in this study, it is appropriate to claim that the reflective assessment technique caused the higher achievement levels among students who practiced it during each lesson.

Research Question 2: Reflection vs. Control Group

As expected, when compared to the control group, which received an alternate and unrelated curriculum to the posttest content, the mean score of students who practiced reflective assessment strategies was significantly higher. Since the control group did not receive instruction on statistics content, which was the substance of the posttest, it would have been surprising if they had out-performed students who were taught statistics content while practicing reflective assessment strategies.

In a posttest-only control group design, the control group serves as a sample of the population for comparison purposes. In this case, it was assumed that the students in the control group performed just as would other students in the population who had not received statistics instruction.

Research Question 3: No Reflection vs. Control Group

A comparison of the test scores of students who received statistics instruction without the addition of reflective assessment strategies and students in the control group resulted in statistical significance. Parallel to the issue presented above, related to Research Question 2, since the control group did not receive instruction in statistics content, their lower performance on the dependent variable measures was expected.

Secondary Analyses

In addition to the statistical procedures that were conducted in relation to the three research questions, three additional analyses were completed. These were included in the study to provide confirmatory results to those found in the primary investigation. An overview and discussion of the retention test, the repeated-measures analysis, and the performance test follows.

Retention After Six Weeks

Six weeks following the completion of the study and posttest assessment, the posttest was re-administered to students to measure their levels of retention. As was expected, the performance of the statistics with and without reflection groups declined over this period of time, however, no significant differences in the amount of decline were found between the two groups. This can be interpreted to mean that while the inclusion of reflective assessment techniques does enhance student performance, it does not appear to enhance long-term memory of the content.

The scores of the control group, however, increased slightly but not significantly on the retention test. Again, this was not surprising since the students in the control group did not have instruction on the content that was measured.

The statistically significant results of the retention test were consistent with those of the posttest. The statistics with reflection group performed significantly higher on the retention test than did the statistics without reflection group, beyond what would be expected by chance. The retention test results echoed the pattern seen in the posttest results, with statistically significant findings on all of the pairwise comparisons examined in the study.

Two-way Mixed-measures ANOVA

The results from the two-way mixed-measures ANOVA indicated that the within subject scores for the posttest and the retention were very similar with means of 26.26 and 26.12, respectively. This analysis showed that there was consistency of performance by subjects on the two test administrations. The between groups results from this analysis were consistent with those found in other statistical analyses conducted in this study. Significantly higher posttest and retention test scores were found in pairwise comparisons between the statistics with reflection group and both the statistics without reflection group and the control group. Significant difference was also found between the posttest and retention test scores of the statistics without reflection and control groups. No significant interaction was found between the two factors, test and treatment, that were examined in this analysis.

Performance Assessment

The results of the performance test appeared to confirm those of the quantitative statistical procedures. Students who received statistics instruction with reflection performed highest, followed in order by students who received statistics instruction without reflection and students who were taught an alternative curriculum. This measure, which was part of the curriculum materials used in the study, was administered at the end of the study on a different day than the posttest administration. The performance test was of a qualitative nature and relied upon the judgment of the rater, which calls to question the reliability and validity of the results. However, it is noteworthy that these results were consistent with the pattern of quantitative findings indicating that reflective assessment techniques are effective in improving student learning.

Effect Sizes

It is worth noting the difference between the posttest performances of students in the two experimental groups. Since effect size indicates the practical significance of findings, these results were of as much interest as were those documenting the exceeding of the alpha level on tests of significance. In addition to reporting overall effect sizes for ANOVAs, effect size calculations were conducted on pairwise comparisons, which provided more precise information about the location and magnitude of differences. In addition, as a response to the issue of homogeneity, effect sizes were reported regarding the results from the nonparametric procedures that were conducted.

The pattern of medium to large effect sizes found between the statistics with reflection and without reflection groups was, perhaps, the most important outcome of this study. The results of this study add empirical evidence to the argument in the metacognitive literature that supports reflective assessment interventions as a best practice. These results provide reason for more research to done on the topic, which could lead to a commonly held belief about the value of student reflection during classroom learning activities.

Homogeneity Issues

As was reported in Chapter Four, the homogeneity of the groups surfaced as an issue in the data analysis of the retention test. Since homogeneity of groups is a fundamental assumption underlying the use of parametric statistical procedures (Green et al., 2000), it was decided to include nonparametric statistical analyses along with the analyses of variance. The related issue of group mortality between the posttest and retention administrations will be included in the discussion on limitations to this study.

Summary of Results Discussion

The results of this study make a strong statement regarding the value of student reflective practice that is seamlessly incorporated into the teaching and learning experience. A pattern of results emerged from this study that consistently showed that students who practiced reflective assessment techniques performed better than students who did not. While there are limitations to this study, that may dampen somewhat its optimistic results, there is reason to speak boldly about the findings.

The primary reason for optimism regarding these results is due to the experimental design of the study. Few empirical studies reported in the literature are of an experimental nature, due to the disruptive impact upon an ongoing school setting. Since this study was conducted in collaboration with a school district pilot of mathematics materials, a rare opportunity was provided to conduct empirical research in the classroom. The tight controls offered by an experimental design make the statistical significance of the findings worthy of attention by both researchers and practitioners.

Limitations

While the statistically significant results of the study offer the promise of meaningful contribution to the body of research on metacognition and of improving classroom practice and student learning, there are also several limitations to this study that deserve attention. These will be discussed in relation to four issues related to the study— the internal validity, the external validity, the measurement of the dependent variable, and the statistical analyses.

Internal Validity

The control of threats to internal validity offered by the experimental posttest-only control-group design is, perhaps, the greatest strength of this study. However, four potential threats to internal validity surfaced through the course of the investigation that may limit the impact of the results. First, the question regarding the homogeneity among the treatment groups needs to be acknowledged and discussed. Posttest measures of homogeneity revealed that the treatment groups were not significantly different; however, the results approached a level of significance. For example, had the alpha level been set at .10, rather than .05, which is sometimes done in program evaluations, the posttest results would not have been homogenous. For the retention test, however, the treatment groups were not found to be homogeneous. The decision to include nonparametric statistical procedures for both posttest and retention test analyses was made in response to the homogeneity issue.

Second, the posttest score distribution for the statistics with reflection group indicated a possible concern with leptokurtosis, or "peakedness of distribution" (Tabachnick & Fidell, 1996, p.71), indicating that the group scores were not normally distributed. This descriptive statistic also supported the inclusion of nonparametric statistical procedures in the data analyses.

Third, mortality of student subjects between the posttest and the retention test administrations, appears to have affected the retention test results. In the six weeks between the posttest and retention test assessments, four subjects moved away and were not available for the retention test. Three of these students were in the statistics with reflection group, each with posttest scores below the group mean, and one student was in the control group, with a posttest score slightly above the group mean. The absence of these scores on the retention test may have compromised the integrity of the random assignment for this phase of the study, which may explain why the treatment groups were found to be non-homogeneous on the retention test. In addition, since the three missing subjects on retention test from the statistics with reflection group each scored below the mean on the posttest, there is a possibility that the absence of their scores on the retention test could have influenced the significance of the findings.

Fourth, the last few minutes of each 50-minute class periods were used differently by the reflective and non-reflective groups. For the reflective group, up to five minutes at the end of each class period was reserved for the reflective assessment activity. Students in the non-reflective group, however, used this time to continue to practice lesson objectives, either individually or with a partner, which was the closing activity in each of the lessons in the study. This meant that over the course of the four weeks students in statistics without reflection group received more practice time than did students in the statistics with reflection group. For the statistics with reflection group, the practice time in each lesson was shortened a few minutes in order to allow time for the reflective assessment activity.

External Validity

This study was conducted in a middle-class to upper-middle class suburban Seattle elementary school. Therefore, any generalization of results must be limited to similar populations of students.

Measurement

The instrument used in the study for the posttest and retention test was researcher-designed, which points to potential questions of reliability and validity. While the 36-item test was drawn from specific learning objectives in the statistics curriculum used in the study and was piloted tested with resulting satisfactory Cronbach's alpha and split-half coefficients, researcher-designed instruments are often fraught with problems. It is only reasonable to identify the instrument as a potential limitation.

The performance assessment, as has already been stated, was a subjective measure, and as such, was a limitation of this study. The reliability and validity of this instrument were unknown. The performance assessment was of less importance to the goals of the study, than was the posttest measure, but it could have provided relevant insights to student performance beyond those inferred by a multiple choice test. Rater reliability problems may have been an issue, due to the absence of a clearly defined scoring guide for this measure.

Statistical Analyses

With the potential problems related to homogeneity and normality of score distribution that surfaced during the statistical analysis of the data, it was decided to supplement the original analysis of variance procedures with nonparametric procedures. While the Kruskall-Wallis and Mann Whitney *U* procedures provided confirmatory findings to those of the one-way ANOVAs, and ruled out the problem of non-homogeneity, their inclusion also resulted in a large amount of statistical data and analysis that could be confusing to a reader.

Recommendations for Future Research

The answers to the three research questions presented in this dissertation make a small contribution to the body of empirical research on the effects of metacognitive strategies on achievement. This investigation, however, has also raised many related questions that might be followed-up on in future research. Three recommendations follow that are of particular importance.

First, further research needs to be done on the effects of a range of reflective assessment strategies in a variety of settings. Students at different grade levels, and using different curricula beyond mathematics, should be studied.

Second, research on the effects of reflective strategies should be conducted with a more diverse sample than occurred in this study. Further research is needed to determine the effects of classroom reflection and the appropriateness of different strategies with students who possess at-risk indicators such as ethnicity, ELL, special education, and poverty.

Finally, it is recommended that researchers interested in conducting experimental research seek to partner with school districts as they conduct curriculum pilots. As was demonstrated in this study, the mathematics pilot provided a vehicle through which the experimental research could be conducted with a minimum of disruption to the school operation. The benefits to a school system of having a researcher lead such a program evaluation has the potential of outweighing the drawbacks inherent in implementing random assignment of students during regular school session.

Implications for Classroom Practice

The implications drawn from this study for the classroom practice of both teachers and students are potentially powerful. As was discussed above related to effect sizes, the practical significance of the results have value for practitioners at the elementary school level, and likely, too, at other levels. The results inform teachers that review, closure, and other reflective activities during a lesson have a positive effect on what students learn. As informal classroom assessments become more strongly emphasized in relation to reaching high standards, the results from this study verify that they are, indeed, useful practice. The provision of staff development opportunities for teachers regarding the use of reflective assessment techniques and other metacognitive approaches will be important for the full impact of this approach to be realized. Therefore, in addition to the classroom teacher, principals, staff developers, central office administrators, and college professors need to be made aware of the effectiveness of bringing reflection to the center of the classroom experience.

For students, reflective assessment has as much to do with ownership of the learning as it does with the mechanics of a particular strategy. As has been emphasized in this paper, if learning is basically a constructivist activity, then students have always owned their learning. However, making students fully aware of this dynamic—leading them to "think about thinking"—is something that teachers and parents must do. As students become increasingly mature as thinkers, they too, need to independently attend to the value of reflective assessment techniques.

Concluding Remarks

From the early days of Plato, on through the ages to the time of John Dewey, and to the present day an individual's thinking about his or her own thinking has been of high interest. The sustained attention to this metacognitive concept over time is an indication to its importance as people have strived to understand the workings of the internal self. In this dissertation, I have drawn upon such deep theoretical roots as a basis for a study of student reflective practice in a classroom setting.

In addition to the theoretical underpinnings of this research, I have been guided by a rich and expanding body of literature on metacognition. As has been stated, the preponderance of this literature is theoretical in nature, rather than empirically based. However, there is a rapidly developing body of empirical research that supports the use of metacognitive or reflective strategies. While I narrowed my inclusion of studies to those that deal specifically with the effects of metacognitive strategies on learning, these were drawn from only one segment of the research literature on the broad topic of metacognition.

The results of this study lend support to the theoretical view that individual student reflection on a concept enhances the chances that the student will learn the concept. The results also provide strong rationale for the incorporation of reflective assessment strategies into the daily activities of the classroom. Consistent with the recently revived momentum toward empirically based research to inform classroom practice, the results of this study offer statistical evidence that such an approach works. It also demonstrates through the classroom interventions how it might be done.

If schools in the United States are to help all students reach high standards, as has been central to the standards movement over the last 20 years, it is essential that the ownership of the learning be placed in the hands of the learners. Reflective assessment approaches have the potential to support this goal, for they can be naturally integrated into the classroom experience. Rather than receiving feedback on what is learned or not learned through external channels, such as tests and teacher praise or criticism, reflective assessment does not rely on outside information to guide learning. The student is the locus of control of the learning, rather than the teacher, the curriculum, or the written test. Even though there is a vital role for teacher guidance on what is reflected upon, since we do not want students focusing upon incorrect or irrelevant concepts, teacher-directed feedback has been over-emphasized in today's public schools. Dewey's vision that the teacher should be the student's "co-partner and guide" has, perhaps, more relevance in today's standards-based environment than it did decades ago when he wrote it.

The continued pressures on students and teachers to improve test scores, while usually well intended, are misguided when they fail to acknowledge the central role of the student. It is time to fully embrace student participation and ownership in a seamless teaching, learning, and assessment process, for in reality they have always owned it. Reflective assessment is an innovation whose time has come.

References

Allen, B. A., & Armour-Thomas, E. (1992). Construct validation of metacognition. *The Journal of Psychology, 127* (2), 203-211.

Andrade, H. G. (1999). *Student self-assessment: At the intersection of metacognition and authentic assessment.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28* (2), 117-148.

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

Biggs, J. B. (1985). The role of metalearning in study processes. *British Journal Educational Psychology, 55*, 185-212.

Biggs, J. B. (1988). The role of metacognition in enhancing learning. *Australian Journal of Education, 32*, 127-138.

Black, P., & Wiliam, D. (1988). Inside the black box. *Phi Delta Kappan, 80* (2), 139-148.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill.

Borkowski, J. G. (1992). Metacognitive theory: A framework for teaching literacy, writing, and math skills. *Journal of Learning Disabilities, 25* (4), 253-257.

Boud, D. J. (1995). *Enhancing learning through self-assessment.* London: Kogan Page.

Boud, D., Keogh, R., & Walker, D. (1985). Promoting reflection in learning: A model. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 18-40). New York: Nichols Publishing Company.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (1999). *How people learn: Brain , mind, experience, and school*. Committee on Developments in the Science of Learning, Commission on Behavioral and Social Sciences and Education, National Research Council (p. xii). Washington, DC: National Academy Press.

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education, 8* (2), 153-169.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Brown, A. L. (1987). Control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert and R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Brown, A. L., & Barclay, C. R. (1976). The effects of training specific mnemonics on the metamnemonic efficiency of retarded children. *Child Development, 47* (1), 71-80.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Buehl, D. (1996). Improving students' learning strategies through self-reflection. *Teaching & Change, 3* (3), 227-244.

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

Campione, J. C. (1987). Metacognitive components of instructional research with problem learners. In F. E. Weinert and R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 117-140). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Candy, P., Harri-Augstein, S., & Thomas, L. (1985). Reflection and the self-organized learner: A model of learning conversations. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 100-116). New York: Nichols Publishing Company.

Conzemius, A., & O'Neill, J. (2001). *Building shared responsibility for student learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Cornoldi, C. (1998). The impact of metacognitive reflection on cognitive control. In G. Mazzoni & T. Nelson (Eds.), *Metacognition and cognitive neuropsychology: Monitoring and control processes* (pp. 139-159). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

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

Costa, A. L., & Kallick, B. (2000). *Developing and exploring habits of mind.* Alexandria, VA: Association for Supervision and Curriculum Development.

Desoete, A., Roeyers, H., & Buysse, A. (2001). Metacognition and mathematical problem solving in grade 3. *Journal of Learning Disabilities, 34* (5), 435-449.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* New York: D. C. Heath.

Dewey, J. (1964). The need for a philosophy of education. In R. D. Archambault (Ed.), *John Dewey on education: Selected writings* (pp. 3-14). Chicago: University of Chicago Press.

Dickson, S. V., Collins, V. L., Simmons, D. C., & Kameenui, E. J. (1998). Metacognitive strategies: Research bases. In D. C. Simmons & E. J. Kameenui (Eds.), *What reading research tells us about children with diverse learning needs: Bases and basics* (pp. 295-360). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Earl, L. M., & LeMahieu, P. G. (1997). Rethinking assessment and accountability. In A. Hargreaves (Ed.), *Rethinking educational change with heart and mind: 1997 ASCD yearbook* (pp. 149-167). Alexandria, VA: Association for Supervision and Curriculum Development.

Elen, J., & Lowyck, J. (1998). Metacognitive instructional knowledge: Cognitive mediation and instructional design. *Journal of Structural Learning and Intelligent Systems, 13* (3-4), 145-169.

Ellis, A. K. (2001). *Teaching, learning, & assessment together.* Larchmont, NY: Eye on Education, Inc.

Ellis, E. S., & Worthington, L. A. (1994). *Research synthesis on effective teaching principles and the design of quality tools for educators* (Technical Report No. 5 for the National Center to Improve the Tools of Educators). Eugene, Oregon: University of Oregon.

Ellison, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development, 14* (4), 272-278.

Feuerstein, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability.* Baltimore: University Park Press.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34* (10), 906-911.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, Jr. and J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman Publishers USA.

Garofalo, J., & Lester, Jr., F. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education, 16* (3), 163-176.

Gredler, M. E. (2001). *Learning and instruction: Theory into practice* (4th ed.). Upper Saddle River, NJ: Merrill Prentice-Hall.

Gibbs, G. & Northedge, A. (1979). Helping students to understand their own study methods. *British Journal of Guidance and Counselling, 7* (1), 91-100.

Green et al. (2000)

Habermas, J. (1996). Dogmatism, reason, and decision. In W. Outhwaite, (Ed.), *The Habermas Reader*. Cambridge, UK: Blackwell Publishers, Ltd.

Hacker, D. J. (1998). Definitions and empirical foundations. In D. Hacker, J. Dunlosky, and A. Graesser (Eds.), *Educational theory and practice* (pp. 1-23). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Haney, W. (1991). We must take care: Fitting assessments to functions. In V. C. Perrone (Ed.), *Expanding student assessment* (pp. 142-163). Alexandria, VA: Association for Supervision and Curriculum Development.

Hartlep, K., & Forsyth, A. (2000). The effect of self-reference on learning and retention. *Teaching of Psychology, 27* (4), 269-271.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22* (3-4), 255-278.

Jans, V., & Leclercq, D. (1997). Metacognitive realism: A cognitive style or a learning strategy? *Educational Psychology, 17* (1-2), 101-111.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83* (3), 307-317.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31* (2), 338-368.

Kirkwood, M. (2000). Infusing higher-order thinking and learning to learn into content instruction: A case study of secondary computing in Scotland. *Journal of Curriculum Studies, 32* (4), 509-535.

Kurtz, B. E., & Borkowski, J. G. (1984). Children's metacognition: Exploring relations among knowledge, process, and motivational variables. *Journal of Experimental Child Psychology, 37* (2), 335-354.

Lappan, G., Fey, J., Fitzgerald, W., Friel, S., & Phillips, E. (1998). *Connected mathematics curriculum: Grade 6.* Menlo Park, CA: Dale Seymour Publications.

Larkin, E., & Chabay, R. (1989). Research on teaching scientific thinking: Implications for computer-based instruction. In L. Resnick and L. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research; 1989 ASCD yearbook* (pp. 150-172). Alexandria, VA: Association for Supervision and Curriculum Development.

Loughran, J. J. (1996). *Developing reflective practice: Learning about teaching and learning through modeling.* New York: Falmer Press.

Main, A. (1985). Reflection and the development of learning skills. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 91-98). New York: Nichols Publishing Company.

Malpass, J. R., O'Neil, Jr., H. F., & Hocevar, D. (1999). Self-regulation, goal orientation, self-efficacy, worry, and high-stakes math achievement for mathematically gifted high school students. *Roeper Review, 21* (4), 281-290.

Manning, B. H., & Payne, B. D. (1996). *Self-talk for teachers and students: Metacognitive strategies for personal and classroom use*. Boston: Allyn and Bacon.

Meichenbaum, D. (1977). *Cognitive behavior modification: An integrative approach.* New York: Plenum Press.

Montague, M. (1992). The effects of cognitive and metacognitive strategy instruction on the mathematical problem solving of middle school students with learning disabilities. *Journal of Learning Disabilities, 25* (4), 230-248.

Moshman, D. (1982). Exogenous, endogenous, and dialectical constructivism. *Developmental Review, 2*, 371-384.

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention in improving arithmetic computation based on the PASS theory. *Journal of Learning Disabilities, 33* (6), 591-598.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory (Volume 26) (pp. 125-173)*. New York: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing About Knowing* (pp. 1- 25). Cambridge, MA: The MIT Press.

O'Neill, P. (1998). From the writing process to the responding sequence: Incorporating self-assessment and reflection in the classroom. *Teaching English in the Two-Year College, 26* (1), 61-70.

Ornstein, A. C., & Hunkins, F. P. (1998). *Curriculum: Foundations, principles, and issues* (3rd ed.). Needham Heights, MA: Allyn and Bacon.

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 25* (1), 21-38.

Osman, M. E., & Hannafin, M. J. (1994). Effects of advance organizers and prior knowledge on science learning. *Journal of Educational Research, 88* (1), 1-13.

Palincsar, A. S., & Brown, D. A. (1987). Enhancing instructional time through attention to metacognition. *Journal of Learning Disabilities, 20* (2), 66-75.

Palincsar, A. S., & Brown, A. L. (1989). Instruction for self-regulated reading. In L. Resnick and L. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research: 1989 ASCD yearbook* (pp. 19-39). Alexandria, VA: Association for Supervision and Curriculum Development.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36* (2), 89-101.

Paris, S. G., & Winograd, P. (1999). The role of self-regulated learning in contextual teaching: Principles and practices for teacher preparation. Contextual teaching and learning: Preparing teachers to enhance student success in the workplace and beyond (Information Series No. 376). Washington, DC: ERIC Clearinghouse on Adult, Career, and Vocational Education.

Pearson, M., & Smith, D. (1985). Debriefing in experience-based learning. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 69-84). New York: Nichols Publishing Company.

Perrone, V. C. (1994). How to engage students in learning. *Educational Leadership, 51* (5), 11-13.

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher, 24* (7), 5-12.

Piaget, J. (1976). *The grasp of consciousness: Action and concept in the young child.* (Translated by S. Wedgwood). Cambridge, MA: Harvard University Press.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated components of classroom academic performance. *Journal of Educational Psychology, 82* (1), 33-40.

Plato. (trans. 1956). Theatetus. In E. Hamilton & H. Cairns (Eds.), *The collected dialogues of Plato.* Princeton, NJ: Princeton University Press.

Powell, J. P. (1985). Autobiographical learning. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 43-51). New York: Nichols Publishing Company.

Powell, S. D., & Makin, M. (1994). Enabling pupils with learning difficulties to reflect on their own thinking. *British Educational Research Journal, 20* (5), 579-594.

Pressley, M., & El-Dinary, P. B. (Eds.). (1993). *The elementary school journal, 94* (2), 105-254.

Pressley, M. Borkowski, J., & O'Sullivan, J. (1984). Memory strategy instruction is made of this: Metamemory and durable strategy use. *Educational Psychologist, 19* (2), 94-107.

Resnick, L., & Klopfer, L. (1989). Toward the thinking curriculum: An overview. In L. Resnick and L. Klopfer (Eds.), *Toward the thinking curriculum: Current* *cognitive research; 1989 ASCD yearbook* (pp. 1-18). Alexandria, VA: Association for Supervision and Curriculum Development.

Robinson, D. H., & Kiewa, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology, 87* (3), 455-467.

Sadler,

Schneider, W., Borkowski, J. G., Kurtz, B., & Kerwin, K. (1986). Metamemory and motivation: A comparison of strategy use and performance in German and American children. *Journal of Cross-Cultural Psychology, 17* (3), 315-336.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematical education* (pp. 185-215). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. Resnick and L. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research; 1989 ASCD yearbook* (pp. 83-103). Alexandria, VA: Association for Supervision and Curriculum Development.

Schunk, D. H. (1983). Progress self-monitoring: Effects on children's self-efficacy and achievement. *Journal of Experimental Education, 51* (2), 89-93.

Schunk, D. H. (1984). Self-efficacy perspective in achievement behavior. *Educational Psychologist, 19* (1), 48-58.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation.* Chicago: Rand McNally.

Scriven, M. (1991). *Beyond formative and summative evaluation.* Chicago: University of Chicago Press.

Schraw, G., Dunkle, M., Bendixen, L., & Roedel, T. (1995). Does a general monitoring skill exist? *The Journal of Educational Psychology, 87* (3), 433-444.

Scruggs, T., Mastropieri, M., Monson, J., & Jorgenson, C. (1985). Maximizing what gifted students can learn: Recent findings of learning strategy research. *Gifted Child Quarterly, 29* (4), 181-185.

Shepard, L. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20* (7), 2-16.

Simmons, R. (1994). The horse before the cart: Assessing for understanding. *Educational Leadership, 51* (5), 22-23.

Simon, et al. (1972)

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, D.C.: National Academy Press.

Soodak, L. C., & Martin-Kniep, G. O. (1994). Authentic assessment and curriculum integration: Natural partners in need of thoughtful policy. *Educational Policy, 8* (2), 183-202.

Starnes, B. A., & Paris, C. (2000). Choosing to learn. *Phi Delta Kappan, 81* (5), 392-398.

Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills*. New York: Harcourt Brace Jovanovich.

Stewart, J., & Landine, J. (1995). Study skills from a metacognitive perspective. *Guidance & Counseling, 11* (1), 16-20.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77* (3), 238-245.

Stiggins, R. J. (1996). *Opening the doors to excellence in assessment.* Portland, OR: Assessment Training Institute, Inc.

Stiggins, R. J. (1997). *Student-centered classroom assessment.* Upper Saddle River, NJ: Merrill Prentice Hall.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins College Publishers.

Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Toward transparency for students and tutors. *Assessment & Evaluation in Higher Education, 26* (6), 605-614.

Tittle, C. K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. *Educational Psychologist, 29* (3), 149-162.

Turner, J. R., & Thayer, J. F. (2001). *Introduction to analysis of variance: Design, analysis, & interpretation.* Thousand Oaks, CA: Sage Publications.

Urbaniak, G. C., Plous, S., & Lestik, M. (2002). *Research randomizer.* Retrieved October 23, 2002, from http://www.randomizer.org/about.htm.

Van Reusen, A. K., & Head, D. N. (1994). Cognitive and metacognitive interventions: Important trends for teachers of students who are visually impaired. *Re:View, 25* (4), 153-163.

Vogt, 1999

von Glasersfeld, E. (1991). Introduction. In E. von Glasersfeld, (Ed.), *Radical constructivism in mathematics education* (xiii-xx). Dordrecht, The Netherlands: Kluwer Academic Publishers.

von Glasersfeld, E. (2000). *Radical constructivism and teaching.* Retrieved February 22, 2003, from *http://www.umass.edu/srri/vonGlasersfeld/onlinePapers/html/geneva/*.

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

Walker, D. (1985). Writing and reflection. In D. Boud, R. Keogh, and D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 52-68). New York: Nichols Publishing Company.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63* (3), 249-294.

Weiner, B. (1980). *Human motivation.* New York: Holt, Rinehart, and Winston.

Weiner, B. (1986). *An attribution theory of motivation and emotion.* New York: Springer-Verlag.

Weinert, F. E. (1987). Introduction and overview: Metacognition and motivation as determinants of effective learning and understanding. In F. E. Weinert and R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 1-19). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Weinstein, C. E., & Van Mater Stone, G. (1993). Broadening our conception of general education: The self-regulated learner. *New Directions for Community Colleges, 21* (1), 31-39.

Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. *Educational Technology Research and Development, 43* (3), 29-48.

White, C. (1999). The metacognitive knowledge of distance learners. *Open Learning,* November, 37-46.

Wiens, J. W. (2001). Metacognition and the adolescent passive learner. *Journal of Learning Disabilities, 16* (3), 144-149.

Wiggins, G. (1993). *Assessing student performance: Exploring the purpose and limits of testing.* San Francisco: Jossey-Bass Publishers.

Williams, M. (2000). The part which metacognition can play in raising standards in English at key stage 2. *Reading,* 4, 3-8.

Wilson, R. C. (1986). Improving faculty teaching: Effective use of student evaluations and consultants. *Journal of Higher Education, 57* (2), 197-211.

Wragg, T. (1997). *Assessment & learning: Primary and secondary.* New York: Routlege.

Yancey, K. B. (1998). Reflection, self-assessment, and learning. *Clearing House, 72* (1), 13-18.

Young, R. E. (1990). *A critical theory of education: Habermas and our children's future.* New York: Teacher's College Press.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81* (3), 329-339.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25* (1), 3-17.

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*, 614-628.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82* (1), 51-59.

Appendix A

Request to Conduct Research in School District

September 11, 2002
John B. Bond
20015 – 65th Ave. SE
Snohomish, WA, 98296

Bob Hamilton, Director of Evaluation, Assessment, and Research
Northshore School District
18315 Bothell Way NE
Bothell, WA 98011

Dear Dr. Hamilton:

Please accept my application to conduct research in the Northshore School District. I am a doctoral student at Seattle Pacific University and will be completing my dissertation during the upcoming year. It is my desire to conduct this research at the fifth and sixth grade level in a Northshore elementary school.

As a Northshore staff member it is important to me for this research to serve the district. With this in mind, I have designed a study that is built around a pilot of Connected Mathematics curriculum, which is our current junior high adoption. The focus of my study will be on the effect of student reflection on mathematics achievement.

Enclosed with this letter are responses to the questions regarding research proposals. For your reference I have also included a copy of a paper on reflective practice that I presented to SPU doctoral faculty in July 2002. This paper includes a description of the research area, a literature review, and a discussion of the theoretical bases for the research.

I look forward to an opportunity to meet with you and the Testing and Research Committee to discuss this proposed study.

Sincerely,

John B. Bond

September 11, 2002

To: Bob Hamilton, Director of Evaluation, Assessment and Research
From: John Bond
Re: Research Proposal

As culmination to a doctoral degree at Seattle Pacific University, I am planning to conduct research regarding the effect of reflective practice on student achievement. This study has been designed to assist Northshore teachers and when complete will provide new information about both the use of reflective instructional strategies and the appropriateness of Connected Mathematics curriculum at the fifth and sixth grade levels.

A brief response to the following questions will provide you with an overview of my proposed research. I look forward to discussing this further with you and the Testing and Research Committee.

1. *In what way(s) is your research compatible with Northshore School District's philosophy and goals?*

My research will focus upon how to increase student retention through student use of reflective practice strategies. Students taking responsibility for their own learning underpins the concept of reflective practice. Guiding students to become independent and lifelong learners is a priority in our district.

2. *Explain the goals and outcomes of your research.*

My goal for this study is to add to the body of empirical research regarding the role and effects of metacognitive strategies. While there has been a great deal of research done on reflection with samples of adults and college students, there have been few empirical studies conducted with K-12 students. It is my intent to contribute to filling this research gap.

3. *Would you describe your research as serving theoretical or applied purposes?*

This research will serve both theoretical and applied purposes. The basis of my study will be theoretical, in that it is related to the effectiveness of metacognitive practice. In conducting the study, however, I will be supporting a district need by investigating the appropriateness of *Connected Mathematics* curriculum in fifth and sixth grades.

4. *How does your research serve current or projected need?*

There have been concerns regarding the appropriateness of some components of our current math adoption for the upper elementary grades. This research will allow fifth and sixth grade teachers in one elementary to try out two instructional units from our 7th grade math curriculum. It will provide a platform for further discussion regarding math curriculum.

5. *What benefits do you believe will result from your research?*

At the minimum, I am confident that this research will benefit the student participants by helping them to develop reflective strategies that will serve them in all academic activities. Hopefully, though, the results will demonstrate a causal relationship between reflective practice and achievement. Should this be the case, it will be a potentially significant contribution to the body of research on this construct.

6. *If requested, will you provide a bibliography of literature and documentation regarding how others have approached this particular topic?*

Yes, I will gladly share any and all of my research materials. Included with this memo is a copy of a paper that I presented in July 2002 that includes a literature review.

7. *Are there legal implications to your research? If so, please describe them.*

There are no legal implications related to this proposed research, other than the need to maintain confidentiality of participants.

8. *Have you taken the legal and humanitarian rights of your subjects into consideration? Please include an example of that process in your proposal summary.*

As a part of the dissertation approval process at SPU, I will be submitting this project to a committee that reviews legal and ethical issues related to research. I will be adhering to American Psychological Association (APA) standards regarding conducting research within a school setting.

9. *Please provide reliability and validity estimates (cite your source) for any tests or other instrumentation you plan to use.*

In collaboration with my dissertation committee chairperson, Dr. Arthur Ellis, I have designed a criterion-referenced test. Reliability assessments will be conducted this fall and will guide refinement of the instrument. Cronbach's Alpha and split-half analyses will be conducted to determine internal validity.

10. *What kind of data are you collecting?*

This is a posttest-only research design. Data will be collected from the test to be administered at the end of the mathematics units.

11. What is your hypothesis? How do you plan to test it?

My research hypothesis is that students who use reflective strategies in math lessons will score significantly higher on the posttest than do students who do not use reflective practices. This hypothesis will be tested using an analysis of variance (ANOVA) statistical procedure.

12. What financial support do you have to conduct this research? What is the source of that support?

I have no financial support for this research.

13. Are you supported in this effort by any organizations? If so, please list them.

My primary supports in this project have been the doctoral faculty at Seattle Pacific University and administrators in our Teaching and Learning Department.

14. How do you plan to publish and use the results of this study?

Should significant findings result from this research, I will consider submitting an article to a professional journal.

15. Do you plan to identify Northshore School District or any of its employees or students by name in any publication?

Consistent with APA guidelines, I will not be identifying the school district, employees, or students by name in the dissertation or any further publication.

16. Will you agree to submit a final copy of your results to the TARC committee prior to any publication?

I will gladly submit a copy of any publication that may result from this research to TARC prior to publication.

17. How much time and effort will your study require for staff and students in Northshore? The Northshore School District will not provide clerical help or use of office equipment for your study.

My intent is that no additional time or effort will be required of staff and students who participate in this study. The math curriculum that will be used for the study is drawn from the adopted 7th grade curriculum and will take the place of corresponding Everyday Math lessons during the course of the four-week unit. Lesson plans and materials will be provided for the participating teachers.

Appendix B

Letter Sent to Parents of Student Participants

October 9, 2002

Dear Parents of Fifth and Sixth Graders:

An ongoing part of maintaining the high standards of the Northshore School District is the regular review and study of curriculum. Occasionally, when new curriculum is being considered, schools participate in pilot studies of the new materials. In the month of November, _______ Elementary School fifth and sixth graders will participate in the piloting of a new math program that is being considered.

Connected Mathematics, which is the current math program in Northshore School District 7th and 8th grades, is being considered for use in the upper intermediate grades. Like Everyday Math, the Connected Mathematics Program is tightly aligned with district and state standards, which helps to prepare students for the WASL.

Fifth and sixth graders will be mixed into six instructional groups for daily lessons of 45 minutes. Each of the 5th and 6th grade teachers will be participating. The Connected Mathematics units will begin in early November and will be complete prior to Thanksgiving Vacation.

This experience will familiarize students with a curriculum which they will be using when they enter seventh grade. The lessons are taken from the sixth grade component of this series. The overlap in the lesson content of Everyday Math and the pilot lessons chosen from the Connected Mathematics curriculum, will ensure that students do not miss essential parts of their regular math program.

In Northshore, we are striving to improve an already strong math program. The transition from elementary to junior high is one of our focuses. The teachers will be able to provide our curriculum and instruction leaders valuable insights regarding how this curriculum compares with the current Everyday Math program and how intermediate students respond to this new program.

We want to assure you that students will benefit from this experience while providing teachers, administrators, and curriculum & instruction leaders with valuable information about the Connected Mathematics curriculum. If you have questions or would like additional information, please contact either the school principal or John Bond at (425) 489-6424.

Sincerely,

_____________, Principal

John Bond, Executive Director of Elementary Education

Appendix C

Overview of Reflective Strategies

Overview of the Reflective Assessment Strategies

The independent variable for this study is the reflective assessment techniques that students in Experimental Group I will practice each day. The same sequence of reflective assessment strategies will be practiced following each of the sixteen statistics lessons. Students will first independently write an "I Learned" statement. Second, in pairs, students will practice the "Think Aloud" strategy, in which they will discuss their written statements. Finally, students will individually write a second "I Learned" statement.

The "I Learned" statements will be collected by the teacher at the end of each class period. Teachers will be provided comments that acknowledge student participation that will be written on student papers before they are returned.

"I Learned" Statements

Description:	*I Learned Statements* are statements of personal learning that are written by learners during the closure of a lesson (Ellis, 2001).
Significance:	This technique facilitates students in reflecting upon what they have learned. It also provides teachers with feedback about what students see as significant, as well as, how successful teachers have been in attaining their instructional objectives.
Implementation:	As part of the closure to each lesson during the study, students in Experimental Group I will write an *I Learned Statement.* Following a short discussion with a partner (the *Think Aloud* strategy), students will write a second *I Learned Statement.* The *I*

Learned Statements will be turned in to the teacher at the end of each class period.

"Think Aloud" Strategy

Description: The *Think Aloud* strategy is a technique in which students verbalize their thinking as they solve problems.

Significance: It is an extension of a reflective strategy that Ellis (2001) refers to as *Talk About It.* By verbalizing their thoughts students "...create a self-feedback mechanism, test one's ideas in public, and make thought processes more intentionally deliberate" (Ellis, 2001, p. 89).

Implementation: In this study the *Think Aloud* strategy will be practiced by students in Experimental Group I by discussing their written *I Learned Statements* with a partner. As explained above, for the purpose of this study, the *Think Aloud* strategy will be used in conjunction with *I Learned Statements.*

Appendix D

Daily Lesson Log

DAILY LESSON LOG

Student Group ______________________________ Date ________________________________

Lesson description	
Attendance (students absent)	
Duration of lesson	Lesson began at: ____________________ Lesson ended at: ____________________
Notes (What went well and not so well ?)	
Comments and Questions	

What was the overall student participation of the class on this lesson?

1	2	3	4	5
Low		Moderate		High

Teacher Signature: ____

Appendix E

Observation by Experimenter

OBSERVATION OF MATH LESSONS

Group ______________________________ **Date** ____________________

Teacher ______________________________

Lesson Description	
Participation Level	
Attention to Lesson Script	
Confounding/intervening Variables	
Classroom Climate	

Experimenter Signature: ____

Appendix F

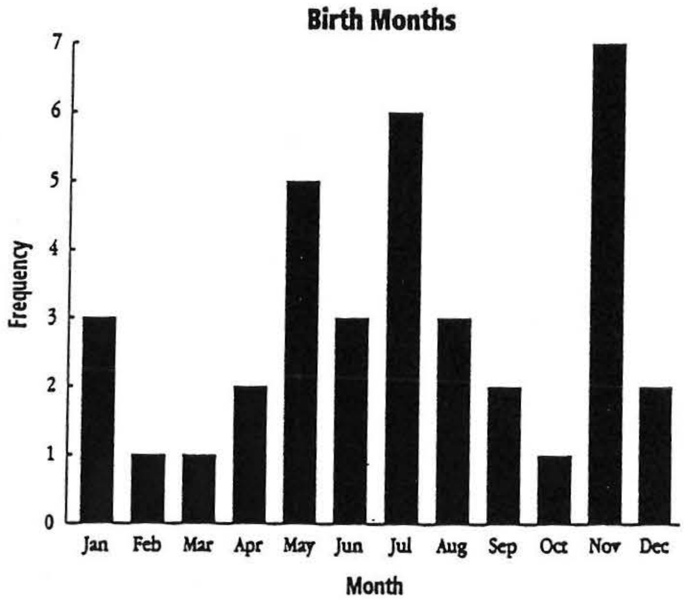
Posttest and Retention Test Instrument

Name ______________________________ Date ________________

1. If a student had six red pencils, four blue pencils, and one green pencil, how many pencils would she have?
   - A. 8
   - B. 15
   - C. 11
   - D. 5

2. A definition of statistics is the study of __________.
   - A. weather
   - B. spelling
   - C. insects
   - D. numbers

For questions 3, 4, and 5, refer to the chart that follows:



3. The distribution showing Birth Months is called a
   - A. bar graph
   - B. data chart
   - C. line plot
   - D. stem and leaf plot

4. The month in which the most births occur is
   - A. November
   - B. May
   - C. January
   - D. July

5. On the Birth Month chart, the *mode* is

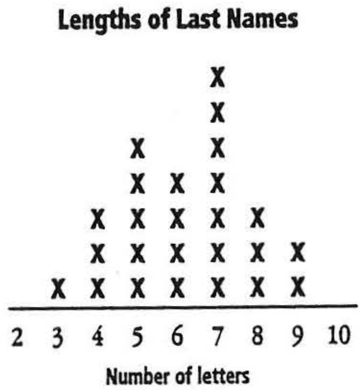
A. October
B. September
C. November
D. May

---

6. Data that are words are called what kind of data?

A. numerical data
B. alphabetical data
C. categorical data
D. range data

For questions 7, 8, and 9, refer to the following chart:



7. The *range* for the length of last names shown on the chart is

A. 10
B. 9
C. 4
D. 6

8. The "typical" name length for this class of students?

A. 2
B. 7
C. 10
D. 4

9. On the Lengths of Last Names chart, the *median* is

A. 6
B. 3
C. 7
D. 9

For questions 10 through 12 see the Class Test Scores data below.

Class Test Scores

0	5
1	
2	4
3	4 9
4	3 7 8
5	7 9
6	1 6 8
7	3 5 6 8 8
8	1 2 2 2 5
9	0 3 9

10. The Class Test Scores chart is called a

A. bar graph
B. distribution
C. stem plot
D. line plot

11. What is the *median* of the Class Test Score data?

A. 48
B. 73
C. 57
D. 5

12. What is the *outlier* score in the Class Test Score data?

A. 99
B. 5
C. 24
D. there are none

---

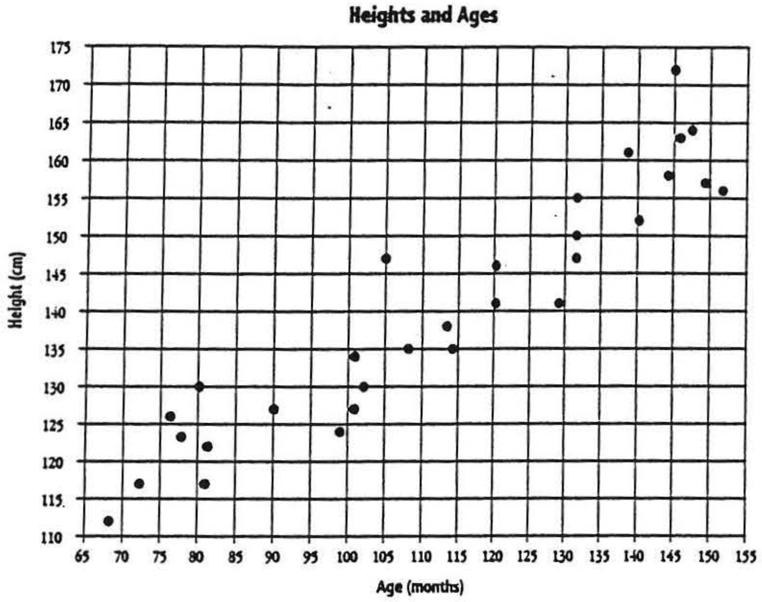
13. A group of 9 students has these numbers of children in their families:
3, 2, 4, 2, 1, 5, 1, 2, and 7. The *median* number of children in the 9 families is

A. 5
B. 3
C. 2
D. 6

14. A distribution of student heights in centimeters would be what kind of data?

A. numerical data
B. measurement data
C. categorical data
D. descriptive data

For questions 15 through 18, refer to the following graph of student heights and ages.



15. How many students are taller than 155 cm.?

A. 8
B. 0
C. 7
D. 11

16. The age in months is found on what axis?

A. diagonal axis
B. vertical axis
C. y-axis
D. horizontal axis

17. Students who are taller, also tend to be...

A. older
B. younger
C. shorter
D. smarter

18. How old in months is the student whose height is 152 cm.?

A. 157 months
B. 132 months
C. 145 months
D. 140 months

For questions 19, 20, and 21, use the following data: A group of 8 students has these numbers of pets in their families: 2, 1, 4, 0, 1, 2, 4, and 2.

19. What is *mean* number of pets in the 8 families?

A. 3
B. 2.3
C. 2
D. 4

20. What is the *mode*?

A. 2
B. 1
C. 0
D. 4

21. What is the *median*?

A. 1
B. 2
C. 3
D. 4

For questions 22-25 refer to the following distribution

Class Test Scores

0	5
1	
2	4
3	4 9
4	3 7 8
5	7 9
6	1 6 8
7	3 5 6 8 8
8	1 2 2 2 5
9	0 3 9

22. What is the *mean* class test score?

A. 50
B. 82
C. 79
D. 65

23. What is the format of this data?

A. line plot
B. bar graph
C. data chart
D. stem and leaf plot

24. What is the *mode* of this data?

A. 57
B. 82
C. 8
D. 78

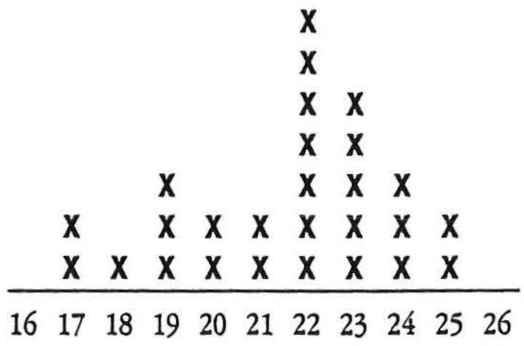
25. If you took out the *outlier* in this distribution, what would happen to the *mean*?

A. it would decrease
B. it would increase
C. it would not change
D. there is no outlier

For questions 26-29, refer to the following situation:

A class investigated the question of how many paces it takes to travel from their class to the gym. They measured the distance by counting the number of paces each student walked. Every step made on the right foot counted as one pace. Here are the results:

Paces to the Gym



26. What is the *median* number of paces the students took to travel the distance?

A. 23
B. 24
C. 21
D. 22

27. What is the *mean* number of paces?

A. 22
B. 21.5
C. 21.56
D. 22.75

28. What is the *mode* in this data?

A. 22
B. 21.5
C. 22.5
D. 23

29. Students who took 17 paces to get to the gym have _________________.

A. the shortest stride
B. the longest stride
C. an average stride
D. the median stride

For questions 30-33, refer to the data that follows:

A group of middle-school students was asked this question: How many movies did you watch last month? Here are a Table and a stem plot of the data:

Student	Number of movies
Joel	15
Tonya	16
Rachel	5
Lawrence	18
Meela	3
Leah	6
Beth	7
Mickey	6
Bhavana	3
Josh	11

Movies Watched

0	3 3 5 6 6 7
1	1 5 6 8
2	

Key
1 | 5 means 15 movies.

30. The total number of students is __________.

A. 11
B. 9
C. 10
D. 18

31. The total number of movies watched is __________.

A. 90
B. 18
C. 100
D. 15

32. The mean number of movies watched is _____________.

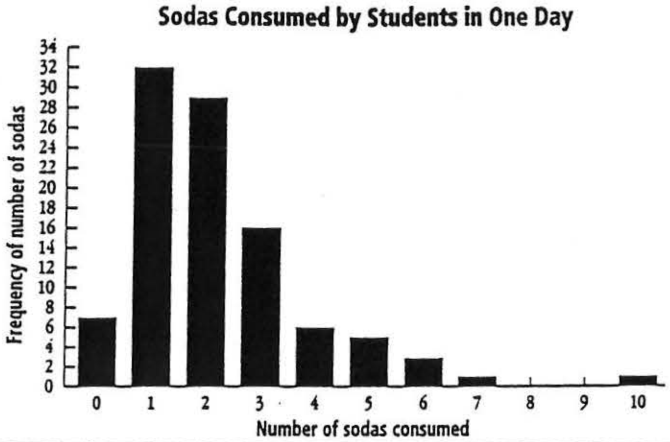
A. 15
B. 6
C. 9
D. 11

33. If a new value was added for a student who watched 42 movies last month, what would happen to the *mean*?

A. it would stay the same
B. it would decrease
C. it would double
D. it would increase

---

34. What kind of data is shown on the Sodas Consumed Chart?



A. categorical data
B. descriptive data
C. numerical data
D. median data

35. Which of the following best describes the *mean* in a set of data?

A. it occurs most often
B. the most typical
C. the "evened out" number
D. it divides the data

36. The *mean* is often referred to as

A. an average
B. a distance
C. a median
D. a statistic

37. If there are six households with a total of 24 people, what is the *mean* number of people in each household?

A. 8
B. 3
C. 6
D. 4

Appendix G

Reliability Analyses of Instrument

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

* * * VAR00001   has zero variance

* * * VAR00003   has zero variance

* * * VAR00004   has zero variance

Correlation Matrix

	VAR00002	VAR00005	VAR00006	VAR00007	VAR00008
VAR00002	1.0000				
VAR00005	-.1627	1.0000			
VAR00006	.1416	.2418	1.0000		
VAR00007	.4524	-.1627	.1416	1.0000	
VAR00008	.6908	-.1124	.0978	-.0658	1.0000
VAR00009	-.0658	.4045	.0978	-.0658	-.0455
VAR00010	-.0658	.4045	.0978	-.0658	-.0455
VAR00011	.0406	.3883	-.5231	.0406	-.1870
VAR00012	.2041	.1195	-.0542	-.1312	.1410
VAR00013	-.0952	.5855	.1416	-.0952	-.0658
VAR00014	.0952	-.5855	-.1416	.0952	.0658
VAR00015	.2114	-.0222	-.0363	-.1627	.4045
VAR00016	.3223	.1284	-.0200	.0134	.2227
VAR00017	-.0658	.4045	.0978	-.0658	-.0455
VAR00018	.4524	-.1627	.1416	.4524	.6908
VAR00019	-.1195	.7348	.1777	-.1195	-.0826
VAR00020	-.0658	-.1124	.0978	-.0658	-.0455
VAR00021	-.1195	.4219	.1777	-.1195	-.0826
VAR00022	.3223	.1284	-.4792	.0134	.2227
VAR00023	-.1195	.1089	-.1629	-.1195	-.0826
VAR00024	.2955	.2936	-.4393	-.0134	.2041
VAR00025	.2041	.1195	-.0542	-.1312	.1410
VAR00026	-.0952	.2114	-.2655	-.0952	-.0658
VAR00027	-.3386	.2041	.1629	.1195	-.5505
VAR00028	-.2041	.7968	.0542	-.2041	-.1410
VAR00029	.0134	.1284	.2097	-.2955	.2227
VAR00030	-.0658	-.1124	.0978	-.0658	-.0455
VAR00031	-.0952	.2114	.1416	-.0952	-.0658
VAR00032	-.2041	.3386	.3035	-.2041	-.1410
VAR00033	-.1627	.4889	.2418	-.1627	-.1124
VAR00034	.0986	.0577	-.1466	.0986	-.1557
VAR00035	.6908	-.1124	.0978	-.0658	1.0000
VAR00036	.5855	.2333	.2418	.2114	.4045

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

Correlation Matrix

	VAR00009	VAR00010	VAR00011	VAR00012	VAR00013
VAR00009	1.0000				
VAR00010	-.0455	1.0000			
VAR00011	.2431	.2431	1.0000		
VAR00012	.1410	.1410	.1989	1.0000	
VAR00013	.6908	.6908	.3519	.2041	1.0000
VAR00014	-.6908	.0658	-.0406	-.2041	-.4524
VAR00015	-.1124	-.1124	-.2496	.1195	-.1627
VAR00016	-.2041	.2227	.2137	.0658	.0134
VAR00017	-.0455	1.0000	.2431	.1410	.6908
VAR00018	-.0658	-.0658	-.2707	-.1312	-.0952
VAR00019	.5505	-.0826	.1812	-.0244	.3386
VAR00020	-.0455	-.0455	-.1870	.1410	-.0658
VAR00021	.5505	.5505	.1812	-.0244	.7968
VAR00022	-.2041	.2227	.5649	.4441	.0134
VAR00023	-.0826	.5505	.1812	-.0244	.3386
VAR00024	.2041	.2041	.8397	.1234	.2955
VAR00025	-.3223	.1410	.1989	-.0268	-.1312
VAR00026	.6908	-.0658	.3519	-.1312	.4524
VAR00027	.0826	.0826	.0793	-.2562	.1195
VAR00028	.3223	.3223	.3729	.2321	.4666
VAR00029	.2227	.2227	-.1374	.0658	.3223
VAR00030	-.0455	-.0455	-.1870	-.3223	-.0658
VAR00031	-.0658	-.0658	.0406	-.4666	-.0952
VAR00032	.3223	-.1410	-.1989	-.3839	.1312
VAR00033	.4045	.4045	.1757	-.3386	.5855
VAR00034	.2919	-.1557	.2802	.2847	.0986
VAR00035	-.0455	-.0455	-.1870	.1410	-.0658
VAR00036	.4045	-.1124	.3883	.1195	.2114

R E L I A B I L I T Y   A N A L Y S I S - S C A L E   (A L P H A)

Correlation Matrix

	VAR00014	VAR00015	VAR00016	VAR00017	VAR00018
VAR00014	1.0000				
VAR00015	-.2114	1.0000			
VAR00016	.2955	.1284	1.0000		
VAR00017	.0658	-.1124	.2227	1.0000	
VAR00018	.0952	.2114	.0134	-.0658	1.0000
VAR00019	-.7968	.1089	-.1124	-.0826	-.1195
VAR00020	.0658	.4045	.2227	-.0455	-.0658
VAR00021	-.3386	-.2041	.1461	.5505	-.1195
VAR00022	-.0134	.1284	.1288	.2227	.0134
VAR00023	.1195	.1089	-.1124	.5505	-.1195
VAR00024	.0134	.0826	.2197	.2041	-.0134
VAR00025	.1312	.1195	.4441	.1410	-.1312
VAR00026	-.4524	-.1627	.0134	-.0658	-.0952
VAR00027	-.1195	-.1089	-.1461	.0826	-.3386
VAR00028	-.4666	.1096	.1234	.3223	-.2041
VAR00029	-.0134	.1284	-.0455	.2227	.0134
VAR00030	.0658	.4045	-.2041	-.0455	-.0658
VAR00031	.0952	.2114	.0134	-.0658	-.0952
VAR00032	-.4666	.1096	-.2550	-.1410	-.2041
VAR00033	-.2114	-.0222	-.0826	.4045	-.1627
VAR00034	-.0986	-.1636	.2145	-.1557	-.2254
VAR00035	.0658	.4045	.2227	-.0455	.6908
VAR00036	-.2114	-.0222	.1284	-.1124	.2114

	VAR00019	VAR00020	VAR00021	VAR00022	VAR00023
VAR00019	1.0000				
VAR00020	-.0826	1.0000			
VAR00021	.2333	-.0826	1.0000		
VAR00022	-.1124	-.2041	-.1124	1.0000	
VAR00023	-.1500	-.0826	.2333	.1461	1.0000
VAR00024	.1124	-.2227	.1124	.5682	.3708
VAR00025	-.0244	-.3223	-.0244	.2550	.2562
VAR00026	.3386	-.0658	.3386	.0134	-.1195
VAR00027	.1500	.0826	.1500	-.1461	.1500
VAR00028	.5855	-.1410	.3050	.1234	.0244
VAR00029	.1461	.2227	.1461	-.3939	.1461
VAR00030	-.0826	-.0455	-.0826	-.2041	.5505
VAR00031	.3386	-.0658	-.1195	-.2955	.3386
VAR00032	.5855	-.1410	.3050	-.2550	.0244
VAR00033	.4219	-.1124	.4219	-.2936	.4219
VAR00034	.2593	-.1557	-.0118	.0318	-.0118
VAR00035	-.0826	-.0455	-.0826	.2227	-.0826
VAR00036	.4219	-.1124	.1089	.1284	-.2041

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	VAR00024	VAR00025	VAR00026	VAR00027	VAR00028
VAR00024	1.0000				
VAR00025	.3125	1.0000			
VAR00026	.2955	-.1312	1.0000		
VAR00027	-.1124	.0244	.1195	1.0000	
VAR00028	.2550	.2321	.1312	.2562	1.0000
VAR00029	.0455	-.3125	.0134	-.4045	.1234
VAR00030	.2041	.1410	-.0658	.0826	-.1410
VAR00031	.2955	.2041	-.0952	.1195	.1312
VAR00032	-.1234	-.1786	.1312	.2562	.1786
VAR00033	.2936	.1195	.2114	.2041	.3386
VAR00034	.1510	.0863	.0986	.0118	.1121
VAR00035	.2041	.1410	-.0658	-.5505	-.1410
VAR00036	.5046	.1195	.2114	-.1089	.1096

	VAR00029	VAR00030	VAR00031	VAR00032	VAR00033
VAR00029	1.0000				
VAR00030	.2227	1.0000			
VAR00031	.3223	.6908	1.0000		
VAR00032	.1234	.3223	.4666	1.0000	
VAR00033	.3395	.4045	.5855	.3386	1.0000
VAR00034	-.1510	-.1557	.0986	.1121	.0577
VAR00035	.2227	-.0455	-.0658	-.1410	-.1124
VAR00036	.1284	-.1124	.2114	.1096	.2333

	VAR00034	VAR00035	VAR00036
VAR00034	1.0000		
VAR00035	-.1557	1.0000	
VAR00036	.2791	.4045	1.0000

N of Cases = 23.0

Statistics for	Mean	Variance	Std Dev	N of Variables
Scale	23.5652	15.6206	3.9523	33

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Reliability Coefficients   33 items

Alpha =   .7246   Standardized item alpha =   .7239

Reliability

****** Method 2 (covariance matrix) will be used for this analysis ****

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (S P L I T)

		Mean	Std Dev	Cases
1.	VAR00005	.7826	.4217	23.0
2.	VAR00007	.9130	.2881	23.0
3.	VAR00009	.9565	.2085	23.0
4.	VAR00011	.5652	.5069	23.0
5.	VAR00013	.9130	.2881	23.0
6.	VAR00015	.7826	.4217	23.0
7.	VAR00017	.9565	.2085	23.0
8.	VAR00019	.8696	.3444	23.0
9.	VAR00021	.8696	.3444	23.0
10.	VAR00023	.8696	.3444	23.0
11.	VAR00025	.3043	.4705	23.0
12.	VAR00027	.1304	.3444	23.0
13.	VAR00029	.5217	.5108	23.0
14.	VAR00031	.9130	.2881	23.0
15.	VAR00033	.7826	.4217	23.0
16.	VAR00035	.9565	.2085	23.0
17.	VAR00002	.9130	.2881	23.0
18.	VAR00006	.1739	.3876	23.0
19.	VAR00008	.9565	.2085	23.0
20.	VAR00010	.9565	.2085	23.0
21.	VAR00012	.3043	.4705	23.0
22.	VAR00014	.0870	.2881	23.0
23.	VAR00016	.5217	.5108	23.0
24.	VAR00018	.9130	.2881	23.0
25.	VAR00020	.9565	.2085	23.0
26.	VAR00022	.5217	.5108	23.0
27.	VAR00024	.4783	.5108	23.0
28.	VAR00026	.9130	.2881	23.0
29.	VAR00028	.6957	.4705	23.0
30.	VAR00030	.9565	.2085	23.0
31.	VAR00032	.6957	.4705	23.0
32.	VAR00034	.6522	.4870	23.0
33.	VAR00036	.7826	.4217	23.0
34.	VAR00001	1.0000	.0000	23.0
35.	VAR00003	1.0000	.0000	23.0
36.	VAR00004	1.0000	.0000	23.0

* * * VAR00001 has zero variance

* * * VAR00003 has zero variance

* * * VAR00004 has zero variance

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (S P L I T)

Correlation Matrix

	VAR00005	VAR00007	VAR00009	VAR00011	VAR00013
VAR00005	1.0000				
VAR00007	-.1627	1.0000			
VAR00009	.4045	-.0658	1.0000		
VAR00011	.3883	.0406	.2431	1.0000	
VAR00013	.5855	-.0952	.6908	.3519	1.0000
VAR00015	-.0222	-.1627	-.1124	-.2496	-.1627
VAR00017	.4045	-.0658	-.0455	.2431	.6908
VAR00019	.7348	-.1195	.5505	.1812	.3386
VAR00021	.4219	-.1195	.5505	.1812	.7968
VAR00023	.1089	-.1195	-.0826	.1812	.3386
VAR00025	.1195	-.1312	-.3223	.1989	-.1312
VAR00027	.2041	.1195	.0826	.0793	.1195
VAR00029	.1284	-.2955	.2227	-.1374	.3223
VAR00031	.2114	-.0952	-.0658	.0406	-.0952
VAR00033	.4889	-.1627	.4045	.1757	.5855
VAR00035	-.1124	-.0658	-.0455	-.1870	-.0658
VAR00002	-.1627	.4524	-.0658	.0406	-.0952
VAR00006	.2418	.1416	.0978	-.5231	.1416
VAR00008	-.1124	-.0658	-.0455	-.1870	-.0658
VAR00010	.4045	-.0658	-.0455	.2431	.6908
VAR00012	.1195	-.1312	.1410	.1989	.2041
VAR00014	-.5855	.0952	-.6908	-.0406	-.4524
VAR00016	.1284	.0134	-.2041	.2137	.0134
VAR00018	-.1627	.4524	-.0658	-.2707	-.0952
VAR00020	-.1124	-.0658	-.0455	-.1870	-.0658
VAR00022	.1284	.0134	-.2041	.5649	.0134
VAR00024	.2936	-.0134	.2041	.8397	.2955
VAR00026	.2114	-.0952	.6908	.3519	.4524
VAR00028	.7968	-.2041	.3223	.3729	.4666
VAR00030	-.1124	-.0658	-.0455	-.1870	-.0658
VAR00032	.3386	-.2041	.3223	-.1989	.1312
VAR00034	.0577	.0986	.2919	.2802	.0986
VAR00036	.2333	.2114	.4045	.3883	.2114

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (S P L I T)

Correlation Matrix

	VAR00015	VAR00017	VAR00019	VAR00021	VAR00023
VAR00015	1.0000				
VAR00017	-.1124	1.0000			
VAR00019	.1089	-.0826	1.0000		
VAR00021	-.2041	.5505	.2333	1.0000	
VAR00023	.1089	.5505	-.1500	.2333	1.0000
VAR00025	.1195	.1410	-.0244	-.0244	.2562
VAR00027	-.1089	.0826	.1500	.1500	.1500
VAR00029	.1284	.2227	.1461	.1461	.1461
VAR00031	.2114	-.0658	.3386	-.1195	.3386
VAR00033	-.0222	.4045	.4219	.4219	.4219
VAR00035	.4045	-.0455	-.0826	-.0826	-.0826
VAR00002	.2114	-.0658	-.1195	-.1195	-.1195
VAR00006	-.0363	.0978	.1777	.1777	-.1629
VAR00008	.4045	-.0455	-.0826	-.0826	-.0826
VAR00010	-.1124	1.0000	-.0826	.5505	.5505
VAR00012	.1195	.1410	-.0244	-.0244	-.0244
VAR00014	-.2114	.0658	-.7968	-.3386	.1195
VAR00016	.1284	.2227	-.1124	.1461	-.1124
VAR00018	.2114	-.0658	-.1195	-.1195	-.1195
VAR00020	.4045	-.0455	-.0826	-.0826	-.0826
VAR00022	.1284	.2227	-.1124	-.1124	.1461
VAR00024	.0826	.2041	.1124	.1124	.3708
VAR00026	-.1627	-.0658	.3386	.3386	-.1195
VAR00028	.1096	.3223	.5855	.3050	.0244
VAR00030	.4045	-.0455	-.0826	-.0826	.5505
VAR00032	.1096	-.1410	.5855	.3050	.0244
VAR00034	-.1636	-.1557	.2593	-.0118	-.0118
VAR00036	-.0222	-.1124	.4219	.1089	-.2041

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (S P L I T)

Correlation Matrix

	VAR00025	VAR00027	VAR00029	VAR00031	VAR00033
VAR00025	1.0000				
VAR00027	.0244	1.0000			
VAR00029	-.3125	-.4045	1.0000		
VAR00031	.2041	.1195	.3223	1.0000	
VAR00033	.1195	.2041	.3395	.5855	1.0000
VAR00035	.1410	-.5505	.2227	-.0658	-.1124
VAR00002	.2041	-.3386	.0134	-.0952	-.1627
VAR00006	-.0542	.1629	.2097	.1416	.2418
VAR00008	.1410	-.5505	.2227	-.0658	-.1124
VAR00010	.1410	.0826	.2227	-.0658	.4045
VAR00012	-.0268	-.2562	.0658	-.4666	-.3386
VAR00014	.1312	-.1195	-.0134	.0952	-.2114
VAR00016	.4441	-.1461	-.0455	.0134	-.0826
VAR00018	-.1312	-.3386	.0134	-.0952	-.1627
VAR00020	-.3223	.0826	.2227	-.0658	-.1124
VAR00022	.2550	-.1461	-.3939	-.2955	-.2936
VAR00024	.3125	-.1124	.0455	.2955	.2936
VAR00026	-.1312	.1195	.0134	-.0952	.2114
VAR00028	.2321	.2562	.1234	.1312	.3386
VAR00030	.1410	.0826	.2227	.6908	.4045
VAR00032	-.1786	.2562	.1234	.4666	.3386
VAR00034	.0863	.0118	-.1510	.0986	.0577
VAR00036	.1195	-.1089	.1284	.2114	.2333

	VAR00035	VAR00002	VAR00006	VAR00008	VAR00010
VAR00035	1.0000				
VAR00002	.6908	1.0000			
VAR00006	.0978	.1416	1.0000		
VAR00008	1.0000	.6908	.0978	1.0000	
VAR00010	-.0455	-.0658	.0978	-.0455	1.0000
VAR00012	.1410	.2041	-.0542	.1410	.1410
VAR00014	.0658	.0952	-.1416	.0658	.0658
VAR00016	.2227	.3223	-.0200	.2227	.2227
VAR00018	.6908	.4524	.1416	.6908	-.0658
VAR00020	-.0455	-.0658	.0978	-.0455	-.0455
VAR00022	.2227	.3223	-.4792	.2227	.2227
VAR00024	.2041	.2955	-.4393	.2041	.2041
VAR00026	-.0658	-.0952	-.2655	-.0658	-.0658
VAR00028	-.1410	-.2041	.0542	-.1410	.3223
VAR00030	-.0455	-.0658	.0978	-.0455	-.0455
VAR00032	-.1410	-.2041	.3035	-.1410	-.1410
VAR00034	-.1557	.0986	-.1466	-.1557	-.1557
VAR00036	.4045	.5855	.2418	.4045	-.1124

R E L I A B I L I T Y   A N A L Y S I S - S C A L E   (S P L I T)

Correlation Matrix

	VAR00012	VAR00014	VAR00016	VAR00018	VAR00020
VAR00012	1.0000				
VAR00014	-.2041	1.0000			
VAR00016	.0658	.2955	1.0000		
VAR00018	-.1312	.0952	.0134	1.0000	
VAR00020	.1410	.0658	.2227	-.0658	1.0000
VAR00022	.4441	-.0134	.1288	.0134	-.2041
VAR00024	.1234	.0134	.2197	-.0134	-.2227
VAR00026	-.1312	-.4524	.0134	-.0952	-.0658
VAR00028	.2321	-.4666	.1234	-.2041	-.1410
VAR00030	-.3223	.0658	-.2041	-.0658	-.0455
VAR00032	-.3839	-.4666	-.2550	-.2041	-.1410
VAR00034	.2847	-.0986	.2145	-.2254	-.1557
VAR00036	.1195	-.2114	.1284	.2114	-.1124

	VAR00022	VAR00024	VAR00026	VAR00028	VAR00030
VAR00022	1.0000				
VAR00024	.5682	1.0000			
VAR00026	.0134	.2955	1.0000		
VAR00028	.1234	.2550	.1312	1.0000	
VAR00030	-.2041	.2041	-.0658	-.1410	1.0000
VAR00032	-.2550	-.1234	.1312	.1786	.3223
VAR00034	.0318	.1510	.0986	.1121	-.1557
VAR00036	.1284	.5046	.2114	.1096	-.1124

	VAR00032	VAR00034	VAR00036
VAR00032	1.0000		
VAR00034	.1121	1.0000	
VAR00036	.1096	.2791	1.0000

N of Cases = 23.0

Statistics for	Mean	Variance	Std Dev	N of Variables
Part 1	12.0870	5.7194	2.3915	16
Part 2	11.4783	4.3518	2.0861	17
Scale	23.5652	15.6206	3.9523	33

R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (S P L I T)

Reliability Coefficients   33 items

Correlation between forms =   .5562   Equal-length Spearman-Brown =
.7148

Guttman Split-half =   .7105   Unequal-length Spearman-Brown =
.7149

Alpha for part 1 =   .6701   Alpha for part 2 =
.4478

16 items in part 1   17 items in part 2

Appendix H

Performance Assessment

Name ______________________________ Date ____________

Unit Test

1. A group of 9 students has these numbers of children in their families:

   3, 2, 4, 2, 1, 5, 1, 2, and 7.

   a. Find the median number of children in the 9 families.

   b. Find the mean number of children in the 9 families.

2. The stem plot below shows the number of minutes it took a class of students to travel to school.

Travel Times to School (minutes)

0	3 3 5 7 8 9
1	0 2 3 5 6 6 8 9
2	0 1 3 3 3 5 5 8 8
3	0 5
4	5

Key
2 | 5 means 25 minutes.

   a. Are these data numerical or categorical?

   b. What is the range of the data?

   c. What is the median of the data? How many students had a score the same as the median?

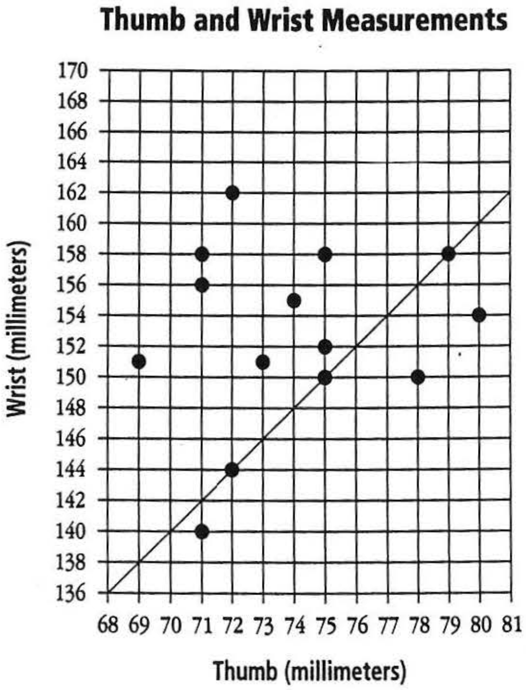
3. Fifteen students read the book *Gulliver's Travels.* In the book, the Lilliputians said they could make clothes for Gulliver by taking one measurement, the length around his thumb. The Lilliputians claimed that

   • the distance around Gulliver's wrist would be twice the distance around his thumb.

   • the distance around Gulliver's neck would be twice the distance around his wrist.

   • the distance around Gulliver's waist would be twice the distance around his neck.

The students wondered whether this doubling relationship would be true for them, too. They measured the distance around their thumbs and their wrists in centimeters, and then graphed the pairs of numbers on a coordinate graph. They drew a line connecting the points that represented wrist measurements that were twice thumb measurements.



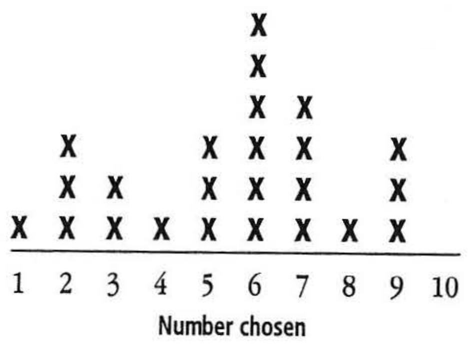
a. How many students' measurements fit the Lilliputian rule that twice the distance around the thumb equals the distance around the wrist?

b. How many students' wrist measurements are less than their thumb measurements?

c. The point for Jeri's thumb and wrist measurements is above the line. If the cuffs of a shirt are twice the measurement around Jeri's thumb, how will the cuffs of the shirt fit her?

d. The point for Robin's thumb and wrist measurements is below the line. If the cuffs of a shirt are twice the measurement around Robin's thumb, how will the cuffs of the shirt fit him?

4. Ms. Snow had her students write down a whole number between 1 and 10 on a slip of paper. Then she collected the papers and displayed the data in a line plot. Use the line plot to answer the following questions.



a. What is the median number chosen by students in this class? What is the mean number chosen? What is the mode?

b. Make a bar graph that displays this information. Explain how the graph is similar to and different from a line plot.

c. If two students were absent on the day Ms. Snow collected the data, how many students are enrolled in the class? Explain your reasoning.