12-2003

An Investigation on the Impact of the Structure of Summative Student Evaluation on Self-Regulated Learning

John Charles Ritzier
Western Michigan University

AN INVESTIGATION ON THE IMPACT OF THE STRUCTURE OF SUMMATIVE STUDENT EVALUATION ON SELF-REGULATED LEARNING

by

John Charles Ritzler

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Educational Studies

Western Michigan University
Kalamazoo, Michigan
December 2003

UMI Number: 3123766

ACKNOWLEDGMENTS

While this project has in one sense been a lonely, individualistic effort, it has also taught me in a new and deeper way the degree to which I am dependent on others. I have been reminded that the degree to which I maximize my individual potential is determined by the degree to which I allow others to meaningfully contribute to my work.

I do not believe it is possible to even quantify the number of people who have contributed to this work much less list each of their names. I sincerely believe there are people who have contributed without my recognition, there are others who have done so without knowing so themselves, and there are still others who have contributed, yet neither I nor they were aware of it. Nevertheless, there are two broad categories of people whose contributions I am quite cognizant of.

The first is my family. From the time of my birth through the present, each family member has made unique and significant contributions to my ability to complete this work and indeed to the work itself.

The second set of people whose contributions were evident includes my professional mentors and colleagues. This includes mentors at Western Michigan University as well as others at other universities, and in other professional settings. Again, each has made unique and significant contributions for which I am grateful.

To all who have blessed me in so many ways throughout my life and have thus contributed to this work, thank you and may God bless you.

John Charles Ritzler

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Leadership and Motivation in the Classroom

As a high school mathematics teacher, I had often reflected on ways to maximize the quality of student learning. During that period of my career, I had attended numerous professional development activities, had subscribed to professional journals, and had maintained dialogs with a variety of professional colleagues. There seemed to be enough suggested solutions to maximizing student learning that I could try a different technique or theory every day of the year. Given this myriad of suggestions, it was important for me to develop a broader view of student learning. I was confident that a broad conceptualization of student learning would lead to a more all-encompassing approach to improving student learning. This type of approach seemed to hold more promise than the many "silver-bullets" that were being pitched by any consultant who lived more than 50 miles away from our school.

During this period of my career, I was also beginning graduate studies in the area of educational leadership. Many of my peers in these classes were either in administrative positions or applying for administrative positions. Most, if not all, of the literature on educational leadership was written from the perspective of school building, central office, or governmental level educational leaders. After digesting

these theories of leadership, I noticed that each involved establishing productive relationships between leaders and followers. Because of this, it was my belief that these same theories of leadership and motivation could be applied at the classroom level with the teacher viewed as an educational leader and the student viewed as the follower. The core concept that productivity is maximized when followers see purpose and take ownership of their work seemed naturally to apply also to maximizing student learning.

The lack of a perceived purpose for students in high school mathematics classes is evident. Over the years, I had seen far too many students who saw no purpose in high school mathematics. I had heard questions like "Why do I have to learn this?" or "When will I ever need to know this?" far too often. Even students who were fairly successful in the classroom, often appeared to be simply going through the motions. They did what needed to be done to learn enough material to get the grade they wanted; however, few students seemed to have a genuine interest in "in-depth" learning.

Similarly, few high school mathematics students take ownership over their learning. Indeed, I and most mathematics teachers define the curriculum as well as the means that will be used to teach that curriculum. Students are seldom involved in decisions regarding the instructional methods used in the classroom and they are even less likely to be involved in decisions regarding the content that is covered. While these decisions perhaps should be made by educational professionals, the fact that they

are certainly limits the ownership students assume for their learning.

This combination of lack of ownership and lack of perceived purpose in learning mathematics clearly signified to me that something drastically different needed to be done not only for the sake of those who were performing poorly in my classes, but also to improve the level of learning for those whose grades suggested they were doing fine. While many curricular changes were available, these seemed unlikely to have the broad impact I was seeking. On the other hand, because a sense of ownership and purpose were missing, attempting to apply the concepts of educational leadership and motivation in my classroom seemed a very promising way to increase the motivation of students and to shift their focus from simply mastering skills to a focus on truly understanding mathematics. In short, applying the concepts of educational leadership and motivation in my classroom seemed a promising way of positively impacting the quality of student learning for all students.

One of the leading theorists in the area of leadership and motivation is Frederick Herzberg. In 1987, the Harvard Business Review republished a 1968 article written by Herzberg entitled "One More Time: How Do You Motivate Employees?" At the time, the article had sold more than a million reprints and was the number-one requested article published in the journal. In the article, Herzberg stated that employees are motivated by factors such as challenging work, responsibility, and opportunities for growth. Surprisingly, they are not motivated by external rewards such as salary, fringe benefits, and incentive programs. These external, non-motivating

factors were labeled by Herzberg as *hygiene factors*. Examples of hygiene factors in a classroom may include teacher assigned marks, gold stars on papers, exemption from an exam based on performance on homework, or any number of other teacher created and teacher imposed rewards. According to Herzberg, hygiene factors like these are necessary to avoid dissatisfaction; however, they do not lead to satisfaction and are not motivating factors. In his article, Herzberg states that the research he used to draw these conclusions has been replicated by "at least 16 other investigations, using a wide variety of populations" (p. 112).

A second authority on the topic of motivation is Abraham Maslow, who has identified five levels of needs that motivate people (1970). Maslow's needs are arranged in a hierarchy. Working from the lowest level of motivational factors to the highest, he identifies physiological, safety, social, esteem, and self-actualization (these levels are shown in figure 1). Maslow's higher levels of needs - social (which is the need to belong or be accepted), esteem (the need to be recognized and respected) and self-actualization (the need to maximize one's individual potential or to become all one is capable of becoming) - are very similar to the factors Herzberg identifies as motivational factors. Unfortunately, I seldom saw evidences of any of these higher order needs being met through student learning in my classroom. To the contrary, the need for learning seemed to be related to survival (physiological and safety needs). In some cases the survival need was related to family pressures, in other cases the student himself felt that his long term survival chances would benefit via the learning of

mathematics. However, I did not frequently see students whose social needs, esteem needs, or self-actualization needs were met through the learning of mathematics.

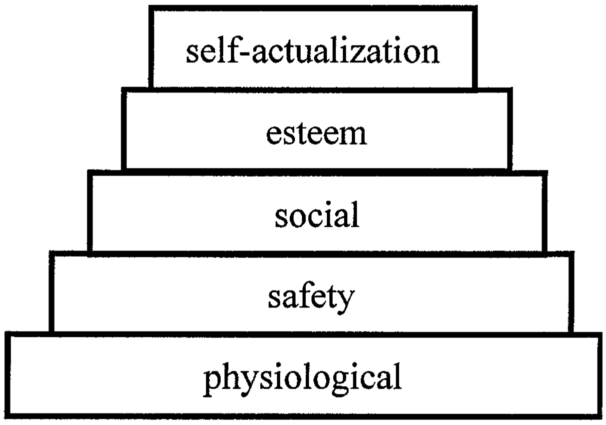

Figure 1. Maslow's Hierarchy of Needs

Motivation theories like those of Herzberg and Maslow fit very naturally into leadership theories. A number of theorists have proposed specific leadership models using these concepts to maximize the motivation and hence the productivity of followers. Because I am conceptualizing classroom teachers as being in a position of leadership over the students (followers) in their classrooms, those leadership theories ought to provide guidance to classroom teachers.

One of the more prominent theorists in the area of leadership was McGregor who identified two forms of leadership, theory x and theory y (1960). Theory y is a participatory style of leadership; in theory x, the leader has control and directs the followers. What is important about this is that, according to McGregor, theory x employees avoid and dislike work, seek direction not responsibility, seek security not

ambition, and must be supervised closely. Theory x describes the structure of many traditional classrooms quite well, because the teachers have control and direct the students. Because of this, students tend to seek direction, not responsibility and students tend to avoid and dislike work. If a teacher's classroom leadership was aligned closer to theory y, it would be more likely that the student's would assume ownership of their learning and approach tasks from a "mastery of the material" rather than a "complete the task" perspective. These qualities that are birthed in followers via a theory y style of leadership have been categorized in educational literature as self-regulated learning qualities. Research has demonstrated that when students practice the art of self-regulated learning, achievement increases (Graham & Golan, 1991; Linder & Harris, 1992; Pintrich & De Groot, 1990; Stipek & Kowalski, 1989; VanderStoep, Pintrich & Fagerlin, 1996; Williams, 1996; Zimmerman & Martinez-Pons, 1988; Zimmerman & Martinez-Pons, 1986). Thus, if classroom leadership models moved to a more participatory or theory y framework, this should draw out the self-regulated learning qualities in the followers and thus improve the quality of their learning.

Argyris (1964) also argued that organizations with autocratic values are characterized by decreased interpersonal competence and hence diminished organizational effectiveness. Organizations that employ democratic values, on the other hand, are marked by increased interpersonal competence and hence increased organizational effectiveness. Again, few traditional classrooms would be characterized

as “employing democratic values.” It follows then, that the organizational effectiveness of traditional classrooms (i.e., the quality of student learning) is diminished. A move to a more participatory or democratic style of classroom leadership ought to increase the quality of student learning.

McGregor and Argyris are far from the only two leadership theorists to recommend a participatory style of leadership. Likert (1967) identified four systems of management, ranging from an autocratic to a participatory style. Similarly, Bennis and Nanus (1985) in their study of leaders in successful organizations found that effective leaders do not control but empower employees. Senge (1990) stated that a key to organizational effectiveness was involving the employees in creating a shared vision and, finally, Block (1993) stated that democratic organizations are the only ones that ultimately will succeed. In other words, classrooms that function under a traditional style of teacher led, autocratic, theory x leadership are destined to be less effective and indeed unsuccessful organizations.

Clearly many theorists already have proposed answers that can provide general guidance to classroom teachers who are wrestling with the question of how they can manage, direct, or lead their students in order to maximize the quality of student learning. If the classroom leadership of teachers were influenced by these theories, and in turn, students were empowered to take ownership over their own learning, student motivation would improve, and, it is probable the quality of student learning would also improve.

Evaluation in the Classroom

While theories of leadership and motivation can provide general guidance to classroom teachers, more specific guidance can be gleaned through an investigation of theories that relate leadership and motivation to the evaluation of employees (or students) within an organization. Like classroom leadership and motivation, the concept of student evaluations can be informed by theorists from the business sector who addressed the issues of employee evaluation as it relates to organizational effectiveness. One of the first theorists to specifically address the issues of employee evaluation as an element of leadership and a tool for motivation was W. Edwards Deming. Mr. Deming, a native of the United States, traveled to Japan after World War II to assist that country in redeveloping its industry. The methods for assuring quality that Deming taught, and that Japan implemented, have been so successful that they have pushed others in the industrial world to follow in an attempt to duplicate his high level of success. There are now a number of organizations around the world that have employed Deming's blend of leadership, motivation, and employee evaluation to improve quality. In his classic book Out of the Crisis (1986), he listed 14 points to which businesses should adhere if they want to develop motivated employees and hence successful organizations.

Not only have private businesses attempted to duplicate Deming's success, but at least one public school has adopted his theories as a means to improving the quality of student learning. David Langford (1997) and his students at Mt. Edgecumbe High

School in Sitka, Alaska, translated Deming's 14 points into 15 analogous points for schools. What is most interesting about this translation from the language of business into the language of education is, when the translation is examined, five of the fifteen points include direct suggestions for how students should be evaluated.

Specifically, point three states that schools should "work to abolish grading and the harmful effects of rating people." This thought is expanded in point four where Langford and his students write "cease dependence on testing to achieve quality. Eliminate the need for inspection on a mass basis (standardized achievement tests, minimum graduation exams, etc.) by providing learning experiences and systems which build in quality." Points eleven and twelve respectively encourage schools to "eliminate slogans, exhortations and targets for teachers and students" and to "eliminate work standards (quotas) on teachers and students, (e.g., raise test scores by 10%, and lower dropouts by 15%)." Finally, point thirteen says that "the responsibility of all educational managers must be changed . . . from results to process." This statement that the responsibility of educational managers (i.e., teachers) must be changed implies that the current focus is results oriented. Indeed, the practice of student evaluation in schools is currently very "results oriented." According to Langford, this core concept of student evaluations must change.

These recommendations for evaluation of students are a direct result of Deming's philosophies on leadership. The empowerment of core workers which leads to a participatory style of leadership ought to have an effect on employee (or student)

evaluations. Student evaluations ought to be participatory, student-centered and non-punitive in a school that is attempting to implement these leadership ideals. As has already been stated, however, what Deming is suggesting for organizations, and Langford is implementing in schools, is a dramatically different approach to student evaluation when compared to the approach that has been traditionally taken.

Langford is not the only educator to adopt Deming's theories as a means to improving the quality of student learning. Deming's ideas are also echoed in The Quality School (Glasser, 1992), in which Glasser argues that if schools are to be truly "quality schools," they may in no way coerce students. Important to this study is the fact that he points out that one of the primary forms of coercion is the practice of giving marks to students. This grading of students shifts a student's focus from learning the material to earning a mark. The classroom teacher has a hammer to hold over the student's head as a means of forcing him or her to produce. Clearly, marks are a classic form of coercion. Glasser proposes to lessen the coercion by changing to a mastery learning system of marking (A, B, or no grade).

Like Glasser, many other educational theorists have specifically advocated that schools move away from the marking system as a motivational device for students (Arnstine & McDowell, 1993; Berliner & Casanova, 1988; Juarez, 1996; Malehorn, 1984). While their alternatives to marks vary, their reasoning is similar to that of Glasser. Juarez and Malehorn each expand upon Glasser's coercion argument by pointing out the anxiety that can be associated with the coercive practice of assigning

grades. Juarez (1996) writes, "Grades not only have the power to 'break students' hearts.' They also can - and often do - 'break' students' lives" (p. 376). In making his argument for alternatives to the marking system, Malehorn (1984) states, "Marks are misleading and incomplete at best. At worst they are inhibiting and traumatizing. If the goal of educational establishments is truly to foster all kinds of learning, school personnel need to examine their assessment methods. Perhaps more than any other element of schooling, grades interfere with pupils' efforts to learn" (p. 257). Arnstine and McDowell expand these points even further. In addition to recognizing the coercive factor and the anxiety associated with marking students, they also point out the fact that a lack of clarity on how grades are determined may actually cause students to strive for undesired goals. Arnstine and McDowell write, "students can be driven to whatever behavior they think will give them a reward" (p. 5). They conclude with the statement, "The preoccupation with such reward systems becomes deleterious to education because all parties to the process - teachers, students, administrators, parents, and the public at large - become absorbed in and anxious about . . . grades and grade-point-averages" (p. 17).

In summary, it is abundantly clear that Deming's principles for leadership, motivation, and evaluation in business organizations can be profitably applied within the educational classroom in general and in the context of student evaluation in specific.

Investigating the Impacts of Evaluation in the Classroom

A beginning step in any research project is to clearly identify the dependent variable. With respect to research on the impact that the structure of summative student evaluation may have in a classroom, the question must be asked, "what are we trying to impact?" In other words, what is the goal of schooling in general, of each individual classroom in particular, and of the process of giving marks within a classroom?

Not surprisingly, in a study investigating the purposes of marks in schools, it was found that there was disagreement among students, parents, teachers, and administrators as to what those purposes were (Gorman, 1989). One could infer that if the purposes of marks were viewed differently, then the purposes of the individual classrooms and of schooling in general are likely viewed differently also. Indeed, numerous purposes of schooling have been suggested by equally numerous theorists. Most all lists include intellectual areas such as "mastery of basic skills" or "learning to learn" as well as emotional, physical, societal, cultural, aesthetic, and, in some cases, moral development areas (Goodlad, 1984; Rothney, 1955; Tyler, 1942). Clearly some decision must be made regarding the purposes of schooling and hence student evaluations within those schools prior to addressing questions regarding the impact of student evaluations. Indeed, one of the difficulties in educational research is narrowing these purposes down to a workable dependent variable.

For the purposes of this study, I have chosen to focus only on the intellectual

purposes of schooling. This is because all theorists agree on intellectual development as a primary purpose of school. More specifically, the focus of this study will be on the success of the classroom teacher in developing students who have "learned how to learn." This construct of "knowing how to learn" has itself been researched significantly (Ablard & Lipschultz, 1998; Corno & Mandinach, 1983; Lindner & Harris, 1992; Pintrich & DeGroot, 1990; VanderStoep, Pintrich & Fagerlin, 1996; Williams, 1996; Zimmerman, 1986). Students who have developed these skills are referred to as "self-regulated learners." Again, I believe few would argue that if student evaluations enhanced the development of self-regulated learning, then the student evaluations, and in turn, the schools themselves had been successful.

While I am defining the primary purpose of schooling to be the development of self-regulated learners, a secondary benefit of this definition is the impact self-regulated learning has on achievement or mastery of basic skills. A number of research studies (which will be addressed in the second chapter of this proposal) have identified a correlation between self-regulated learning and student achievement. In other words, indicators of students' engagement in learning can be viewed as indicators of actual learning. Because of this, it will not be necessary to specifically investigate achievement outcomes. It is important to note that this correlation to achievement is only a secondary benefit of choosing the development of self-regulated learners as the dependent variable. The primary reason is the researcher's bias that the development of self-regulated learners truly is the primary purpose of schooling.

Statement of the Problem

In this study, I will explore the possible educational utility of the leadership and motivational theories discussed above by investigating the following research questions:

1. How does the method of summative student evaluation impact self-regulated learning in low-achieving high school mathematics students?

2. How does the cognitive level of the task mediate the relationship between method of summative student evaluation and self-regulated learning?

3. How does the length of time students are exposed to a method of summative student evaluation impact self-regulated learning?

The methods of summative student evaluation that are considered here include: (a) traditional grading (A, B, C, D, or F), (b) credit/no-credit, and (c) student-developed summative evaluation.

The dependent variable, self-regulated learning, will be conceptualized to consist of four factors. They are: (1) self-efficacy, (2) motivation, (3) mastery orientation to learning, and (4) use of cognitive strategies. These factors were drawn from previous research studies regarding the larger construct of self-regulated learning. Some of that research is reviewed in Chapter Two of this study.

Significance of the Study

For educational practitioners to investigate the impact particular programs or

procedures have on students is often difficult. Even if the goal of the program is seemingly clear, such as increased student achievement, the reluctance to create an experimental setting, the large number of variables affecting student achievement, and the potentially complex interactions between those variables, all contribute to the challenge educators face in drawing valid conclusions. A school may have direct evidence, like increased test scores, that student achievement is improving; however, they may not be able to point to exactly what impact, if any, a given piece of their program is contributing to that growth.

For example, this study was set in the context of a remediation summer school program in an urban school district. Students were mandated to participate in the program based on the results of standardized testing and teacher recommendations. Once the students were identified, all students were mandated to participate. No attempt was made to randomly assign some students to a nonparticipating control group. The following year, the district saw an increase in their standardized test scores. Some concluded the remediation summer school program must have been successful. However, this was also the first year after a new superintendent assumed the leadership of the district. During that year, the superintendent had made numerous efforts to raise the level of attention given to standardized tests. In addition, this was the first year after implementation of a spring, during the school day, remediation program. There were also a number of building level programs that were either implemented or eliminated during that year. Clearly, it is not possible to conclude the

increase in test scores was a result of only the summer remediation program.

One significance of this study is the fact that it is conducted within the context of actual classrooms, yet, by controlling many confounding variables, valid conclusions regarding self-regulated learning can be drawn.

A second significance of this study is its direct contribution to the empirical research regarding structure of summative student evaluation. Indeed, while there is a significant amount of theoretical literature regarding summative student evaluations, there is very little empirical research relating summative student evaluation to any traditional goals of education. This study will assist administrators and teachers to make informed decisions regarding the structure of summative student evaluations.

Limitations of the Study

A number of limitations characterize this study. One is its restriction of context to low-achieving high school mathematics students. The impact of summative evaluation methods in other disciplines, at other levels (i.e., at the primary or post-secondary level), or with other types of students (i.e., average or high-achieving students) may well vary. For example, primary level students have had relatively limited experiences with summative evaluations. Post-secondary students, on the other hand, have been exposed to summative evaluations on numerous occasions and, likely, in a variety of forms. Clearly, the background, and hence baggage, that high school students possess relative to summative evaluations lies somewhere between

these two. It is reasonable to assume that the impact of varying forms of summative evaluation would be mediated by previous evaluation experiences and hence vary for students who are at differing stages of their educational career. Similarly, low-achieving students have likely had many negative past experiences with summative evaluations. Again, the impact current varieties of summative evaluation may have will almost certainly not be independent of those previous experiences. Hence, the impact on average or high-achieving students may well be different.

A second limitation of the study is its focus on the four characteristics used to operationalize self-regulated learning. Other authors may well look for slightly different manifestations of self-regulated learning.

Finally, as with any qualitative study, this study may indeed suggest relationships between the constructs of summative student evaluation and self-regulated learning; however, it would be limited in its generalizability. Qualitative studies in general, and the grounded theory approach that has been modified for use in this study specifically, are designed to unveil potential theories. They are not designed to prescribe future outcomes.

Outline of the Dissertation

The dissertation will be organized as follows:

Chapter II. Literature Review

Chapter III. Method

Chapter IV. Results

Chapter V. Discussion

CHAPTER 2

REVIEW OF LITERATURE

Formative Student Evaluations

The concepts of assessment and evaluation are difficult to define. Further delineation to ideas such as formative and summative evaluation methods are equally difficult to make. In writing about these concepts in the context of educational classrooms, Black and Wiliam (1998) write:

> We use the general term assessment to refer to all those activities undertaken by teachers - and by their students in assessing themselves - that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs. (p. 140)

Black and Wiliam footnote this definition with the comment that "there is no internationally agreed-upon term here" (p. 148). Conner et al. (1985) write "formative measures are ones used during instruction to improve learning and teaching rather than after instruction to grade and rank" (p. 63).

While this study is focused on summative student evaluations, there are at least two reasons for completing a brief review of the literature related to formative student evaluations. First, as is implied by Black and Wiliam, the line between formative and

summative is not clearly defined. The second reason, which is dependent on the first, is that the use of what may be termed formative techniques within the context of summative evaluations has been shown to have a positive effect on student achievement. For example, Butler and Nisan (1986) found that the inclusion of task-related comments as a piece of the summative student evaluation led to an increase in student performance as opposed to giving marks alone. Indeed, this type of finding is one motivation for the design of this study which investigates alternatives to traditional, purely summative, classroom evaluation practices.

Formative evaluation techniques and their impact on students have been researched rather extensively. In 1998, Dylan and Wiliams conducted a meta-analysis summarizing 250 sources. They concluded "that innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains" (p. 140). Of even greater pertinence to this study was their conclusion regarding the impact on low-achieving students. They write, "improved formative assessment helps low achievers more than other students and so reduces the range of achievement while raising achievement overall" (p. 141). In addition, to the study by Dylan and Wiliams, Fuchs and Fuchs (1986) conducted a meta-evaluation of 21 studies investigating the effects of formative evaluation on special education students. They conclude "that the use of systematic formative evaluation procedures, within a group of studies that employed predominantly mildly handicapped subjects, significantly increased students' school achievement, both statistically and practically"

(p. 205). These conclusions suggest that any inclusion of formative techniques in summative classroom evaluation practices may well be particularly beneficial to the low-achieving students who are the focus of this study.

Summative Student Evaluations

Focusing now on summative student evaluations, clearly the primary summative student evaluation structure in traditional K-12 schools has, for at least the last 100 years, been to assign marks (or grades). According to Malehorn (1984), "grades have been a part of schooling for generations" (p. 256). Juarez (1994) reports that while most other aspects of the educational system have changed over the last two centuries, "there have been only two basic types of grading systems" (p. 37). Cizek (1996) echos this finding writing "despite all the other changes, a student's educational performance is still primarily reported using grades" (p. 103).

Although there have certainly been individual teachers and schools who have done otherwise, these marks were generally assigned, hence externally imposed upon, students. What follows is a summary of the research literature on the summative student evaluation practice of assigning marks to students.

The Relationship of Marks to Motivation

The basic question of interest here is whether marks motivate students, are neutral, or hurt student motivation? Ebel (1974) argues that "grading does tend to

motivate and direct study" (p. 3). On the other hand, Kirschenbaum (1973) writes "pupils learn to perform for the grade and as a result show less initiative, independence, self-motivation and creativity" (p. 46). Given these polar differences of opinion, it is not surprising to find the empirical evidence somewhat limited and unclear. Stiggins (1999) perhaps summed this up best. In explaining how grades tend to motivate some students while leading to the disengagement of others he writes, "the relationship between assessment and student motivation is complex indeed" (p. 197).

Brookhart (1997) has proposed a model to investigate the impact of classroom assessment on student effort and achievement. Given that effort is an indicator of motivation, this model is of interest in discussing the relationship between marks and motivation. The intervening variables in Brookhart's model involve student perceptions; they are (a) perceived task characteristics, and (b) perceived self-efficacy. Brookhart argues that classroom assessment impacts these perceptions, and the perceptions in turn impact effort and achievement. In a recent study designed to investigate this theoretical framework the researchers confirmed the complexity of the relationship between marks and motivation (Brookhart & DeVoge, 1999). The results of this study generally provided support for Brookhart's model. In other words, perceived self-efficacy did impact effort and achievement. Regarding marks specifically, they were found to impact self-efficacy. Hence, marks impacted self-efficacy, which in turn impacted effort.

In further discussing the role of marks, Brookhart and DeVoge (1999) note the

distinction between informational feedback and controlling feedback described by Ryan, Connell, and Deci (as cited in Brookhart & DeVoge, 1999). While "informational feedback" provides information students can use to improve their achievement, "controlling feedback" simply provides a judgement. Marks would certainly fall under the category of controlling feedback. In general, Brookhart and DeVoge (1999) found that controlling feedback "may influence future learning through students' use of it as evidence for their capability to succeed at a particular kind of assessment" (p. 422). The role marks played in Brookhart's model, however, was dependent on the types of marks students had received in the past. Brookhart and DeVoge found that "students' self-efficacy judgments about their abilities to do particular classroom assessments were based on previous experiences" (p. 422). In other words, students who had done well on previous assessments, had increased self-efficacy and hence increased effort. Students who had done poorly, had decreased self-efficacy and hence decreased effort. In summary, while marks clearly played a role in impacting motivation (as measured by effort), the direction of the impact was dependent upon the previous marks.

Yet another finding from the study by Brookhart and DeVoge contributes to the complexity of this issue. They found that in some cases, students who had performed well on previous assessments demonstrated less effort. They postulated this may be due to a degree of confidence that led them to believe effort was not necessary. Whether or not these students are still motivated, would likely be

dependent on the researcher's operationalization of motivation.

While Brookhart's model of classroom assessment and it's impact on motivation is dependent on using effort as a proxy for measuring motivation, it is also possible to examine the role of marks and motivation by using a proxy for marks. For example, any type of extrinsic reward that is offered to students may be viewed as a proxy for marks. Indeed, Arnstine and McDowell (1993) write:

> Because students believe they are assigned on the basis of merit, grades function for them as rewards. Because they don't always care about the subject under study, students learn to study "for the grade." Thus, grades can be conceived as a form of merit pay for children. (p. 12)

If one accepts this argument, then the literature relating extrinsic rewards (like merit pay) to motivation would be applicable to the discussion of the impact of marks on motivation. Interestingly, those who have researched extrinsic rewards and motivation, like marks and motivation, have found the relationship to be complex. After conducting a meta-analysis of 96 experimental studies, Cameron and Pierce (1994) conclude that rewards do not necessarily decrease intrinsic motivation. However, in discussing the research that claims rewards do negatively impact motivation, they write, "while several researchers agree with this conclusion, . . . others continue to favor the use of reinforcement principles in applied settings" (p. 364). They also point out that some researchers have "noted the contradictory nature of the findings and have attempted to identify the conditions under which extrinsic

reward produces decrements in intrinsic motivation" (p. 368). The fact that questions like this exist, provides clear evidence that there is not a direct, unmitigated relationship between the assignment of marks and motivation.

A final method for investigating the relationship of marks to motivation is to compare the practice of assigning marks to other practices of summative evaluation. Butler and Nisan (1986) designed a study to test the impact on motivation when sixth grade students were given either marks, task-related comments, or no feedback. Clearly, task-related comments or no feedback are two possible alternatives to marks. They found that motivation was impacted positively by the receipt of task-related comments and negatively by the receipt of marks or no feedback. In summarizing the effects of marks on motivation, Butler and Nisan write:

> Our results suggest, as some critics argue (Holt, 1964; Silberman, 1970), that the information routinely given in schools - that is grades - may encourage an emphasis on quantitative aspects of learning, depress creativity, foster fear of failure, and undermine interest. (p. 215)

In summary, depending on how both marks and motivation are conceptualized, as Butler and Nisan concluded, the assignment of marks *may* be detrimental to motivation. The key word in this conclusion is the word "may." The relationship between marks and motivation is neither direct, nor applicable in all settings or with all students.

The Relationship of Marks to Achievement

To some degree, the relationship between the assignment of marks and achievement is itself related to the relationship between marks and motivation. Recall the model for understanding the role of classroom assessment proposed by Brookhart (1997). In this model, she suggests marks impact perceptions, which in turn impact effort, which in turn impacts achievement. Indeed, it seems reasonable that increased motivation or effort would lead to increased achievement. Hence, based on the literature relating marks to motivation, our initial conclusion should be that the assignment of marks *may* be detrimental to increased student achievement.

In spite of the fact that educators have been arguing for and against the use of marks for nearly 100 years, there are very few empirical studies that investigate their relationship to achievement. There have, on the other hand, been a number of studies that have attempted to investigate the relationship between the general construct of classroom assessment and student achievement. Richard Stiggins (1999) aptly summarizes the findings of those studies when he writes, "we have reason to believe that an investment in the quality of classroom assessment will pay major dividends" (p. 193). Again, however, this positive finding is relating classroom assessment to achievement. It is not specifically relating the use of marks to achievement.

The studies that have come the closest to relating marks to achievement are those that have investigated the impact of marks as opposed to other types of feedback. For example, Page (1958) compared the effects when secondary students

were given only a mark, when they were given a mark and a predefined comment that matched the mark, and when they were given a mark and a "free comment" determined by the teacher and related to the student's work. He concluded that "Free Comment students achieved higher scores than Specified Comment students, and Specified Comments did better than No Comments" (p. 180). The design of Page's study was very similar to the study in which Butler and Nisan (1986) concluded task-related comments increased student motivation. In fact, as another example of the dependent nature of achievement and motivation, Page, in his conclusions, writes:

> When the average secondary teacher takes the time and trouble to write comments (believed to be "encouraging") on student papers, these apparently have a measurable and potent effect upon student effort, or attention, or attitude, or whatever it is which causes learning to improve. (p. 180)

In other words, while Page's study demonstrated that achievement could be raised by augmenting assigned marks with teacher comments, he speculated that the relationship between the summative assessment and achievement was moderated by some other variable that was closely related to motivation.

Stewart and White (1976) conducted an analysis of 13 studies that attempted to replicate Page's findings. Unfortunately, the majority of the studies were unable to show a statistically significant effect. Nevertheless, Stewart and White make the relatively modest claim that "the depression of student performance under the no-comment treatment in the present study and other statistically nonsignificant

replications (Moody, 1970; Simons, 1971; Sweet, 1966) represents a trend toward comment effectiveness which is more consistent with Page's finding than with the null hypothesis" (p. 497).

Kluger and DeNisi (1996) conducted a meta-analysis of 131 studies investigating the relationship between feedback in general and performance. Consistent with the research that has already been discussed on marks and motivation as well as marks and achievement, they found the relationship to be mixed and complex. While their overall analysis showed that feedback improved performance, in more than a third of the studies feedback was found to diminish performance. In discussing the conditions under which feedback may lead to increased performance, they suggest the feedback must be "containing cues that support learning, attracting attention to feedback-standard discrepancies at the task level . . . , and is void of cues to the meta-task level (e.g., cues that direct attention to the self)" (p. 278). This does not describe the nature of marks. Indeed, marks alone are unlikely to provide any cues that support learning, they provide no specific evidence of discrepancies between the mark and the standard, and, they frequently do direct attention to the self via their normative nature. Hence, the research on feedback in general suggests that there are likely better methods of summative student assessment.

A final line of study which is somewhat related to the relationship between marks and achievement would be studies that are designed to investigate the relationship between the frequency of assessments and achievement. In a meta-

analysis of 40 such studies, Bangert-Drowns, Kulik, and Kulik (1991) found that a higher frequency of assessment activities was positively related to student achievement. The type of feedback provided on these assessments was not investigated, simply the frequency of the assessments. Hence, the conclusion should not be drawn that the giving of marks impacted achievement, indeed, other types of summative assessments methods, if applied more frequently, ought also to provide the intended effect of increased achievement.

It is important to note that these studies are investigating the impact of classroom assessment on achievement in general. In the cases where a positive relationship was found, it was always when some form of summative assessment other than marks alone was used. Unfortunately, as has already been discussed, marks alone are the most common form of summative assessment in use in schools today. Perhaps this is what prompt Hin-Wai Yung (2002) who writes, "the forms that examinations and assessment take are widely recognised as determinants of educational practice" (p. 97).

In summary, classroom assessment has the potential to positively impact student achievement; however, if all forms of classroom assessment serve a summative role, and, the only feedback given to students is a letter grade (or a mark), this may actually inhibit student achievement. While this study will not investigate the impact of assigning marks on achievement directly, it is anticipated that, if the assignment of marks impacts self-regulated learning, that will in turn impact achievement.

Reliability of Marks

As early as 1912, Starch and Elliot conducted a study investigating the reliability of marks. In their study, 200 English teachers were each asked to assign a percentage grade to each of two student papers. Seventy-five percent was considered to be the lowest passing grade. The two papers were graded and returned by 142 teachers. On one of the papers the scores ranged from 64% to 98%. The range on the other was from 50% to 97%. The following year, Starch and Elliot repeated the study with geometry papers and found an even wider range of scores.

Starch and Elliot framed their study by investigating the impact of different teachers assigning marks to the same paper. Using this design, they were able to raise serious concerns regarding the reliability of marks. According to Cizek (2000), "although these studies [referring to the work of Starch and Elliot] identified the problem, little progress has been made in improving it" (p. 17). Cizek concludes, "the bottom line is that, since the introduction of grades, little progress has been made in improving their reliability" (p. 17).

In contrast to this conclusion, some researchers have attempted to show that grades are reliable (Smith, 1992; Werts, Linn & Joreskog, 1978). These studies investigated the correlations between students grades across a discipline or across the entire curriculum. For example, Smith (1992) found that the median correlation between marks in pairs of classes within the same discipline ranged from a low of 0.32 to a high of 0.57. He notes that each of these was significant at the alpha = 0.05 level.

He concludes, “the relationship between pairs of courses also indicated the reliability of grades” (p. 335).

There are at least two flaws in a design that claims to provide support for the reliability of marks based on correlations between courses. First, testing a null hypothesis that the correlation between the two groups of grades is zero is not a very significant (practically speaking) statistical test. Smith may well be able to claim that there is a non-zero correlation between marks assigned in pairs of college courses within a discipline. However, I would claim that the correlations he reported are not significant when the term retains its traditional, non-statistical, meaning. The second problem with this design is even more serious. Even if one could show a high level of correlation between marks in two classes, that would not necessarily lead to a conclusion that the marks were reliable. Indeed, it may be true that certain students generally earn high marks and others generally earn low marks. However, how these overall, end-of-term marks were determined may very greatly. It seems the design employed by Starch and Elliot, namely, allowing different teachers to assign marks to individual assignments, is a much more telling technique for investigating the reliability of marks. In summary, the reliability of marks appears to be questionable at best, even after considering studies that claim to be supporting their reliability.

Validity of Marks

Questions regarding the validity of marks are even more important than those

pertaining to their reliability. Indeed, even if one were to conclude that marks are assigned in a consistent manner, this would be of no value if there was no clear consensus regarding the meaning of those marks.

In summarizing the opinions of experts in the field of educational measurement regarding marks, Friedman (1998) writes, "attitude, effort, ability, behavior, and the like should *not* affect grades at all, because grades should reflect the level of student achievement in each course" (p. 78). In spite of this, studies have consistently shown that these factors *are* affecting the marks students receive (Blount, 1997; Frary, Cross & Weber, 1993; Friedman & Manley, 1992; Friedman & Troug, 1999; Stiggins, Frisbie & Griswold, 1989).

Effort appears to be the most prominent factor for inclusion beyond student achievement. In a survey of teachers, administrators, counselors, parents and students, Friedman and Manley (1992) found that all groups believed motivation/effort should play some part in determining a student's final grade. Friedman and Troug (1999) investigated the written grading policies of 53 high school teachers. They found that 9% explicitly included effort and for another 23% the policy was too vague to tell if effort was included. Although that leaves 68% for whom the written policy does not include effort, there is no guarantee that those teachers strictly follow their written policies. Indeed, Blount (1997) conducted one-on-one interviews with 58 practicing teachers. He writes:

These teachers were clear that they considered student effort in the grading

> process, even though authorities consider including effort in the grade as grade contamination. Many school districts have regulations against it (Anderson and Wendel 1988). In practice, however, these teachers reward students who try, seeking ways to raise their numbers. (p. 332)

Blount concludes that fully 86% of the teachers include effort as a determining factor in marking students. In their study of high school teachers, Stiggins, Frisbie, and Griswold (1989) found that 80% of the teachers believed effort should be included as a factor in calculating a student's grade.

In addition to effort, other non-achievement factors are also frequently included when determining student's final marks. In their review of written policies, Friedman and Troug (1999) found that 17% of the policies explicitly included "attitude." Another 32% included "behavior" and 43% included "attendance." Again, these numbers reflect instances where non-achievement factors were specifically included in written grading policies. It is reasonable to assume that the numbers would be even higher had the investigation examined teachers actual grading practices. Blount (1997) concludes, "effort, behavior, and attitude were clearly part of the grading process" (p. 331). Friedman and Manley (1992) found that teachers believed "learning ability should be taken into account when assigning quarter/semester grades" and "attitude should be considered" (p. 102). Friedman and Troug (1999) summarize the concerns regarding the validity of marks when they pose the following rhetorical questions:

> How is validity affected by some of the grading practices used here? If teachers are using grades as a way to manage student behavior - i.e., being in one's seat when the bell rings or following directions - to what extent are grades describing actual student performance? (p. 41)

To summarize, consider the following example. If the problem is relatively limited, and marks are used to measure only one additional construct beyond achievement, say effort for example, this can still cause serious problems with the validity of the mark. Let's say the teacher has assigned 50% of the mark based on effort and the other 50% based on achievement. A final grade of "C" may mean the student has an "A" level of achievement and an "E" level of effort, or, it may mean an "E" level of achievement and an "A" level of effort, or, the student may be at the "C" level in both categories. Clearly, the most prevalent form of summative student assessment in secondary schools today, namely the assignment of marks to students, is extremely limited in terms of its use for drawing valid conclusions regarding anything.

Summary of the Literature on Marks

The mixed nature of the literature regarding the relationship of marks to motivation, the limited amount of empirical data demonstrating its relationship to student achievement, and the agreement of researchers regarding the problems in developing a reliable and valid marking system, all provide strong justification for investigating alternatives to the use of marks for summative student evaluation. This

study proposes to investigate two such alternatives. The first, is to completely eliminate the use of marks by allowing students to develop their own summative evaluations. The second, will attempt to minimize the concerns regarding marks by using a credit/no-credit structure of summative student evaluation.

Mathematics Education

This study will be limited to an investigation within a mathematics classroom. Thus, a brief review of literature related to mathematics education is appropriate. Mathematics education may be dichotomized as (1) the content (or curriculum) of mathematics education, and (2) the delivery of instruction. While the debate over the content of mathematics is indeed interesting, it is beyond the scope of this study. The content within the classrooms involved in the study was clearly defined based on the Indiana Graduation Qualifying Exam (GQE). The GQE is in turn based on the Indiana State Standards for Mathematics. Again, the nature of these standards and their appropriateness in terms of defining essential mathematics content is beyond the scope of this study. On the other hand, the delivery of instruction is a critical variable that is not yet clearly defined for this study. A brief review of literature related to the delivery of mathematics instruction follows. Based on that review, a decision is made regarding the delivery of mathematics instruction for this study.

The delivery of mathematics curriculum has received at least as much attention as the curriculum itself. According to the Mathematical Sciences Education Board

(1991) goals for teachers are shifting from authoritarian models based on "transmission of knowledge" and "drill and practice" to student-centered methods featuring "stimulation of learning" and "active exploration." The same report states that classrooms need to:

- Encourage students to explore.
- Help students to verbalize their mathematical ideas.
- Show students that many mathematical questions have more than one right answer.
- Teach students, through experience, the importance of careful reasoning and disciplined understanding.
- Provide evidence that mathematics is alive and exciting.
- Build confidence in all students that they can learn mathematics. (p. 7)

Teaching mathematics from the above perspective has sometimes been labeled as a constructivist approach to mathematics instruction. According to Jones and Wilson (1997), the core belief of constructivism is "that knowledge is not transmitted directly from the teacher to the student. Instead, the learner constructs knowledge through active engagement in the process of assimilating information and adjusting existing understandings to accommodate new ways of knowing" (p. 160). This implies a need for active discussions in mathematics classrooms. Discussions not just between the teacher and the students, but within subgroups of students. Students need to test their mathematical ideas against those of others so that they can be confident

that they are correct. The only way we have of knowing what our thinking is like is by comparing it with that of others. The major pedagogical implication of this view is that mathematics classrooms should be communities of learners (Silver, Kilpatrick, & Schlesinger, 1990).

More recently, data from the videotape portion of the Third International Mathematics and Science Study (TIMMS) has been used to argue for a more constructivist approach to mathematics education (Bracey, 2000; Geist, 2000; Stigler & Hiebert, 1997). Stigler and Heibert (1997) used the videotape study to analyze the differences between mathematics instruction in Japan, Germany, and the United States. They had previously concluded that the achievement level in the United States was lower than either Japan or Germany. This led to an attempt to uncover instructional practices that may account for the differences in achievement. They found significant differences among the three countries. In the United States, topics tended to be taught simply by stating the information. In both Japan and Germany, topics were developed. Seatwork in the United States was spent primarily on practice, with almost no time on inventing or thinking activities. In Germany, the vast majority of seatwork time was also spent on practice; however, some time was spent on inventing and thinking. In Japan, students actually spent more time on inventing/thinking activities than they did on practicing procedures. This mindset was also evident in the goals teachers had for lessons. In the United States, almost three times as many teachers said the goal of the lesson was to develop skills as opposed to thinking. In Japan, approximately three

times as many teachers believed the goal of the lesson was thinking as opposed to skills. These results suggest a more constructivist approach to teaching in Japan as opposed to the United States. Bracey (2000) summarizes the differences between mathematics classrooms in the two countries by describing the Japanese classroom in this way,

> the teacher would review the previous day's problem, present the problem of the day, and set the students to working on its solution either individually or in groups. The whole class would then discuss the solutions, usually with students who thought they had solved the problem leading the discussion from the blackboard - something that almost never happened in American classes. (p. 474)

Beginning with the publication of the National Council of Teachers of Mathematics Standards for School Mathematics in 1989 and continuing through recent analysis of the Third International Mathematics and Science Study, mathematics education experts have consistently argued for a more constructivist approach to mathematics instruction. Clearly, whichever approach to instruction a teacher chooses to apply in his or her classroom, this approach will quite likely, in and of itself, have an impact on self-regulated learning. It is thus important to insure the type of instruction remain constant throughout this study. It also seems wise to implement the most widely recommended instructional approach. For these reasons, this study will attempt to investigate the impact of structure of summative student evaluation on self-

regulated learning within a secondary mathematics classroom in which the method of instruction, or the delivery of the curriculum, is aligned with the recommendations of the above theorists.

Specifically, the classroom will function much like Bracey's summary of the Japanese classroom cited above. This researcher will be the instructor in each experimental classroom. I possess a masters degree in mathematics education and have experience in implementing a constructivist approach to mathematics education. By implementing the most widely recommended approach to mathematics instruction and holding it constant throughout the study, it is anticipated that the research design will effectively examine the impact of types of summative student evaluation within an "ideal" mathematics classroom.

Self-Regulated Learning

The concept of empowering core workers is one that crosses a number of disciplines. In each case, the theory posits that empowered workers will be more productive. In educational literature, this concept has been researched extensively and categorized under the construct of self-regulated learning. While self-regulated learning always deals with the degree to which the learner is empowered, or takes ownership of his or her learning, authors have operationalized the construct in a wide variety of ways.

Corno and Mandinach (1983) offer one of the earliest efforts at

operationalizing self-regulated learning. They write, "self-regulated learning will be shown to consist of specific cognitive activities, such as deliberate planning and monitoring, which learners carry out as they encounter academic tasks" (p. 89). They further categorize these cognitive activities into five processes which they view as "both necessary and sufficient for defining self-regulated learning in complex tasks" (p. 94). Their five components of self-regulated learning are alertness, selectivity, connecting, planning, and monitoring. It is important to note that Corno and Mandinach distinguish between cognitive activities the learner initiates and carries out entirely on his or her own and activities the learner may initiate that require others (i.e., peer or expert assistance). They view the latter as "resource management," which they claim is distinct from self-regulated learning.

Barry Zimmerman (1986), another foundational researcher on self-regulated learning, expanded the concept to include what Corno and Mandinach had separated out as resource management. Zimmerman categorized these strategies under "seeking social assistance," and he further delineated seeking assistance from peers, teachers, and adults. These three types of seeking social assistance were three of the fourteen strategies of self-regulated learners Zimmerman identified in his model. Although Zimmerman's model expanded the concept of self-regulated learning to some degree, it still limited it to a learner's use of cognitive strategies.

Zimmerman's model has been cited by many researchers since. Some authors use this model directly (Ablard & Lipschultz, 1998; Williams, 1996). In other cases,

where authors have constructed their own model, pieces of Zimmerman's model can often be found. VanderStoep, Pintrich and Fagerlin (1996), for example, conceptualize the construct as having three facets. From Zimmerman, they identify one facet as a learner's use of cognitive strategies. To this they add the learner's motivational beliefs (including intrinsic orientation, self-efficacy, and task-value beliefs) and the learner's prior knowledge base. In an earlier model, Pintrich & DeGroot (1990) had conceptualized self-regulated learning as consisting of cognitive strategies, metacognitive strategies (both from Zimmerman) and management of effort. Management of effort referred to a learner's willingness and ability to persist in a task. Likewise, Lindner and Harris (1992) identify five facets of self-regulated learners. From Zimmerman, they borrow cognitive and metacognitive strategies, and like VanderStoep, Pintrich, and Fagerlin they include motivation/self-efficacy. Their unique contribution to the model is an identification of contextual facets. They identify these as contextual sensitivity and environmental utilization/control.

Although there is no universally accepted operationalization of self-regulated learning, the primary factors that are included in each model are relatively consistent. Based on these models, I have identified four aspects which will constitute an operationalization of self-regulated learning for this study. Like all models, Zimmerman's use of cognitive strategies is one aspect. I am also including self-efficacy, motivation, and mastery orientation to learning. Further elaboration on each of the four aspects follows.

Self-Efficacy

Significant research has been done regarding the concept of self-efficacy as one component of self-regulated learning. Dale Schunk (1984) defines self-efficacy as "personal judgments of one's capability to organize and implement behaviors in specific situations" (p. 48). His research found that higher levels of self-efficacy were associated with higher achievement (Schunk, 1984). Other researchers have also found that a learner's perception of his or her academic efficacy either positively affects achievement or is a necessary prerequisite for achievement (Bong, 2002; Covington & Omelich, 1985; McCombs, 1984; Pajares, 1996; Pajares & Valiante, 1997; Shen & Pedulla, 2000).

The positive relationship between self-efficacy and achievement is likely neither direct nor simple. According to Pajares (1996),

> Self-efficacy beliefs act as determinants of behavior by influencing the choices that individuals make, the effort they expend, the perseverance they exert in the face of difficulties, and the thought patterns and emotional reactions they experience. It is for these reasons that high self-efficacy is likely to promote stronger academic performances . . . (p. 325)

Indeed, in many cases where a positive correlation was found between self-efficacy and achievement, at least one mediating variable was present. Schunk (1999) reported that while self-efficacy had a direct impact on achievement, it also had an indirect impact through the mediating variable of persistence. Similarly, in their study of

elementary students' writing achievement, Pajares and Valiante (1997) found that "students' self-efficacy beliefs about their own writing capability directly influenced their writing apprehension, perceived usefulness of writing, and essay-writing performance" (p. 357). As another example, Shen and Pedulla (2000) used data from the Third International Mathematics and Science Study to investigate the relationship in the disciplines of math and science. They write, "in terms of self-efficacy theory, our within-country data analyses support the theory defining the self-evaluation - motivation - academic performance relationship" (p. 250). In other words, their analysis of the data suggested that positive self-evaluations led to increased motivation which in turn led to increased achievement.

As an example of a study that did not specifically consider mediating variables between self-efficacy and achievement, Bong (2002) attempted to investigate the relationship between subject specific self-efficacy and achievement in other subjects. Of interest to the argument being made here, she concluded that, "self-efficacy perceptions in a specific school subject indeed linked most strongly to achievement measures in the same subject" (p. 158). One of the subjects Bong included in her study was mathematics.

In investigating the relationship between self-efficacy in mathematics and either mathematical achievement in general or mathematical problem solving ability in specific, a number of research studies have found that self-efficacy has direct positive effects (Hackett, 1985; Pajares, 1996, Pajares & Kranzler, 1995; Pajares & Miller;

1994; Schunk, 1981). Of particular interest to this study is the work of Pajares and Kranzler (1995) with high school students within the public school setting. Given that the group of students that will be investigated in this study is a subgroup of that (namely low-achieving public high school mathematics students), a conclusion by Pajares and Kranzler that self-efficacy produces positive outcomes would provide strong justification for the inclusion of self-efficacy in this study's operationalization of self-regulated learning. Indeed, Pajares and Kranzler (1995) concluded "that students' self-efficacy beliefs about their math capability had strong direct effects on math anxiety and on mathematics problem-solving performance even when general mental ability was controlled" (p. 437).

The research regarding the relationship between self-efficacy and achievement provides clear evidence that a positive correlation exists. For this reason, the learner's self-perception of his or her mathematics efficacy is being included as one component of self-regulated learning in this study.

Motivation

The relationship of motivation and self-regulated learning, or the role motivation plays as a part of self-regulated learning, has been discussed by a number of researchers. Corno and Mandinach (1983) wrote that self-regulated learning (cognitive strategies in their view) yielded motivated students. Garcia, McCann, Turner, and Roska (1998) argued the converse, namely, that motivated students

impacted the learner's use of cognitive strategies. In either case, a relationship was shown to exist between motivation and the use of cognitive strategies. This should not be surprising in that many theorists now include both motivation and the use of cognitive strategies as factors within the construct of self-regulated learning (as has already been discussed). I too will include motivation as one factor in the operationalization of self-regulated learning for this study.

The next question of interest is then, "what is motivation?" In their study, Garcia, McCann, Turner, and Roska (1998) conceptualized motivation as consisting of the following four facets: 1) intrinsic goal orientation, 2) task value, 3) self-efficacy, and 4) test anxiety. In addition to Garcia, et. al., a few other researchers have conceptualized self-efficacy as a piece of the construct of motivation (Lindner & Harris, 1992; Pintrich & De Groot, 1990; VanderStoep, Pintrich & Fagerlin, 1996). While self-efficacy may be an indicator of motivation in some situations, it may not always be so. For example, a student may be very confident in his or her ability to complete a task yet still lack the motivation to persevere in the task. For this reason, this study will consider self-efficacy as a separate factor, distinct from motivation. However, the two factors are clearly not independent. Indeed, as was discussed in the section on self-efficacy, some researchers view motivation as a mediating variable between self-efficacy and achievement. While there is likely a complex interaction between these two factors, the qualitative design proposed in this study should allow appropriate flexibility for identifying their presence as indicators of self-regulated

learning.

Returning to the question of the operationalization of motivation, there are now three factors left to consider, 1) intrinsic goal orientation, 2) task value, and 3) test anxiety. In this study, the independent variable is summative student evaluation. Hence, the method of testing varies among classrooms. For this reason, the primary evidences that are considered when interpreting data regarding student motivation are intrinsic goal orientation and task value.

Mastery Orientation to Learning

The third factor that is being used to operationalize self-regulated learning in this study is the learner's approach to a task. The phrase "approach to the task" in this context is used to refer to the ultimate goal the learner has for the learning task. If the goals of students could be placed on a continuum ranging from "pass the class" to "understand in-depth all of the mathematics that is presented as well as related topics," students whose goals are closer to the latter statement would be viewed as stronger self-regulated learners.

Researchers have used varying terms to define student approaches to learning. Ames (1992) contrasted a "mastery orientation" with a "performance orientation." Students who approach a task from a mastery perspective "are oriented toward developing new skills, trying to understand their work, improving their level of competence, or achieving a sense of mastery based on self-referenced standards" (p.

262). This contrasts with a performance orientation where "attention is directed toward achieving normatively defined success" (p. 262). Other authors who have conceptualized the same dichotomy, however, have used different terms. Dweck (1986) referred to "(a) *learning goals*, in which individuals seek to increase their competence, to understand or master something new, and (b) *performance goals*, in which individuals seek to gain favorable judgments of their competence or avoid negative judgments of their competence" (p. 1040). Maehr and Nicholls (as cited in Ames, 1992) labeled the two approaches "task-involvement" and "ego-involvement." Although the name attached to the construct varies from "mastery orientation," to "learning goals," or "task-involvement," the definition is relatively consistent. For the purposes of this study I have chosen to use the term "mastery orientation to learning."

Both Stipek and Kowalski (1989) and Graham and Golan (1991) have shown that a mastery orientation (or task-involvement as each refers to it) leads to increased student achievement. The work of Stipek and Kowalski is particularly pertinent to this study as they distinguished between "low-effort" and "high-effort" students in their investigation. "Low-effort" students were defined as students who tended to believe that increased effort would not lead to increased achievement. In other studies, students who hold to these beliefs are referred to as "helpless" (Diener & Dweck, 1978; Licht & Dweck, 1984). It is reasonable to assume that the characteristics associated with "low-effort" or "helpless" students may well be associated with the low-achieving students of this study. In their study, Stipek and Kowalski provided

some students with task instructions that focused on mastery of the task while providing other students with instructions that focused on the outcome or performance level. In effect, they were attempting to influence the orientation or approach to learning of these students via the instructions they were given. They then investigated the students' use of problem solving strategies, hypothesizing that the students who were given instructions to focus their attention on mastery of the task would use effective strategies more frequently than those who were given performance oriented instructions. While Stipek and Kowalski were not able to attribute an effect based on mastery orientation in the case of "high-effort" students, they did detect an effect in the case of "low-effort" students. They write, "children who tended to de-emphasize effort as a cause of achievement outcomes were more likely to use effective strategies consistently when they were given instructions designed to allay concerns about performance and to focus their attention on the task" (p. 389). This result provides justification for including a mastery orientation to learning as an appropriate indicator of self-regulated learning.

In the study conducted by Graham and Golan (1991), they use the term "task involvement" to refer to a mastery orientation to learning. However, while Stipek and Kowalski referred to the alternative orientation as "performance oriented," Graham and Golan label it "ego involvement." They write,

> Task-involving states are those in which one's goal is to master the task: Greater understanding or acquisition of new skills is considered an end in itself.

> This is contrasted with ego-involving states in which one's primary goal is to demonstrate high ability relative to others or to conceal low ability. (p. 187)

They provided students with either a "task-involving" motivational condition, an "ego-involving" motivational condition, or neither. They then investigated the students' performance at both "deep levels of processing" as well as "shallow levels." They found that students who were given the ego-involving motivational condition performed more poorly relative to the deep levels of processing than either the task-involving group or the control group. These results not only provide justification for a mastery orientation to learning, they also suggest that traditional forms of summative student evaluation (like marking students) may well be detriments to deeper levels of learning. Graham and Golan write:

> If this analysis is accepted, then it appears that ego involvement might well interfere with the cognitive effort needed for deeper levels of information processing. These findings have important implications for both learning and instruction. With its emphasis on competition, grades, and various other kinds of normative evaluation, much of school learning becomes synonymous with an ego-focused context, the very motivational state that we and others have found to have detrimental consequences for performance. Furthermore, contemporary approaches to instruction, guided by cognitive views of learning, attach ever greater importance to meaningfulness, elaboration, and various other cognitive activities that require deep rather than shallow processing (see

> Glover, Roning, & Bruning, 1990). As our data indicate, these are the kinds of learning activities that might be most vulnerable to the effects of a maladaptive motivational state. (p. 193)

In summary, a mastery orientation to learning is a desirable student outcome, and marking students may well serve as an impediment to a mastery orientation to learning.

Ames (1984) investigated the hypothesis that marking students impedes a mastery orientation in a study of fifth and sixth grade students. Students were placed in either a "competitive" setting (analogous to marking students) or an "individual setting." According to Ames, "the individual setting involved no external performance criteria. Instead, children were encouraged to try and were provided with an opportunity for self-improvement over a series of trials" (p. 484). She concludes:

> The individualistic structure has been shown as a sharp contrast to the competitive structure in that the individual setting appears to facilitate a task focus evidenced by the greater use of self-instructions and to reduce the ego focus that is endogenous [sic] to the competitive structures evidenced by greater ability attributions. Whereas children in the competitive structure exhibited a strong ability focus, children in the individualistic structure appeared more achievement or mastery oriented. (p. 486)

The conclusion reached by Ames, suggests that a summative evaluation structure that de-emphasizes marks will lead to a mastery orientation to learning.

Clearly, the inclusion of a mastery orientation to learning as one aspect of self-regulated learning is appropriate. Indeed, there is evidence to suggest the treatments offered in this study may positively impact self-regulated learning.

Use of Cognitive Strategies

Early models of self-regulated learning focused exclusively on the learner's use of cognitive or meta-cognitive strategies. While more recent models have expanded the definition to include additional factors, the use of cognitive strategies is still an integral factor in all models. Indeed, much of the research that suggests self-regulated learning can lead to higher achievement is, in fact, either in whole or in part, suggesting that a learner's use of cognitive strategies can lead to higher achievement (Linder & Harris, 1992; Pintrich & De Groot, 1990; VanderStoep, Pintrich & Fagerlin, 1996; Williams, 1996; Zimmerman & Martinez-Pons, 1988; Zimmerman & Martinez-Pons, 1986).

Based on the abundance of evidence linking the use of cognitive strategies to student achievement, and based on the fact that all other models include this factor when operationalizing self-regulated learning, the use of cognitive strategies will be the fourth and final factor included in the operationalization of self-regulated learning in this study.

Interaction of Self-Regulated Learning Factors

The four factors that are being used to operationalize self-regulated learning are not independent of one another. The nature of some of the dependencies was discussed in each individual section; however, some researchers have specifically studied these dependencies. For example, Zimmerman and Martinez-Pons (1990) conducted a study with students in grades 5, 8, and 11 that was designed to determine if self-efficacy impacted self-regulated learning. Their operationalization of self-regulated learning was limited to the use of cognitive and meta-cognitive strategies; hence, their study was actually investigating the relationship between self-efficacy and the use of cognitive/meta-cognitive strategies. They found that the former can at times be used to predict the latter, hence the two factors are dependent. Of even greater interest is the following quote drawn from their discussion of the results:

> These developmental data suggest that instructional procedures that draw on or enhance students' perceptions of self-efficacy, such as participant modeling or mastery learning (Bandura, 1986; Schunk, 1984; Zimmerman, in press), may hold particular promise for motivating junior and senior high school students. (p. 57)

It is extremely interesting to note that while their study was designed to investigate the relationship between two of the factors used in this current study (namely self-efficacy and the use of cognitive strategies), when summarizing the results, Zimmerman and Martinez-Pons pull in both of the remaining two factors (namely motivation and a

mastery approach to learning). Note they state that self-efficacy may be bolstered through mastery learning and that this may then be "motivating" for students. Presumably, the motivation would then lead to the use of cognitive strategies.

In addition to Zimmerman and Martinez-Pons, other researchers have identified relationships among the four factors of self-regulated learning. Pintrich and De Groot (1990) found that self-efficacy was positively correlated with the use of cognitive strategies. Both Hagan (1994) as well as Shih and Alexander (2000) found that self-efficacy was positively correlated with the goal orientation of students. Finally, Malpass, O'Neil, and Hocevar (1999) defined self-regulation to include the use of cognitive and meta-cognitive strategies as well as motivational strategies. They then conducted a study with mathematically gifted high school students and found that both self-efficacy as well as a mastery goal orientation were positively related to self-regulated learning. Hence, they related all four of the factors used in this study to define self-regulated learning.

Clearly, interactions between the factors used to define self-regulated learning exist and are complex in nature. Depending on which study is referenced, any one factor may be cited as influencing any of the other factors. These interdependencies are shown graphically in figure 2.

The dependent nature of the four factors which define self-regulated learning is an important point in the design of the current study. While operationalizing self-regulated learning into four distinct factors provides a level of clarity regarding the

construct and a degree of confidence in approaching the data analysis, in some senses, this is a false confidence. The complex dependencies between the factors must still be accounted for. For these reasons, it will be necessary to design a study that is extremely flexible regarding the analysis of data. The design will need to allow each of these factors to be examined in the context of one another. A complete description of the design follows in chapter three.

Self-Regulated Learning

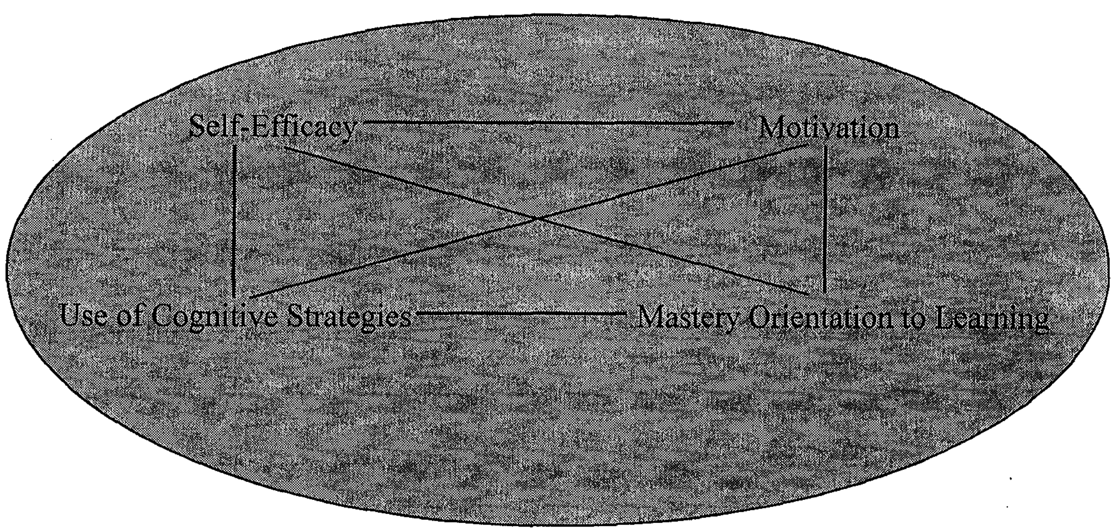

Figure 2. Self-Regulated Learning

CHAPTER 3

METHOD

The study is designed to investigate the following research questions:

1. How does the method of summative student evaluation impact self-regulated learning in low-achieving high school mathematics students?

2. How does the cognitive level of the task mediate the relationship between method of summative student evaluation and self-regulated learning?

3. How does the length of time students are exposed to a method of summative student evaluation impact self-regulated learning?

The focus of the study will be narrowed in three ways. First, because of the researcher's expertise and experience in the area of mathematics education, the study will investigate each of the research questions within the discipline of mathematics. Second, and again as a result of the researcher's thirteen years of experience as a secondary school teacher, the study will investigate the impact on secondary school students. Finally, due to the fact that the study is being conducted during a remediation summer school program, the study will investigate the impact on low-achieving secondary school mathematics students.

The following describes the method that will be used to answer these questions.

Methodological Assumptions

According to Glesne and Peshkin (1992), the choice of a research methodology depends on the researcher's view of the nature of reality. They suggest that quantitative methods are "supported by the positivist or scientific paradigm, which leads us to regard the world as made up of observable, measurable facts. In contrast, qualitative methods are generally supported by the interpretivist paradigm, which portrays a world in which reality is socially constructed, complex, and everchanging" (p. 6). The positivist tends to believe that the social sciences can be explained in the same way as the natural sciences. The interpretivist, on the other hand, believes that social phenomena cannot be studied outside of the context in which they occur. Miles and Huberman (1994) elaborate on the interpretivist perspective, suggesting that, "in this view, social processes are ephemeral, fluid phenomena with no existence independent of social actors' ways of construing and describing them" (p. 2).

Attempting to design an educational study that fits neatly within the scientific paradigm is risky at best. This is because of the fact that educational outcomes are indeed dependent on the context in which they occur. A study which does not take that into consideration lacks validity. In this study in particular, as was discussed in chapter two, the dependent variable, namely self-regulated learning, is a very complex construct. In addition, the relationship between self-regulated learning and structure of summative evaluation is very likely dependent on contextual issues. For example, if a student is given a low mark on some task in a class and his or her peers tease him

regarding that mark, the impact of the mark on self-regulated learning for that student may vary significantly from the impact of the same mark, on the same task, with the same student, in a class where no teasing occurs. There are likely an infinite number of examples of contextual issues like this that would affect the relationship between the structure of the summative evaluation and self-regulated learning.

It seems clear that studies which are designed to investigate the impact of particular interventions within schools and on students, ought to take into account complicated contextual factors. While the scientific paradigm is not well suited for this, the interpretivist paradigm is. Indeed, a study that is designed from the interpretivist paradigm would have the potential of discovering connections that may not even have been imagined prior to the study. For these reasons, a qualitative research methodology will be employed in this study. Details of the methodology as well as the research design follow.

Research Design

Data related to summative student evaluation and self-regulated learning was collected during the month of August 2001 in the context of existing summer school classrooms at an urban Midwestern high school. Approximately twenty students in each of three classrooms participated in cognitively varied tasks at each of three sessions. Different methods of summative student evaluation were assigned to each of the three classrooms. The details of the research design will be explained in turn.

Participants

The participants were drawn from a summer school program in mathematics for incoming 10th grade students who had been identified as not yet mastering the 9th grade mathematics curriculum administered by the South Bend Community School Corporation in South Bend, Indiana. The students were mandated to participate in the summer school program based on below standard scores on the Mathematics Achievement Level Test published by the Northwest Evaluation Association and teachers' recommendations. The researcher selected three classrooms of students within this summer school program to participate in the study.

The nationally normed test scores of participants in this study were all below the 30th national percentile. Eighty-six percent of the participants scored below the 20th national percentile.

Classrooms

Classroom 1, named traditional, was set up to simulate a traditionally graded classroom. Classroom 2, student-developed, was set up to simulate a classroom where the only summative student evaluation was student-developed. Classroom 3, named mastery, simulated a classroom that was graded using a credit/no-credit approach.

Method of Instruction

Each lesson was taught from a constructivist perspective. Students were given the opportunity to learn the material through a problem setting. Students were also given the opportunity to receive assistance in a variety of ways. A complete description of each lesson is included in Appendix A.

Lessons

Students who were required to participate in the South Bend Community School Corporation's High School Summer Remediation Program were scheduled into classrooms each morning of the week for three consecutive weeks. During each of these morning sessions the students were provided breakfast, mathematics instruction, English/language arts instruction, and lunch. The classrooms that participated in this study received mathematics instruction provided by the researcher for one hour on each of three mornings. The researcher is a certified mathematics teacher with a master's degree in mathematics education. The same four lessons were taught to each of the three classrooms during the three days of researcher led instruction. The first three lessons were drawn from a single topic from the Indiana State Standards for Mathematics. The fourth lesson involved a combination of two topics. A unique summative evaluation matched to each classroom was associated with each lesson.

The lessons were administered in three different sessions over a period of nine calendar days. Lesson 1 was administered to all three groups on Monday, lesson 2 to all three groups on Friday and lessons 3 and 4 to all three groups on the following Tuesday.

Lesson 1

The first lesson was designed from the topic of the "Pythagorean Theorem" under the Geometry standard. Students were given rulers and scratch paper. They were asked to measure and mark 3 inches to the right of the bottom left hand corner of their paper and then 4 inches above the bottom left hand corner. They were then asked to connect the marks forming a right triangle with legs of length 3 and 4 inches. Finally, they were asked to measure the hypotenuse and record its length. On the second sheet of paper, students were asked to repeat the task measuring 6 centimeters to the right and 8 centimeters up. Students were told that the numbers they were given were chosen to create "nice" integer answers. Students were also given the option of creating their own numbers to use when measuring to the right and up. Students who chose to exercise this option were told that the length of the hypotenuse may not be an integer. When students completed the measurements, the class came to consensus regarding the length of the two hypotenuses and recorded those on the board. They also recorded the information from any right triangles that were created

by students who chose to use alternative lengths for the legs of the triangle. Using the data recorded on the board, the researcher led a class discussion designed to identify patterns. Once the pattern was identified and summarized on the board as the Pythagorean Theorem, students were given a worksheet to complete. They were told that they may work alone or with one other person.

Lesson 2

The second lesson was designed from the topic of "measures of center" under the Statistics standard within the Indiana State Standards for Mathematics. Students were given rulers and asked to measure the length of each of their two pinky fingers. As measurements were completed, results were listed in a table on the board. A class discussion then ensued regarding how the class might summarize the results with one statistic. From the class discussion, the statistics of mean, median and mode were identified and defined. Students were then given a worksheet to complete and told they may work alone or with a partner.

Lesson 3

The third lesson was designed from the topic of "equivalent representations," a subtopic of the Indiana State Standard on Functions. For this lesson, students were placed in groups of three and given an activity to complete that required them to

complete a table of values and a graph of values for a problem setting involving a linear function. The researcher circulated among groups assisting and summarizing concepts as needed.

Lesson 4

The fourth lesson was taught immediately following the third (during the same class period). Students were again placed in groups of three and given an activity to complete. This activity asked the students to apply two topics under the standard of Algebra. Students were required both to write and to solve algebraic equations in one variable. Due to the fact that students were asked to apply two algebraic concepts, it is believed that the cognitive difficulty of this fourth lesson was greater than that of the first three lessons.

Summative Evaluation Conditions

Traditionally Graded Classroom

The first class simulated a traditionally graded classroom. Students were given a brief quiz at the end of each lesson and a letter grade was assigned. Report cards were not issued at the end of the summer school program, therefore, it may be argued that the grade given on the quiz did not have the same value as a grade given in a class during the traditional school year. To address this issue, and in an attempt to ensure

that the grade functioned as a reward, students' names were entered into a raffle four times for each A they earned, three times for each B, twice for each C, and once for each D. After all four lessons were completed and graded, a drawing was held with the winner receiving a $20.00 gift certificate to a popular local restaurant.

The intention of this experimental classroom was to model as closely as possible the summative student evaluation structure that is present in the majority of secondary school mathematics classrooms.

Student-Developed Summative Evaluation Classroom

The second class simulated a classroom in which the only type of summative student evaluation was student-developed. Students were not given a quiz at the end of each lesson. Nor were they given the associated rewards of raffle tickets. These students did not participate in a drawing for a gift certificate.

The summer school program was intended to prepare students for the Indiana Graduation Qualifying Exam in mathematics. As a prerequisite for a high school diploma, all Indiana students must either pass the exam or demonstrate that they know the state standards in an alternative way. Students in the experimental classroom that was simulating student-developed summative assessments were asked to think about alternative ways in which they might demonstrate their knowledge of the mathematical material that was being taught.

The intent of this experimental classroom was to create a setting in which traditional summative student evaluation had been completely eliminated. The goal was to model a classroom in which there would be no grades or other externally imposed rewards for success at all.

Credit/No-credit Classroom

The third classroom simulated a classroom that was graded using a credit/no-credit approach to grading. Students were given the same quiz as the first experimental class at the end of each of the four lessons. These quizzes were given a score of either credit or no-credit. If the quiz had five possible points, credit was given for scores of 3, 4 or 5. If the quiz had only four possible points, credit was given for scores of 3 or 4. Like the first group, for each credit a student earned, his or her name was placed in a drawing; however, no name was put in the drawing for students who did not earn credit on the evaluation. Again, like the first group, after the fourth lesson a drawing was held with the winner receiving a $20.00 gift certificate to a popular local restaurant.

The intention of this experimental classroom was to model a classroom setting in which the method of summative student evaluation fell somewhere between traditional grading and the complete elimination of grades. By assigning grades of only credit or no-credit some externally imposed measure was imposed on the

students, however, it was theorized that the potential detrimental effects of this grading would be lessened through the contraction of the traditional grading system into simply two categories.

Data Collection Methods

Three methods of data collection were used in the study. First, each classroom was videotaped as students completed each lesson. Due to the size of the classrooms and the large number of students in each, only a subset of student behaviors are visible and hence analyzable through the video. However, all verbal responses were recorded on the video and hence were available for analyzing. Both the video and audio portions of the tape were reviewed, coded, and analyzed as described in the following section.

The second method of data collection was the collection of student work that had been completed during each lesson. These artifacts provided a framework for conducting personal interviews with selected students from each classroom.

The final method of data collection was personal interviews conducted with four students from each classroom. Due to the relatively high absentee rate in each class, the first narrowing criteria for selecting interviewees was to identify students who were in attendance for all four lessons. Secondly, Limited English Proficient students and students with extensive special education needs were not considered for

interviewing. Finally, if possible, two girls and two boys were selected from each class.

The interviews were conducted with individual students on the two days following the completion of the fourth lesson. The interview was designed to gather information regarding self-regulated learning in the context of the four factors included in the operationalization of self-regulated learning. The interview protocol is included in Appendix B.

Prior to conducting the first interview, a draft of the interview protocol was piloted with a student from one of the three classrooms who met the interviewing criteria. The interview protocol itself was not changed as a result of the pilot interview; however, this researcher was made aware of the importance of allowing sufficient wait-time for the student to respond. During the pilot interview, the student was frequently slow to respond, therefore, the researcher interjected suggested responses. This led to very limited and shallow remarks from the student. Student responses from the pilot interview were not used in the data analysis. However, based on the pilot interview, the researcher was careful to remain quiet and allow the students to generate their own responses during the twelve interviews which were used. The interviews were reviewed, coded, and analyzed as described in the following section.

Data Analysis Methods

Research methods are frequently categorized as either quantitative or qualitative. Quantitative methods tend to be the most clearly defined and frequently implemented. When investigating the relationship between variables using quantitative data analysis methods it is necessary to randomly assign students to experimental and control groups. The only difference between the two groups should be the variable of interest (for example, the method of summative student evaluation). After the treatment has been applied to both groups, some type of measurement is taken (for example, a standardized test). If one group has a higher measure than the other, a statistical analysis is then performed to determine whether the difference is likely due to chance alone, or whether the difference is likely due to the treatment applied. The inappropriateness of quantitative methods for this study has already been discussed.

When rejecting the quantitative paradigm, great care must be exercised not to simply gather qualitative data and then attempt to apply quantitative concepts to the qualitative data. The following example, which I will refer to as a pseudo-qualitative design, demonstrates the care that must be taken in designing a qualitative study. In a pseudo-qualitative study, the researcher acknowledges that the assumptions necessary for quantitative analysis are not present and determines not to conduct the final step of a quantitative study, namely a statistical test. However, he or she still attempts to quantify the qualitative data. For example, if a pseudo-qualitative design were applied

in this study, the researcher may suggest that time on task is a measure of motivation. He or she would then attempt to review video evidence to determine how long each student remains on task. The researcher may even calculate a mean time on task for each group and compare those. Figure 3 demonstrates what this design would look like.

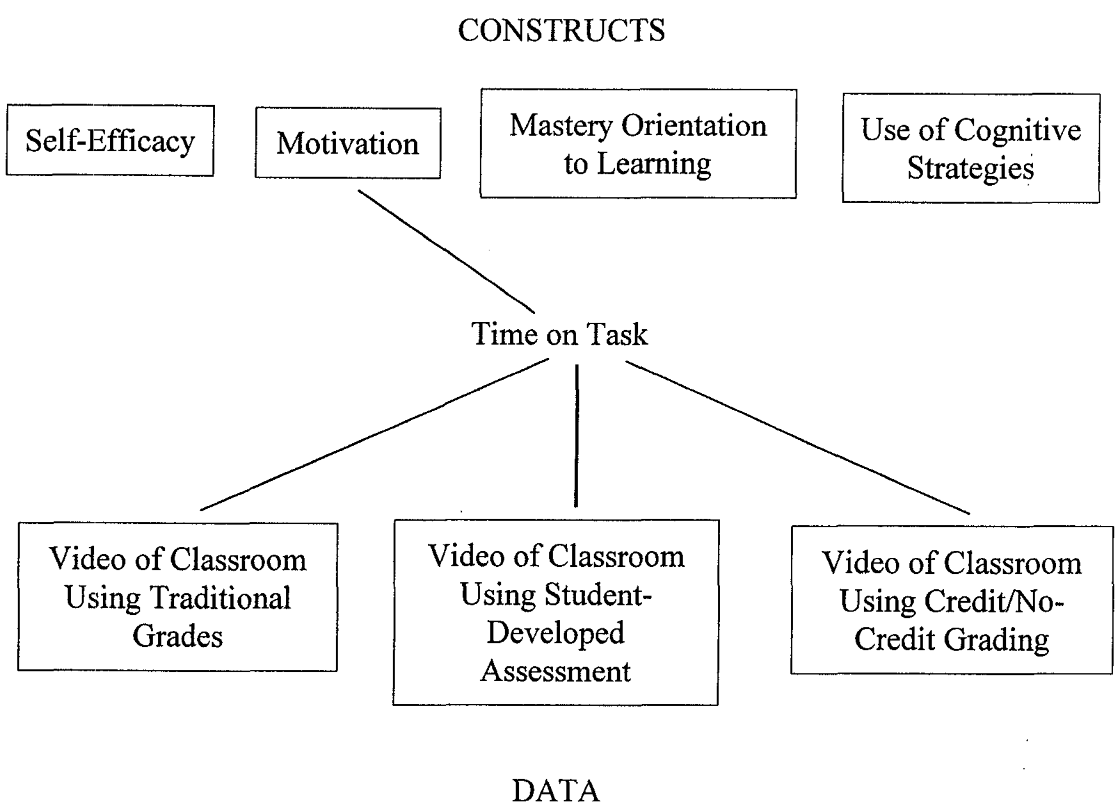

Figure 3. Pseudo-Qualitative Design

In figure 3, I have only demonstrated how the researcher's work would

progress from data collected from the videotapes to the construct of motivation (which is operationalized as time on task.) To complete the pseudo-qualitative design each of the four constructs would need to be operationalized in a similarly quantifiable manner.

There are a number of flaws with this design. First, the comparison of quantities lacks meaning without some discussion of the difference that may be expected due to chance alone (some statistical analysis). Second, the ability of a researcher to adequately identify a numerical representation of time on task from reviewing a video is questionable. Students may appear to be on task, yet they are actually working on a letter to their grandmother, not their math assignment. Likewise, other students may appear to be daydreaming, yet they might be pondering the complexities of the problem they are supposed to be solving. Other students may be talking to one another: Are they discussing mathematics or weekend activities? Or, are weekend activities maybe even connected to the mathematics lesson? It is likely very difficult to distinguish and hence be confident of the numerical value assigned in each of these cases. Yet, once a numerical value is assigned, it tends to carry tremendous weight at least in the minds of the readers if not the researchers themselves. For all of these reasons, great care has been taken to avoid the pitfalls of a pseudo-qualitative design when conceptualizing the data analysis in this study.

True qualitative research ought to make every attempt to grasp the complex

attributes or qualities of each variable in question as well as the complicated web of relationships that may be present. Miles and Huberman (1994) have written an extensive sourcebook detailing how qualitative data analysis methods can best be used in order to accomplish this task. They emphasize that an important component of qualitative analysis is to make sure the qualitative nature of the data is not lost. In other words, it is not sufficient simply to count how many times one student raises his or her hand in a classroom and then compare it to the number of times other students raise their hands. In fact, they write:

> We argue that although words may be more unwieldy than numbers, they render more meaning than numbers alone and should be hung on to throughout the data analysis. Converting words into numbers and then tossing away the words gets a researcher into all kinds of mischief. You thus are assuming that the chief property of words is that there are more of some than of others. Focusing solely on numbers shifts attention from substance to arithmetic, throwing out the whole notion of "qualities" or "essential characteristics." (p. 56)

A second key component of qualitative research is making sure there is a systematic way in which the data is being approached. In some senses, these two key components are in tension with one another. It would be relatively easy for a researcher to make significant mistakes in one area in the name of adhering to the

other. I believe the advice of Miles and Huberman regarding coding is extremely helpful in assuring both of the two key components are included in data analysis.

Miles and Huberman contrast two approaches to creating codes for data. In describing their stated preference, they advise the creation of a "start list." This list is generated from the conceptual framework of the study. While the list may grow in some areas and be reduced in others during the coding process, as the name implies, the researcher has a starting point to begin coding.

On the other end of the coding spectrum, Miles and Huberman reference the grounded theory approach of Glaser and Strauss. According to Glaser and Strauss, theory can be constructed from an analysis of qualitative research. They suggest that this inductive approach is in contrast to the idea that theory must be stated and then be deductively proven (as cited in Lewis, 1992). Researchers who take this approach to coding data would not want a predetermined "start list" of codes. In fact, researchers who espouse a grounded theory approach would argue that a "start list" is inappropriate as it is dependent on the researcher's conceptual framework as opposed to the actual data. The grounded theorist believes that coding should evolve out of the data. According to Miles and Huberman, this method of coding is "more open-minded and more context-sensitive" (p. 58).

For this study, I used a method that lies somewhere between the two. While I appreciate the relatively unlimited potential of the grounded theory approach, it is also

true that I had to some degree developed a "start list" through the four factor conceptualization of self-regulated learning. When reviewing the videotape, artifacts of student work, and tape-recorded interviews, I looked for the types of qualities that are shown in figure 4. It should be noted that, unlike the pseudo-qualitative design, this figure does not imply a one-to-one correspondence between the data and the four constructs that encompass self-regulated learning. Many of the qualities may well be evidence of multiple constructs. In addition, it is possible that some qualities may emerge from the data which were not anticipated or included in figure 4. This is both typical and appropriate in effective qualitative research.

In addition to qualities that may be indicators of self-regulated learning, I was also looking for indicators regarding the impact of the cognitive difficulty of the task as well as the time students had been exposed to the particular method of summative student evaluation. Again, the qualities shown in figure 4 were often tied to these two questions. For example, a student's persistence or types of questions asked was frequently found to be related to the cognitive difficulty of the task. The ability to look at the data in this truly qualitative manner was invaluable in identifying qualities that related to the three research questions proposed for this study.

Miles and Huberman (1994) also discuss three different types of codes. In the initial analysis of data, a researcher is likely to code either descriptively or interpretively. The third type of code identified by Miles and Huberman is a "pattern

code." They suggest that pattern codes are similar in concept to a statistical factor in that they group "disparate pieces into a more inclusive and meaningful whole" (p. 58).

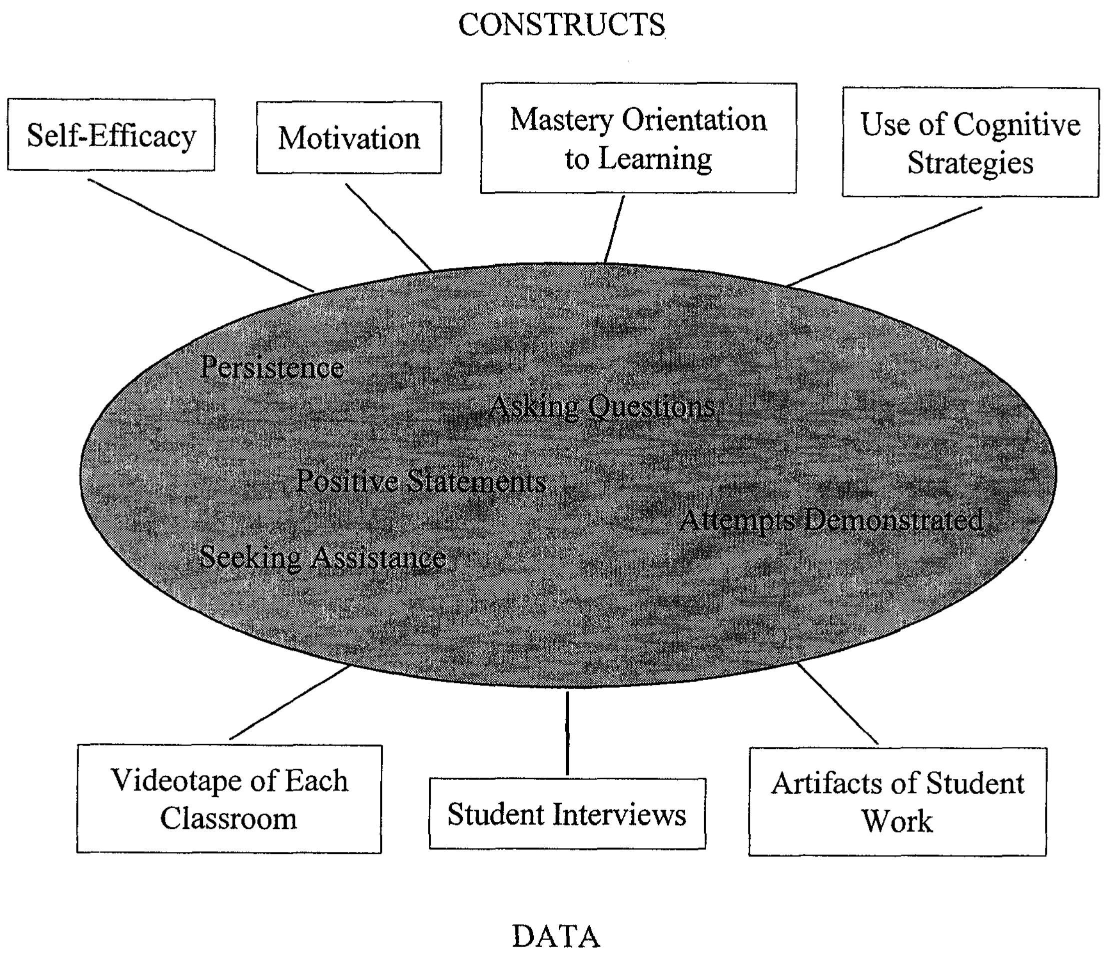

Figure 4. Qualitative Design

According to Miles and Huberman, pattern codes are typically assigned later in the analysis process. This is particularly true of the grounded theory approach. Data

is initially coded and then physically "cut up" so that it can be sorted into related groupings. These groupings are then assigned a pattern code.

For this project, the videotaped classrooms, the tape-recorded interviews and the pieces of student work that have been collected will be reviewed with the intent of identifying qualities like those suggested in figure 4. The qualities will be coded using a process of both descriptive and interpretive coding. Once this coding has been completed, the researcher will compare and sort the qualities of student behavior that have been identified in an attempt to further identify patterns of relationships between summative evaluations and the four factors of self-regulated learning. During this final stage of identifying patterns, the researcher will also be looking for connections between the cognitive difficulty of the tasks involved and the evidences of self-regulated learning as well as the length of time the students have been exposed to the particular method of summative evaluation and the evidences they display of self-regulated learning.

Once completed, the data analysis section will obviously not yield a conclusion that one group of student's level of performance was statistically significantly higher than the other two groups. It will, however, yield a thorough description of the qualities found in each group. The meaningfulness of this analysis will be bolstered in two ways. First, by the thoughtful and systematic design of the study. A design that carefully held many difficult to control variables (i.e., curriculum materials,

instructional strategies, formative evaluation strategies, school climate, classroom climate, teacher nuances, etc.) constant and varied only the type of summative student evaluation. Secondly, the meaningfulness of the analysis will be enhanced by the rich and full description of qualities found in each of the three groups of students.

In summary, the contribution to our current understanding of summative student evaluations and how they impact self-regulated learning should be significant. Consider the opinion of Charles Lewis regarding the grounded theory research methodology. Lewis writes, "While elements of the social world cannot be precisely mapped out, understandings of the interplay of experience, practice, and social structure can be developed, especially if historical research is used to provide context for the interpretive process" (p. 287). As has been outlined, in the current study, numerous variables have been held constant and only a few manipulated, and in both cases, the variables were chosen based on an extensive review of the historical research. The research design that has ultimately been developed for this study will ensure that the researcher is well positioned to develop "understandings of the interplay of experience, practice, and social structure" within high school mathematics classrooms regarding summative student evaluations and self-regulated learning.

CHAPTER 4

RESULTS

Traditionally Graded Classroom

As I have written earlier, most secondary school classrooms continue to assign either letter or numerical grades to students when completing summative student evaluations. I have also outlined a number of potential detriments to this traditional process of summative student evaluation. Thus, the first experimental classroom in this study, which I have labeled a "traditionally graded classroom," was somewhat analogous to a control group in an experimental study. In this classroom students were given regular quizzes and each quiz was graded and assigned a letter grade of A, B, C, D, or F.

The most notable finding regarding this classroom was the fact that the inclusion of graded quizzes had almost no impact on the existence of self-regulated learning. Across all four constructs, there were evidences of self-regulated learning; however, in each case, there was little to no evidence that the graded quizzes were impacting self-regulated learning. Given this fact, it was reasonable to look for other factors that may have contributed to the evidences of self-regulated learning. The most reasonable factor appears to be the constructivist nature of the instruction. This pedagogy was held constant across all three classrooms, and indeed, the evidences of

self-regulated learning were relatively constant as well.

A second notable finding from this classroom involved the very cautious manner in which students were demonstrating self-regulated learning. In almost all cases, it appeared as if there was a contradiction at play within each individual student. They could see evidences of self-regulated learning within themselves; however, this contradicted their beliefs regarding their abilities. Hence the evidences were demonstrated in subtle ways. This inner contradiction is likely a product of the limitation of the study to low-achieving students. It is certainly reasonable to assume the evidences may have been stronger with average or high-achieving students.

A final finding from this classroom, which is worth highlighting, involved the degree to which the constructs were dependent on each other as well as the degree to which these dependent constructs interacted with the cognitive difficulty of the tasks. What was clear in this traditionally graded classroom was that 1) the perception of cognitive difficulty of the tasks varied among students, 2) statements made during interviews as well as activities recorded on videotape were frequently evidences of multiple constructs, and 3) these dependent evidences interacted with cognitive difficulty in a complex manner.

Specific examples of the evidences of impact on self-regulated learning in this classroom are discussed in the following sections.

Self-Efficacy

For the traditionally graded classroom, as well as the other experimental classrooms, self-efficacy turned out to be the most difficult of the four constructs to gather evidence about. Only a small amount of data was found to be directly related to self-efficacy.

One evidence of self-efficacious behavior was seen in the analysis of the videotape of this classroom. During whole class activities and discussions students did, at times, volunteer responses. Certainly, some degree of self-efficacy is demonstrated by the relatively courageous act of orally responding to a math question in front of one's peers. One example of this was during the second lesson which covered measures of central tendency. The class had gathered data on the lengths of their fingers and summarized those data on the board. When the teacher asked for suggestions on how to summarize the data with a single statistic, one student volunteered "add them up and divide by how many numbers you have." A little later in the lesson, when the teacher was attempting to ascertain the level of knowledge the class had on the statistics of median and mode, another student responded. However, the response was given in a quiet, unsure, almost questioning voice. While the response itself indicated a degree of self-efficacy, the tone demonstrated the opposite. This tension was also seen in a discussion between the teacher and one of the students during the fourth lesson. The teacher complemented the student for helping another student with a particular algebra problem. The student responded, "But I didn't do it

like that." The teacher encouraged her by stating, "It doesn't matter how you do it. As long as you can do it." After the teacher had moved away from this group, the student was seen confidently explaining her methods to the other two students in her group.

Yet another example of this complex demonstration of self-efficacy was seen in the following response given by one student during the interview that was conducted after the completion of the activities. Speaking of the four activities that were taught by the researcher, the student said, "it has helped me try harder to do stuff and to ask more. Because usually I'm kinda embarrassed to ask, cause all the other kids know most of the math stuff, and they know it easily, and I'm a little late catching on, and it's kinda embarrassing to be the only one in the class that asks questions. [But in these activities] actually, I rose my hand and I asked for help. But it didn't seem too bad since I wasn't the only one doing that. I like the remediation program because we're going just a nice pace." This tension which was represented by behavior that was both self-efficacious as well as doubtful at the same time may be related to the fact that the study focused on low-achieving students. It would be reasonable to assume these students had a very limited amount of self-efficacy coming into the study.

The interview question that was designed to gather evidences about self-efficacy was "Thinking only about the lessons in which I (the researcher) worked with your class, have these activities made you any more or less confident in your ability to pass the math portion of the Graduation Qualifying Exam?" Responses from the four

students interviewed from this class indicated a general increase in self-efficacy, although not an overwhelming increase. The exchange between one student and the researcher went as follows:

> Student: I feel better.
>
> Researcher: You do, you think this has helped you to get more confident?
>
> Student: Um, umm.

Another student commented, "They made me more confident, because, it seemed like, at first when you started doing it, it seemed like I didn't know it, but then my brain just refreshed it and then it seemed easier so I just felt more confident."

This interview question also elicited a response related to the effect cognitive difficulty had on self-efficacy. The conversation between one student and the teacher went like this:

> Student: On some pages, like some of the things we did, made me more confident in that area. But, like the mean, median and mode, that didn't really. It kinda made me change my mind a little bit because it was harder for me to do than the rest of them.
>
> Researcher: So the more challenging ones kinda hurt your confidence.
>
> Student: Yeh.

Again, this type of response may be related to the restriction of the study to low-achieving students.

Another interesting and unanticipated finding from this student's response was

the fact that the student viewed the second activity on measures of central tendency as the most challenging. In the design of the study, the fourth activity which integrated two algebraic topics was created to gather evidence about the impact of changing the cognitive difficulty of the material. However, for this student, the second activity was the most challenging. Indeed, through analysis of other student responses it was found that a number of students believed the third activity to be the most challenging. The qualitative nature of both the data collection and the data analysis allowed for this observation. This discovery meant the researcher could not attribute responses that were unique to the fourth lesson to the cognitive difficulty of the lesson. It was necessary to dig a little deeper into the data to ascertain in which cases cognitive difficulty played an intervening role.

In summary, some degree of self-efficacy was demonstrated by the students in the traditionally graded classroom. Potentially as a result of the focus of the study on low-achieving students, these evidences were often mixed with behaviors or comments that reflected a lack of self-efficacy. Regardless of the complexities of the evidence on self-efficacy, it did not appear to be directly related to the fact that quizzes were given or that grades were assigned as a method of summative student evaluation. Indeed, there was no evidence that the assignment of grades had either a positive or negative impact on self-efficacy.

Motivation

As can be seen in the interview protocol, there were two questions asked of students that were specifically aimed at gathering evidence regarding the level of motivation they demonstrated during these activities. One of those questions asked in a very direct manner, "When you were trying to complete the activities in these lessons, why were you trying, what motivated you to try?" The other question attempted to get a measure of the motivational level of the student by investigating the value the student assigned to the task. Students were asked, "Did you find the activities you were involved in valuable, why or why not?"

Student responses to these two questions, as well as the classroom activities which were captured on videotape, seemed to indicate that they possessed at least a surface level of motivation. However, that motivation did not appear to be related to the quizzes that were given in the class or the grades that were assigned to those quizzes. For the most part, when reviewing student behaviors on the video, the students in this class demonstrated a fair amount of effort on each activity. They were not just sitting in class and doing nothing. One exception to this rule involved an instance where the teacher was helping one group, with his back to two other groups in the back of the room. During this time, one of the other two groups appeared to be socializing more than they were when they were being monitored more closely. The fact that a quiz was coming in a few minutes did not seem to have as strong of an impact on their efforts as teacher monitoring. This was also during the fourth and

most challenging lesson, hence, it may be that either over time, or as the difficulty of the activity had increased, the motivational level of the students began to deteriorate. However, in either case, the fact that a quiz was about to be taken did not appear to be a motivational factor.

As stated, the example given above was a rare example of students who appeared to have little motivation for completing the activity or lesson in this class. For the most part, students were actively engaged. However, when students who were interviewed were probed as to why they were trying hard, or why they placed value on the activities, their responses not only showed that quizzes and grades had no impact on their motivation, the responses indicated that students were sometimes not sure themselves as to why they valued the activities. As was the case with evidences regarding self-efficacy, the students seemingly needed to convince themselves that they really were motivated. It was almost as if they were struggling to accept the fact that they were motivated. For example, when one girl was asked whether or not she thought the activities were worth doing she responded, "You gotta learn how to do those stuff." When probed further she simply said, "Cause, (pause) you just have to. I guess." Another girl responded with yes, "Because I may use it sometime more in life." When she was asked to elaborate, she responded firmly with "Because you have a lot of math in your life and it's just something you need to know."

In some cases students were able to identify more specific reasons for either valuing the activities or simply trying hard on them. Although, even in these cases,

students still did not mention the quizzes or grades as a reason. Instead, the reasons identified were outside the realm of this specific class. Furthermore, these students continued to demonstrate a lack of conviction or clear commitment to these motivating factors. For example, students identified the need to pass the state Graduation Qualifying Exam (ISTEP+/GQE) or to be prepared for future schooling as motivational factors. In doing so, one student stated, "I need to learn the things so I can pass my test [the ISTEP+/GQE] . . . [I need to] learn how to do it so I can get the test done and over with." She continued with, "Maybe you have to do that in higher math skills, you have to know how to do that to finish another problem. Or, it is connected in some way and it would help out or something." Another student said, "Because I'm going to have to use that on the ISTEP, so yes, it is useful." A third student stated "I wanna go to COLLEGE, and like all that stuff, I just don't wanna be like a nobody and like just get a regular job, I wanna DO something. So I was trying to like make sure I can do better, and remember what I'm doing on that. So I just wasn't trying to go do it sloppy . . . I was just trying to do it right, and not just try to hurry up and get it done."

Many of the responses cited as motivators for these students showed a dependency between the constructs of motivational factors and a mastery orientation to learning. However, even when students talked about being motivated simply to understand the material, these students still demonstrated a degree of caution in their response as well as a lack of clarity regarding why they wanted to learn the material.

For example, one of the students who had initially said he was motivated to learn the material for the ISTEP+/GQE continued with "Cause I guess I probably will need it later on in life . . . I just think of it as more math for me to work at and to get better at, I wouldn't know whether I would need it later on."

Relative to the questions of how the effects on motivation changed over time or across cognitive difficulty of the task, little evidence was found. There was no evidence to suggest a change in motivational level across time. Regarding cognitive difficulty of the task, the study was designed to look at changes during the fourth task, which was theorized to require a higher level of cognitive processing than the first three tasks. While a change in motivational level was not apparent from this perspective, the following comment made by one student during the interview process suggests there may be some relationship between cognitive difficulty of tasks and motivational level of the students. In response to a question regarding why the student was putting forth some effort in the class, he commented, "I just figured I needed to try something, . . . I really didn't have that much to go on with math, and it was kinda hard . . . everyone else can easily understand it better than I can . . . and me, . . . I don't even know what I'm doing." This response may indicate that an increase in cognitive difficulty will spur motivation. However, the student's response is far from convincing. Again, this type of response may be a factor of the selection of low-achieving students for the study. Regarding ties to the grades assigned in this classroom, the motivational factors that were seen do not appear to be related to or

impacted by the method of summative student evaluation which was used in this traditionally graded classroom.

Mastery Orientation to Learning

The question that was posed to students in the interview regarding this construct was, "During each of the lessons, did you feel you were concerned with really understanding the problems or were you only trying to get the activity done and over with?" This question elicited responses that demonstrated complex relationships between the cognitive difficulty of the task and the construct of mastery orientation to learning, as well as a complex relationship between a mastery orientation to learning and the prior construct of motivation.

For example, one student responded "I really wanted to understand because I knew that ISTEP [the Graduation Qualifying Exam] was coming up and I was trying to like get it into my mind so that I could remember, ... I wanna just pass that and get it over with, so I was really trying to remember what this had to do with." Although the question was aimed at the student's orientation to learning (mastery versus product), the student's response demonstrated the connection between that and motivation. Clearly, the student was motivated. However, as motivation was conceptualized for this study, the researcher was looking for either intrinsic motivation or motivation that was based on the value of the task. The response of this particular student clearly did not demonstrate an intrinsic motivation. It is less clear, however,

this student's response likely did not demonstrate a valuing of the task either. To the degree the student did believe the task had value, the value was directed to passing the Graduation Qualifying Exam, not the task itself. In other words, the motivation that was demonstrated by this student was not the motivation that was conceptualized in this study to be an indicator of self-regulated learning.

Nevertheless, it is important to note that the task was not valued for the grade assigned in this class. It was valued for potential success on the Graduation Qualifying Exam. This is similar to students working hard for an opportunity to go to college, or get into a particular college, or land a particular job, etc. While the motivation is neither intrinsic nor based on task value, it is based on something outside of the specific learning activity. Therefore, the argument may be made that this was indeed an indicator of self-regulated learning in spite of the fact that it did not fit neatly into the four hypothesized constructs for this study.

The analysis of the orientation to learning of this student was similar to the analysis of the motivation. Again, the student, at some level, was demonstrating a desire to master the material. He was not simply interested in completing a product. This was indeed an indicator of self-regulated learning. However, it was not a result of the grades that were being assigned in the class.

The response of another student indicated some of the same feelings, as well as a connectedness to the cognitive difficulty of the task. Initially, when questioned about the first lesson, the student responded with, "I really wanted to understand,

because, I just never really worked with anything like that before and it was kinda fun to know a new thing about it. So if I ever have it in the future I know how to do it and I can get it done." As with the previous student, the degree to which this student approached the task with an orientation toward mastery of the material was linked to the degree to which the student perceived a future need for the material.

When this last student commented, "I just never really worked with anything like that before," it appeared as if the mastery orientation the student was showing may be impacted by the cognitive difficulty of the task. As the researcher repeated the interview question for each of the specific lessons this student had worked on, the fact that the cognitive difficulty of the task was playing a role in the mastery versus product orientation became more apparent. Recall the interview question was asking the student if he was "just trying to get the problem done" or if he "really wanted to understand how to do the problem." When this question was posed to this student regarding the first worksheet, he responded in part with "I just never worked with anything like that before and it was kinda fun to know a new thing about it." When the same question was posed to the student regarding the worksheet in the second lesson he responded with, "I was trying to get that done, it was a little harder than the other ones and it's just a lot more to think about, but, if I have to learn it I will, but it wouldn't be something I choose to learn." Unlike the first response which suggested that increased cognitive difficulty may lead to an increased level of mastery orientation to learning, this response suggested the opposite.

When the interview question was posed to the student regarding the worksheets from the third and fourth lessons respectively his responses were, "The top part, I didn't really have problems on, so that was something I just wanted to get done, but then the bottom part, the graph, it was sorta harder . . . so I wanted to concentrate on that and get that right" and, "it was sorta complicated, but . . . I wanted to learn how to do it, but it was just another one I just wanted to get done." These responses, when considered with the responses to the first two worksheets, suggest the difficulty of the task did play a role in the student's orientation to learning the task, however, that role is not simplistic. It appears as if a task that is more challenging is on one hand a stimulus for a mastery orientation to the task. Yet, on the other hand, his closing comment that "I wanted to learn how to do it, but it was just another one I just wanted to get done" suggests that more challenging tasks may also stifle the student's attempts toward achieving a deep understanding of the task.

This seemingly conflicting evidence may actually be consistent in light of the theory Vygotsky (1978) suggested regarding the "zone of proximal development." Vygotsky theorized that students' capabilities for development were evidenced by activities which they could not perform on their own, however, when given assistance, they could perform the activity. In other words, if an activity was just slightly above a student's actual developmental level, it may well be within his or her zone of proximal development. This would explain the cases in which the increases in cognitive difficulty led to an increase in mastery orientation. If, on the other hand, the increase

in cognitive difficulty was significantly beyond the student's actual developmental level, the activity may now fall outside of the student's zone of proximal development. This would then explain the cases in which the increase in cognitive difficulty led to a decrease in mastery orientation.

Given the low-achievement history of the students involved in this study, the difference between a difficult task having a positive impact on the students or having a negative impact is likely a very fine distinction. One piece of supporting evidence for this viewpoint is the fact that this student found the second worksheet to be "a little harder than the other ones." This second worksheet was not conceptualized to be more challenging than the others. It was the fourth worksheet that was conceptualized to be at a higher level of cognitive difficulty. However, for this low achieving student, the second worksheet was the most challenging. For him, it may have been at this point that the fine line had been crossed and the increased cognitive difficulty was now functioning as an impediment to the student's quest for a mastery orientation to learning.

Further evidences that the cognitive difficulty of the task may ultimately hinder a student's mastery orientation to the task are seen in the simple responses of yet another student to the same interview questions. When asked about the worksheet given over the first lesson, this student responded, "I wanted to understand it." When asked the same question regarding the fourth lesson (which was conceptualized to be at a higher level of cognitive difficulty) the student responded that she was trying to

"just get it done and over with." When the researcher probed into the reasoning for this, the student confirmed that it was due to the increased difficulty of the fourth task.

For the purposes of this study, it is important to note that the assignment of grades to these students did not seem to play a role in the student's orientation to the task. Yes, the students frequently demonstrated a mastery orientation to the task, and, that orientation was impacted by the cognitive difficulty of the task. However, the students gave no indication that this was either stimulated or hindered by the assignment of grades within the class.

The findings from an analysis of the videotapes of the classroom activities seem to corroborate the interview responses of these students. For the most part, there are evidences of a mastery orientation to learning. Some examples of this include, after the teacher has gone over a particular problem with one student in a small group, another student was seen copying the information over onto their worksheet (recall the worksheets were not graded, only the quizzes). In another group, after the teacher had shown the group how to complete a particular problem, one group member continued to probe for understanding by asking "So, you just put the $10 here and then the price of the three videos here plus the $1.35 for the drinks and then the change here?" At yet another time, when the teacher had begun to explain a problem to one member of a group using that student's worksheet, the student across from her made a significant effort to lean forward to see and follow along. Similarly, in yet another situation, one girl actually got up from her group and moved to other groups

checking her responses against theirs. As one final example of a mastery orientation to learning, after the teacher had been working with one particular student and then left that area, the student was seen explaining her methods to the other two students in the group.

As was suggested by the responses students gave during the interviews, there is some evidence on the video that the cognitive difficulty of the tasks may have impacted the students' orientation to learning. During the fourth and most challenging lesson, while the teacher was helping one group, with his back to two other groups in the rear of the class, those other two groups appeared to be socializing more than they were when they were being monitored. The fact that a quiz was coming in a few minutes did not seem to impact their efforts. By contrast, when the teacher was assisting them, or even just monitoring them, they seemed to be more interested in learning the material. It could be argued that this was an indication of a change over time. In other words, in may be argued that the students' mastery orientation to learning decreased over time. However, there was no evidence that this was gradually diminishing over the first three lessons. In addition, there was no evidence from the interview data that students' orientation to learning was changing over time. Hence, it appeared more likely that this increase in socialization and decrease in attention to learning was related to the cognitive difficulty of the fourth lesson.

Unlike the interviews, there was a small amount of evidence that the mastery orientation students demonstrated in this classroom (or the lack of it) may have been

related to the grades they were assigned. During quizzes, some students attempted to ask questions regarding how to do particular problems. The teacher informed those students that he could not assist them during the quiz. However, he encouraged them to ask those questions the next day when he returned the quiz and went over it with them. However, after the quiz was completed, neither of the two students who had attempted to ask questions during the quiz went back to their notes and tried to figure out how to do those problems. In fact, both simply put their heads down and rested. In addition, when the quiz was returned on the following day, the teacher encouraged students to ask questions. He reminded them that some of them had questions during the quiz. However, no one asked questions at this point. These actions suggested that the students may have been more interested in getting a good grade on the quiz, than they were in actually mastering the material.

In summary, the interviews and videotape taken together suggested that in many instances the students in the traditionally graded classroom were demonstrating a mastery orientation to learning. That orientation may well have been impacted by the cognitive difficulty of the task. Tasks which were somewhat challenging may have increased the student's desire to fully understand the task. However, if the task was too difficult, the student's desire to understand tended to disappear completely. There was no credible evidence that these relationships changed over time. In addition, there was only limited evidence that these attitudes were related to the assignment of grades in the classroom.

Use of Cognitive Strategies

Evidences either for or against the use of cognitive strategies were difficult to spot in the videotape of the classroom. The one evidence that was apparent, was a very consistent pattern of requesting assistance, primarily from the teacher, but also from other students. Specific instances of these requests were cited in the previous section as evidences of a mastery orientation to learning. Clearly, these two constructs, namely a mastery orientation to learning and the use of cognitive strategies, are not independent of one another.

More specific examples of the students' use of cognitive strategies were seen in their responses to the interview question that was directed to this topic. The question posed to students was, "How did you go about thinking through the problems, or trying to figure them out?" To assist students in responding to this question, copies of the worksheets were used as prompts. In most cases, the researcher had a copy of the student's actual responses. This served as an excellent prompt for discussing the cognitive strategies that were employed by students while they completed the worksheet.

For this group of students, the most evident use of cognitive strategies clearly fell in the category Zimmerman and Pons (1986) defined as the seeking of social assistance. When the interviewees were probed regarding the assistance they received, it was clear they had sought out the assistance from both the teacher as well as other students. Consider the following quotes and short extracts from some of the

interviews:

- "Then we got stuck, and we called the teacher over and she helped us out. She told us a way and we tried it and that's what we got."
- "To get the final answer [the teacher] came over and helped us, because neither one of us could get it."
- "But I didn't know how to figure out what to use, so [the teacher] came over and helped me."
- Researcher: What makes you so sure you got these right?

  Student: Because you told me.

  Researcher: Did you call me over to look at them?

  Student: Yes.

As is inferred in the above quotes, the move to seek assistance from the teacher was directly related to the cognitive difficulty of the task. When the task difficulty was such that other cognitive strategies failed, assistance was sought. One student summarized it this way, "I called [the teacher] over. We couldn't figure it out, we forgot to use a step for number eight. . . . because we always had to get side c, but it switched over when I had to get side b. So, that was what really messed me up that time."

Students did not only seek assistance from the teacher. Many sought the help of their peers. In the following extract, we see an example where students worked first within a group to which they were assigned, but then expanded their seeking of

assistance to other students in the class:

Researcher: Then you did the graph together. Did you have any problems? It looks like there was a lot of erasing.

Student: We did a bar graph there first.

Researcher: How did you figure out that you shouldn't do a bar graph?

Student: I went to ask another group, and then I was like, I knew we had to do it this way, and I told them, and we changed it.

Again, the cognitive strategy of seeking assistance was more likely to be employed with tasks that had a higher level of difficulty. Consider the following extract:

Researcher: Were you working on this by yourself?

Student: No, I was working with a friend. On this one we were working together.

Researcher: Why?

Student: This one, I guess it was more complicated. You had more to think about, more to find out, more questions and stuff.

Researcher: The person you were working with, was that any benefit to you?

Student: Yeh, because we knew it was right if we got the same answer or if it was wrong if we got two different answers or something.

Another student flatly stated, "I didn't know how to do these. I just worked with other kids that were in my group, and that's how I figured out that one." This student

also showed a link between self-efficacy and the use of cognitive strategies when he continued with, "I have kinda like a real problem with math, . . . it's kinda difficult to me than it would be to other people, cause I don't catch on as fast as what they do." Nevertheless, although her self-efficacy was low, this did not stop her from employing the cognitive strategy of seeking assistance from her peers.

A third cognitive strategy that might also fall within the broad category of seeking assistance was students simply listening to the teacher and one another at appropriate times. For example, one student commented, "When you [the teacher] did it on the board, I got a different answer than you so I just corrected it." When another student was describing the method he used to solve a particular problem he stated, "I think you [the teacher] were just talking aloud about it, you were telling everybody about it [the method for getting past a particularly challenging part of one problem]." Later in the interview when this same student was explaining how he calculated the median on the second worksheet he said, "I thought that was the middle number, but I didn't know you were supposed to like count down . . . you [the teacher] told me. I think you were doing an example on the board."

In addition to these three types of seeking assistance, other examples of the use of cognitive strategies were evident. When asked how she figured out that a particular problem was missing $20, one girl responded, "I guess we just kept re-reading it, cause we didn't understand. So we just put that [the wrong answer] and then I guess we re-read it again and figured out that we had to add the $20 for gas." It was

interesting to note that this student was both seeking assistance (she refers to "we" because she is working in a small group) and was also employing the strategy of re-reading the problem.

Two other students offered evidence that they had at least mentally checked their answers for reasonableness. When asked why she had erased a particular answer on her worksheet, she replied, "Cause that's supposed to be four something and that was all the way up to eight." When describing how he calculated a mean, another student stated, "The first time I did it I think I added too many 30,000's and I ended up getting a bigger number than what I was supposed to. I ended up with like 500,000 or something. I ended up with a really big number and I think I added too many 30,000's together. Then I asked some of my classmates and they did it right and they told me I didn't do something right. So I went back and I did it again and I got the right number." Like some of his peers who were cited earlier, this latter student was exercising one cognitive strategy, namely checking one's answer for reasonableness, within the context of another cognitive strategy, namely seeking assistance.

In yet another example of the use of cognitive strategies, there was a combination of "listening/seeking assistance" and "checking for reasonableness." This can be seen in the following exchange between the researcher and a particular student:

Researcher: How did you know you had some wrong?

Student: They just, I guess they just didn't look right, [pause] or no, when we were doing this the numbers came out wrong or

something.

Researcher: Let's talk about the graph. You say it didn't look right. Is this one that is erased the one that didn't look right?

Student: Yeh.

Researcher: How did you know that it wasn't right?

Student: Cause you [the teacher] said it was supposed to be a straight line and it wasn't.

As I have stated, this particular student had employed the strategy of listening to the teacher (a version of seeking assistance) and was also employing the strategy of checking her response for reasonableness.

There were other examples of students using multiple cognitive strategies. Consider the following statement made by a student who was explaining why he chose not to use the method the teacher had explained on the board. He said, "I tried to figure out my OWN way to do it and I think I kinda messed up on it, because, I got kinda stuck after I got here. When I asked for help from one of my classmates, I couldn't really figure out what they were trying to say on it. That's why I had to wait for when you [the teacher] could help us out." In this one statement we see a student who listened to the teacher explain a method on the board, attempted to develop his own method, asked other students for help, and then asked the teacher for additional assistance. It is also interesting to note that this student's responses were not only an evidence of the use of multiple cognitive strategies, but they were also evidence of a

mastery orientation to learning. The student sincerely wanted to understand the activity in depth, not just get it done. As has been discussed, this was yet another evidence of the dependent nature of the constructs that were being used to operationalize self-regulated learning.

Another example of the use of multiple cognitive strategies was seen in the following interchange between the researcher and a particular student who was explaining her strategies for completing the fourth worksheet.

Researcher: Were you just kinda making some guesses?

Student: Not really. First thing I just reasoned it out. I was trying to find something that multiplied into seven that is like in the range of 30's. I just thought of five and I knew that was 35 plus two equals 37.

Researcher: Now it looks like on number 2 you were doing something different, why?

Student: Yeh, because I couldn't find . . . what would equal 23. So I just tried a different step that I remembered like subtracting from each side and then bringing it down there and dividing.

This student was first employing a strategy of "guess and test." In fact, the student suggested that he was not guessing randomly, but rather, had thought through what a reasonable guess would be. As the cognitive difficulty increased to the point that he could no longer successfully use the guess and test method, the student attempted to

recall algebraic methods for solving the problem.

In summary, a number of evidences were found where students were employing the use of cognitive strategies. Increases in the cognitive difficulty of the task (as perceived by the participants rather than conceptualized by the researcher) appeared to stimulate the use of cognitive strategies. There were no evidences that the use of these strategies changed over time. In addition, there were no evidences that the use of these strategies was either encouraged or discouraged by the grades assigned to the quizzes.

Credit/No-credit Classroom

The first experimental alternative to the traditionally graded classroom was a classroom in which summative student evaluations were assigned only credit or no-credit. It was hypothesized that this reduction of emphasis on letter grades would also reduce negative impacts on self-regulated learning. In this classroom, students were given regular quizzes and each quiz was graded and then assigned a score of either credit, or no-credit.

In examining the data collected from this classroom two observations stand above the rest. The first, and perhaps most important, is the fact that a move away from traditional grading did not appear to have a detrimental effect on any of the four constructs that encompass self-regulated learning. In fact, when examining the construct of self-efficacy, there was evidence that the move away from traditional

grading may have increased the self-efficacy of students. This suggests that at the very least, there is nothing to be lost from a move away from traditional grading and potentially something to be gained.

The second notable finding from this classroom is related to the fact that significant evidences of self-regulated learning were found in both the traditionally graded classroom as well as this credit/no-credit classroom. Hence, it was reasonable to look for similarities between these two classrooms that may have contributed to, or at least allowed for, these evidences to be displayed. As was the case in the traditionally graded classroom, there was substantial evidence in the credit/no-credit classroom that the constructivist pedagogy employed served as a catalyst for self-regulated learning. Specific examples of this and other evidences of impact on self-regulated learning in the credit/no-credit classroom are discussed in the following sections.

Self-Efficacy

The design of the study incorporated a reward in both the traditionally graded classroom as well as the credit/no-credit classroom. At one level, the reward was simply the grade that was assigned. However, because those grades were less rewarding in a summer school program than they would have been in a regular school program (where report cards are issued and grade point averages calculated), an additional reward corresponding to each grade was given. In the traditionally graded

classroom, for each "A" a student earned his or her name was placed in a box four times for a raffle to be held at the end of the summer school session. For each "B" the student's name was entered three times. For each "C" the student's name was entered twice and for each "D" his or her name was placed into the box one time. In spite of this extra assurance that grades were rewarding, there was absolutely no evidence in the traditionally graded classroom that these rewards impacted self-regulated learning.

In the credit/no-credit classroom, the grades and hence the raffle were simplified. If a student earned "Credit" on a quiz, his or her name was placed in the box for the raffle. This move to what students may have perceived as a more attainable goal may have had a positive impact on self-efficacy. As evidence of that, consider the following description of events from the videotape of the credit/no-credit classroom.

While the teacher was passing back the 2nd quiz, a student looked at hers and said in a somewhat surprised voice, "Oh, I passed," then after a brief pause she continued, "So my name goes in the box?" After the teacher responded affirmatively, another student excitedly said, "My name goes in the box!" and she slapped her desk. She continued in a firm voice saying, "My name up in the box this time, I got it RIGHT!"

It is unlikely that these low-achieving students had experienced much positive reinforcement regarding their academic performance in the past. Their reactions to passing this particular quiz certainly seem to indicate that. Perhaps the use of a

credit/no-credit grading approach allowed for more students to believe they were successful and thus enhanced their mathematical self-efficacy.

It is important to note here that while the credit/no-credit approach seemed to elicit some positive evidence regarding self-efficacy (as compared to the traditionally graded classroom), this self-efficacious behavior was also tied to the student's orientation to learning, and, the evidence on this construct is of a negative nature. In other words, the comments students were making were, at the same time, both positive evidences of self-efficacy and evidence of a product orientation toward learning (as opposed to a process orientation). In the videotape source cited above, the students continued to ask a variety of questions. The majority of the questions they were now asking seemed to focus on their score on the quiz. Some examples were "So is this right or not?" and "It was worth two points right?"and "If the answer was 8.5, isn't 7 OK?" These questions, along with their earlier excitement regarding receiving credit on the quiz, did indeed indicate a level of self-efficacy, however, they were also product oriented as opposed to process oriented. This was one more example of the complex dependent relationship among the constructs of self-regulated learning.

An additional factor that appeared to be playing a role regarding self-efficacy was the instructional techniques used by the teacher. These techniques were not a variable under review in this study. In fact, each classroom was taught by the same instructor using a constructivist approach to mathematics instruction. There was

evidence, however, that this approach may have been a factor in promoting self-efficacy.

In each classroom, students were given the opportunity to "discover the mathematical ideas" through directed activities. They were able to discuss these activities with their peers as well as the teacher. In addition, there were opportunities to discuss the activities as a whole class. In the credit/no-credit classroom, examples were seen of students volunteering to show their work on the chalkboard. In addition, while one student was working alone on one of the activities he could be heard saying, "I didn't do it right." The teacher responded with "Yours looks perfect. That looks perfect." To this the student responded, "Oh, thank-you." Both the student who volunteered to show his work on the board as well as the student who was ultimately encouraged by his efforts at developing a solution on his own, were demonstrations of self-efficacy. However, it is possible that these demonstrations were tied more to the constant of a constructivist approach to mathematics than they were to the variable of a credit/no-credit approach to summative evaluation. As one additional evidence to this point, consider the following conversation between the researcher and one student during the interview conducted with the student after the lessons were completed:

Researcher: Thinking only about the lessons I worked with you on, have these made you any more confident in your ability to pass the math portion of the Graduation Qualifying Exam?

Student: Yeh, a little bit. Cause last year I wasn't too good at math. My

teacher, she never helped me out.

Researcher: So, when I was working with you, you got more help?

Student: Yes.

Researcher: How about from other students?

Student: [In my class last year] she wouldn't let nobody talk.

Researcher: So you preferred this way when you could get help from other students.

Student: Yes.

Researcher: And you think because of that you're a little more confident now?

Student: Yeh.

As was the case in evaluating self-regulated learning patterns in the traditionally graded classroom, there was again evidence that the cognitive difficulty of the task played a role. When one student was asked whether her confidence in her ability to do the mathematics had improved, she commented, "In some of it I think so. Because some of it was easy, and it was like easy to remember, but some of it, it was like hard." Clearly, at least in this low-achieving student, an increase in the cognitive difficulty of the task had the potential of negatively impacting self-efficacy. It was, however, interesting to note that the student did not refer to the credit or no-credit grades which were assigned to her. Indeed, it appeared as if the cognitive difficulty of the task played the role it did in self-efficacy independently of the summative

evaluation that was used in the classroom.

Motivation

For this study two factors were conceptualized as indicators of motivation. They were the value students associated with a task and the level of intrinsic motivation they demonstrated. Certainly, like the four constructs that help us understand self-regulated learning, these two sub-constructs that help us view motivation were likely dependent variables. In addition to this fact that they did not stand as clear and distinct entities from one another, they were in and of themselves not always clearly identified. Nevertheless, there did appear to be some evidences of motivation in the students in this credit/no-credit classroom.

An early example of this was seen on the video at the very beginning of the first class. One student asked, "Can I take notes? So I can learn?" Later, during the student interviews two different students were asked the question, "What motivated you to try to learn the mathematics that was being taught? Why did you try?" The first responded in the following manner:

Student: Because, if I don't learn how to do it, then later on, the person that I copy off, they won't be there, so then I won't know how to do it. When that problem comes up again, in like a test or something, I won't know how to do it and then I'll get it wrong, so that's why I want to learn how to do it.

Researcher: You mean like the quizzes that I gave?

Student: Yeh, or another test at school.

Researcher: Let's assume you pass the Graduation Qualifying Exam, will you take any more math classes?

Student: Yeh.

Researcher: How come?

Student: Just to keep refreshing my memory up, until I graduate.

Researcher: Why do you want to keep refreshing your memory?

Student: Because you can forget the stuff sometimes. Like I just forgot that, that fast, and we haven't even been out of school that long. . . .

Researcher: Why do you want to know it?

Student: Because it's good to learn math. You know, even though you pass, it's not just to forget it, you just keep learning more stuff, even though you pass the test, you just wanna keep learning more and more.

Researcher: You just like learning things?

Student: If I can, and I understand it, yeh.

While this student began by referring to tests as a motivational factor, as the discussion progressed, it became clear she had at least some level of intrinsic motivation. Similarly, a second student responded to the same question:

Student: Because I want to pass that GQE test mostly. Learn more of it.

Researcher: Do you think if you passed the GQE, and you were sitting in a math class and this stuff came up, would you still do it?

Student: Yeh, I'd probably still do it.

Researcher: How come?

Student: Because I wanna know more.

Clearly, these students' desire to "know more" was mixed with a desire to be successful on the Graduation Qualifying Exam. Each of the other two students who were interviewed from this class also identified the Graduation Qualifying Exam as a motivator. This was indeed an indication of the students believe that the task had value. Hence, these were evidences of motivation. These particular motivational evidences were also tied to a mastery orientation to learning. This gave additional strength to the argument that they were indicators of self-regulated learning.

Although many of the evidences of self-regulated learning in general and motivation specifically were mixtures of two different constructs, the following exchange between the researcher and one of the students from this class was an example of a student who was at best struggling to see any value in the tasks, however, she did demonstrate a level of intrinsic motivation:

Researcher: Do you think this stuff was valuable? The Pythagorean Theorem, do you think that was worth learning?

Student: Not really, I don't see what it's gotta do with (pause) I don't

know.

Researcher: How about the mean, median, mode?

Student: Not really, it is, but it's not, I don't know why.

Researcher: What makes you think it might be?

Student: I guess cause that worksheet has to do with money.

Researcher: How about the third one on tables and graphs? Is that of any value?

Student: They all value, but some of them, it's like not worth doing. You just gotta do them.

Researcher: What do you mean you gotta do it? Why did you gotta do it?

Student: You ain't GOT to.

Researcher: You could have sat there and did nothing all day.

Student: I didn't want to do that. I wanted to learn it, but it's like, I'll probably not even use it when I get older. But, I MIGHT, so why shouldn't I just go ahead and learn it?

As in earlier cases, there was at least one evidence that the motivation demonstrated in this class may have been tied to the constructivist methodology employed by the teacher. During the lesson on the Pythagorean Theorem, students were given a right triangle with legs of lengths six and eight. The task was to measure the length of the hypotenuse and then use this triangle and others to discover a pattern (namely the Pythagorean Theorem). Students were also given the option of creating

their own triangles. In the video of this classroom, two different students could be seen creating their own triangles and then choosing to share those with the class as a whole when the teacher went over the problems on the board. While this was an indicator of motivation, and it was encouraging to note the summative evaluation in the class did not stifle this motivation, it also appeared as if the motivation was tied closely to the experimental constant of instructional methodology.

Also, as in earlier cases, there was some evidence that this construct was impacted by the cognitive difficulty of the task. When a particular student was asked if he would still be motivated to study math even after he passed the Graduation Qualifying Exam, he responded, "I'd still work on it . . . cause I'd know a little more about it so it would be a little easier." By omission, this statement implied that an increase in cognitive difficulty may stifle motivation. Clearly this was far from a definitive conclusion, however, it did suggest a tie may exist to cognitive difficulty.

In summary, there were evidences of student motivation in this classroom. There was little evidence that the credit/no-credit approach to summative evaluation that was used in this class was either a catalyst or a roadblock to this motivation. Increases in the cognitive difficulty of the tasks may be detrimental to motivation. Finally, no evidences were seen that the motivational level of students changed over time in this class.

Mastery Orientation to Learning

As discussed in the section on motivation, evidences of motivation and a mastery orientation to learning were often linked. However, in addition to the evidences already discussed in earlier sections, a variety of evidences pointing more specifically to a mastery orientation to learning were seen in this classroom.

One example was seen in the videotape of the classroom during the third lesson. After the teacher had returned the quiz from the previous lesson to all students, one of the students remembered a question she had asked during the quiz and then asked if the teacher would explain it. The quiz was at this point obviously already graded. There would be no final exam, hence, the student was clearly interested in understanding or mastering the topic.

A second example could be seen in the videotape of the first lesson. When students were given their first set of problems in a worksheet format, the teacher told them that they may work alone or with one other person. He also stated that he had already worked out the first problem and therefore, if anyone was unable to figure it out, he would give them a copy of the solution complete with work shown. However, no students requested this. All students began working on the problem. Some worked alone, some asked for help either from one another or the teacher. However, none requested to be simply given the solution. A student with a purely product orientation to learning would have simply asked for a copy of the solution to the problem so that the product would be complete.

Yet another example of a mastery orientation to learning was seen in the interaction between students and the teacher during the time they continued to work on the first worksheet. A number of the students were asking the teacher for assistance. Indeed, some students could be seen expressing frustration that the teacher was unable to get to them quickly enough. Given the fact that the worksheet itself was not graded (only the quizzes were), this seemed to be an indication of the desire of these students to master the material.

Clearly a question of interest in this section, as well as in each of the other sections, was the degree to which the credit/no-credit approach to summative evaluation may have contributed to the evidences of a mastery orientation to learning. An interesting interchange was seen in the videotape during the class discussion of a previous quiz. First, the teacher passed back the quizzes and encouraged the students to ask questions about the problems they got wrong. The first response from a student was actually directed to the class when the student looked around and said, "So who got the best grade?" Although the actual grade was only credit or no-credit, this student was apparently interested in the number of points that some other students may have received which then led to their grade of credit. Indeed, in the review of literature it was pointed out that some argue the competitive nature of grades make them a motivator while others argue they stifle true motivation.

After this first student had spoken, a couple of other students asked to go over some of the problems on the quiz. One student said, "I want to see what I did

wrong." Another responded to that with "Me too." The teacher did indeed go over one of the problems. While most students appeared to be paying attention, few were seen taking notes. In addition, a couple of students questioned whether their quiz may have been graded incorrectly. When the correct answer was shown to be "a," one student said "I had 'a' here." The teacher responded by stating, "But over here you had 'b'. I wasn't sure which you wanted as your answer."

While the initial desire to go over problems on the quiz seemed to indicate a mastery orientation to learning which was positively impacted by the fact that a quiz had been given, a deeper listening to students suggested otherwise. It appeared by the lack of significant note-taking and the concerns over who got the best grade and whether their grade should have been higher, that the students may well have been more interested in the competitive game than they were in actually understanding the material.

As in earlier sections, when students did appear to be interested in truly understanding the material, an argument could be made that the interest was impacted more by the constructivist methodology than by the summative evaluation. For example, during the third lesson the teacher required students to work in small groups on the problems. During this time each of the groups appear to get to work quickly. A number of the students requested that the teacher come over and assist their efforts. One example of a question was, "So we have to add the gas price in, right?" Another example was "Does it matter if I put the total cost right here and the number of tickets

right here?" In addition, one group was seen having a good discussion regarding how they should label the axis on the graph they were making. In yet another instance, a student said to one of their group members, "You have to add the 20 in." The other group member argued, "It's still 60!" The first responded "No it's not." At this point the second recognized his error and said in an understanding voice, "Ohhhhhhhh."

Clearly this was an example of a very active engagement by students in an attempt to master the material. It also seemed likely that this mastery orientation was at least stimulated by the instructional methodology of the teacher. For the purposes of this study it is important to note that, at least for this group of low achieving students, the fact that a graded quiz was coming up did not appear to stifle the collaborative engagement of the students. It was unclear whether the quiz provided any stimulus to the group, however, at the very least it did not appear to hinder their mastery orientation.

Again, as in previous sections, student interview responses suggested a rather tenuous tie to the cognitive difficulty of the material. The interview question that was posed to these students in an attempt to gage their level of mastery orientation to the material was, "During each of the lessons I taught, did you feel you were concerned with understanding the problems in depth or were you only concerned about being able to get the correct answers?" The responses of three of the four interviewees in this class indicated the tie to cognitive difficulty. The first student's response seemed to indicate that an increase in the cognitive difficulty of the task also led to an increase

in her mastery orientation. She responded, "I wanted to make sure I knew how to do it, and sometimes, I just try to get them done. But, if I don't know how to do it, then I really try to learn how to do it. . . . At these first three, I already knew how to do them so I was just trying to get it done, but close to the end it started getting confusing, I wanted to learn how to do it." While her response was somewhat difficult to navigate, ultimately, one must conclude that this student's mastery orientation to learning increased as the cognitive difficulty increased.

A second student responded to the same question in this manner:

Student: I wanted to understand them.

Researcher: [What about] the third one, the table and the graph?

Student: That one I kinda just wanted to get done.

Researcher: Why?

Student: This one, I thought it was more harder, it was harder.

In this case, it appeared as if an increase in the cognitive difficulty leads to a decrease in the mastery orientation of the student. Similarly, when a third student was asked the same question, his response was that he wanted to understand them except for the fourth one. When probed on why that was he responded, "It's just confusing to me . . . it's harder."

This inconsistency in the impact of cognitive difficulty is in fact in line with the apparent role of cognitive difficulty in prior sections. At least for the low-achieving students in this study, an increase in the cognitive difficulty may in some cases have

caused an increase in evidences of self-regulated learning. However, it is very possible that any significant increase in cognitive difficulty may have snuffed out self-regulated learning evidences all together.

Of the four constructs that were conceptualized as evidences of self-regulated learning, the final one to be examined within the credit/no-credit classroom was the student's use of cognitive strategies. That examination actually began by noticing situations where the use of cognitive strategies overlapped with evidences of a mastery orientation to learning. For example, during the videotape of the classroom, a student could be seen walking over to another group of students apparently for assistance. When the teacher questioned the student, the response was "I'm trying to see if I'm doing it right." This was an excellent example of a student who was evidencing a mastery orientation to learning in the midst of an application of a cognitive strategy (namely seeking peer assistance).

Another example of this mix of the two constructs that could be seen in the videotape of this classroom was shown in the following comments by a number of students during a class discussion in which the teacher was attempting to guide the students to a recollection of the Pythagorean Theorem:

Student 1: It's about triangles.

Student 2: [Later when discussing the triangle on the worksheet] It has a right angle.

Student 3: [When probed regarding the triangle and the formula] two to

the 2^{nd} power.

Teacher: Ah, ah, it's about powers, specifically, the 2^{nd} power.

[The teacher calls the sides a, b and c on the board]

Student: a + b = c

Teacher: That's real close.

Student 2: a - b = c [A number of students are "thinking out loud" at this point, eventually the teacher writes the theorem on the board using the sides of the triangle currently under discussion.]

Teacher: If a is three, what does it mean to square it?

[A number of students say things like "nine," "times it," "take it to the 2^{nd} power," etc.]

Student: It will be 9 inches plus 16 inches equals 25.

[The next triangle involves squaring 18. Many if not all students are engaged in this as evidenced by them picking up their calculators to figure it out. When one student gives an incorrect answer, a number of others correct him. The problem involves sides that were measured by students and therefore the Pythagorean Theorem is not an exact equality. Upon seeing this]

Student: Something ain't right there. One of those sides ain't right.

The involvement of these students in the discussion certainly evidenced a mastery

orientation to learning. Clearly it was also evidence of motivation and, their involvement in the discussion itself was an appropriate use of cognitive strategies.

To summarize, consider the following evidence from the videotape of the fourth lesson. As the lesson began, groups got to work quickly and many asked the teacher for assistance. Groups continued to ask questions throughout the entire work time. During this time, the teacher explained that the problems could be solved by a "guess and test" method, because some students had not yet had an algebra class. After saying this a student still asked to see how to do it using the algebra method, saying "I wanna know how to DO it."

Use of Cognitive Strategies

There were a number of examples of the use of cognitive strategies in this class. As has already been discussed, some of these overlapped with a mastery orientation to learning as well as motivation. In this section I will highlight examples which were either unique to the construct of use of cognitive strategies, or, although dependent with one other construct, the specific example has not yet been discussed.

There were numerous examples of students who employed the cognitive strategy of seeking assistance. Throughout the video of each lesson students could be seen working in small groups, asking questions of one another, asking questions of the teacher, or looking to explanations that were written on the chalkboard. In addition, in the interviews students made comments like, "I got stuck, but after I had some help,

I understood," or "But I had help right there. I got that answer from somebody else in class," or "As a group we were talking that we'd just do 40 times the two and then [the teacher] told us that you had to add the 20 too," or "I called her over . . . she showed me how to do them," or "I didn't know what to do on that and [a student name] helped me with just that right there," or finally "I asked somebody in my group for help, because I didn't get it." Explicit in the last quote and implicit in many of the others was the tie to cognitive difficulty. Clearly, the need to seek assistance arises as the cognitive difficulty of the task increases. In this class, students did seek assistance.

A second cognitive strategy that was seen in this class was that of note-taking. In one of the interviews, the researcher asked the student about what appeared to be definitions that were written at the top of one of the student's worksheets. The student responded, "I wrote these down just so I'd remember what they mean." This same student described a couple of other strategies he had employed. In a derivation of seeking assistance, the student described how he and another student divided the task among the two of them. He said, "He come up with the equation, . . . he got stuck on it, and I figured out how much the price of the movies were, then he put it in the equation." Later in the interview this same student described how he used a cognitive strategy of guessing and testing to find a solution to a particular problem.

A second student who was interviewed described yet another derivation of seeking assistance. She related how, after she had completed a problem, she checked her answer with another student. Furthermore, when the problem was presented on

the chalkboard, she checked it again there.

Each of these examples demonstrated that there was indeed evidence of the use of cognitive strategies in this classroom. As stated, this was tied to the cognitive difficulty of the task. As one student summarized when responding to a question regarding his choice to work in a small group, "Sometimes, if I know I can do it, I'll work by myself. But if I know I need help, then I need to work with somebody else." In addition, as in the first classroom, there appeared to be some evidence that the use of cognitive strategies may have been tied to the instructional methodology employed. During the interview with one student, he described how he worked on the worksheets for the first two lessons by himself. Although he was allowed to work with others he did not. During the third lesson this student was mandated to at least sit in a small workgroup (although he could still choose to work by himself). In this setting, the student confirmed that he did indeed get assistance from another member of his group. He was then asked if working in a small group was therefore helpful and he confirmed that it was. Again, the constructivist methodology of instruction was not a variable under study. It was held constant throughout each classroom. The fact that there are evidences this methodology may have impacted certain aspects of self-regulated learning across the classrooms is worth noting. However, it is also important to note that the method of summative evaluation employed in this class did not counteract the instructional methodology. In other words, while the credit/no-credit approach to summative evaluation may not have had a direct positive impact on the students' use

of cognitive strategies in this classroom, it did not appear to have a negative impact either.

Student-Developed Summative Evaluation Classroom

The second experimental alternative to the traditionally graded classroom was a classroom in which the only summative student evaluations were student-developed. It was hypothesized that this complete elimination of letter grades would also eliminate any negative impacts on self-regulated learning that were associated with either a traditionally graded classroom or a classroom that was graded with simply credit or no-credit. In this classroom, students were not given regular quizzes. Copies of the quizzes that were administered in the other two experimental classrooms were given to the students and it was suggested that students (a) assure themselves of their ability to complete problems like those on the quiz, and (b) develop evidences of their mastery over each individual topic.

The most encouraging finding from this classroom was the fact that evidences of self-regulated learning continued to be consistently demonstrated in spite of the fact that there were no summative evaluations imposed on students and, hence, no grades given to students. Indeed, there were some particularly strong evidences in the area of self-efficacy, specifically toward the end of the summer school class after students had been exposed to learning without grades for a period of time. The evidences displayed in the areas of motivation, a mastery orientation to learning, and the use of cognitive

strategies were also consistent with those seen in the first two classrooms. This indicates, that at the very least, moving away from traditional grades does not harm students, and, at best helps.

The demonstrations of self-regulated learning in this classroom, which were consistent with the first two classrooms, provide support for at least two points that began to emerge as data from the three classrooms that were analyzed. First, it demonstrates that the assignment of grades is not a prerequisite to learning. Second, it continues to suggest that the constructivist pedagogy that was employed in all three classrooms served as a catalyst to self-regulated learning. The constructivist pedagogy provided opportunities for students to demonstrate self-efficacy, motivation, a mastery orientation to learning, and the use of cognitive strategies. For example, consider specifically the use of cognitive strategies. In this classroom, as was the case in the first two classrooms, the most frequently occurring cognitive strategy was that of seeking assistance. Students were consistently seeking assistance of the teacher as well as one another; which is a great accomplishment in the classroom. A traditional, lecture based, mathematics classroom would not have provided the kind of opportunity for deploying this strategy as was possible in these three classrooms using a constructivist pedagogy. A complete discussion of the evidences of self-regulated learning in this classroom as well as how those evidences are related to one another and the constructivist pedagogy follows.

Self-Efficacy

As in the first two classrooms, evidences of self-efficacy were rather sporadic. However, the evidences that were seen in this classroom were particularly strong evidences compared to the evidences found in the other two classrooms. For example, in reviewing the videotape of this classroom at the conclusion of the third lesson, the teacher could be seen asking if anybody would be willing to put their table or their graph on the board. In this same class, only 20 minutes earlier, when asked simply if they had any ideas or questions, there was no response. Now after working in groups for awhile, two students quickly volunteered to put both the table and the graph respectively up on the board. The student who had volunteered to put the graph up said in a confident voice "I can do that" and he quickly got up and began a cocky strut to the board. When he had finished, he returned to his seat, however, before sitting down, he said to the class, "My graphs probably the hardest thing up there." He then danced in a manner similar to a professional football player who has just scored a touchdown, said "Yeh, yeh," pointed to his work on the board, and sat down. Finally, when the teacher began reviewing the problem as a class, the teacher said, "Now let's look at [a student name]'s graph." At that point the student said to the class "You hear that, [the student's name]'s."

During the fourth lesson, the teacher again asked for volunteers to show their work on the board. In addition to the student who did the "touchdown dance" after showing his work in lesson three, two different students quickly volunteered to put up

their solutions.

In spite of these particularly impressive displays of self-efficacy, there were other instances where the evidence is less convincing. For example, during the second lesson the teacher was leading a class discussion over the activity on measures of central tendency. The teacher asked the class, "Is that the mean or median?" A few students could be seen whispering answers. It appeared as if they were only half hoping the teacher would hear them. Another example of this somewhat shaking self-efficacy was seen in the video of the first lesson. As in lessons three and four, the students were being asked to show their work on the board. However, in this lesson the student who volunteered had to be coaxed to do so. In addition, while the student was putting his solution on the board, he could be heard saying in a not so confident voice, "Yeh, I think that's it."

Each of these two examples demonstrates the potentially thin line between self-efficacious behavior and a lack of such behavior in low-achieving students. As further evidence of this, consider the following response given by one of the interviewed students when she was asked, "Have these lessons made you any more or less confident of your ability to pass the Graduation Qualifying Exam?" She responded, "In a way it has. See, math is like not a very good subject for me and it has kinda pushed me towards, you know, I need to start focusing in math. I need to start learning, taking notes, whatever in class. So, in a way it has."

Clearly there were examples of positive as well as questionable self-efficacy. It

was interesting to note, however, that the latter two examples of less self-efficacious behavior occurred during the first two lessons while the first two examples of extremely self-efficacious behavior occurred during lessons three and four. This may suggest that the self-efficacy of students improved over time. As an additional evidence that self-efficacy may have improved over time consider the following interview response. This student was also asked, "Have these lessons made you any more or less confident of your ability to pass the Graduation Qualifying Exam?" She responded, "Well, there is one thing I have to admit. Before this whole thing I had this huge hatred for math. But after awhile, just having someone sit me down, and have me work through problem after problem after problem, and actually explain to me how to do it; it actually helped me a lot better, and now I don't hate math as much as I used to."

In the other two experimental classrooms, similar evidences that self-efficacy may have improved over time were not seen. However, it is important to note that like the other two classes, there was some evidence that this behavior may have been associated with the instructional methodology. Clearly the methodology used across classrooms allowed for students to work together, present problems on the board, and discuss the solutions as a class. Nevertheless, this was the only class in which a potential improvement in self-efficacy was seen over time. Perhaps the graded quizzes in the first two classrooms contributed to pushing those students toward the negative side of the thin line between positive and negative self-efficacy.

Motivation

As with each of the first two experimental classrooms, evidence related to the conceptualization of motivation, namely task value and intrinsic motivation, was limited. It was possible that these low achieving students, across classrooms, struggled to manifest these qualities. Consider the following responses from interviews with students in this class:

Researcher: When you were trying to solve the problems and learn the material in these lessons, what motivated you? Why did you try?

Student: Well, because I came to this remediation to learn, I mean because of the ISTEP . . .

Researcher: Let's say you passed the ISTEP, and you're sitting in math . . .

Student: I'd still do them for the grade, yeh. Cause if it was getting graded yeh. But, I mean every once in a while it's good to do it for some practice because you kinda forget after awhile.

Researcher: Now, this one wasn't getting graded. You weren't getting any grade. . . .

Student: A lot of this, like the Pythagorean Theorem, I didn't understand how to do that, so I wanted to do that. And the rest of these, like the algebra and then this stuff, I kinda knew, but it's better to reinforce it so I tried to do all of them, that's why.

Researcher: If there was no ISTEP, . . . do you think you would have done this, or would you have just sat there and done nothing?

Student: Well, I probably would have done it, just, well I don't know, if I was in that situation. I mean some parts, yeh I might, but some parts I might not. I'm not sure if I would or not. . . . I'd probably do the stuff that I don't know. I mean if they're going to send me to remediation, than I might as well do the stuff that I don't know so that I won't get sent back or something. So, I'd just probably try really hard on the stuff that I didn't know. And I might go along with the stuff that I knew already just for extra practice. I don't know, I think maybe.

The confusion in this student's response was overwhelming. At some points he wanted to admit that there was some valuing of the task and/or intrinsic motivation. However, he seemed incapable of communicating that. Recall this was a low-achieving high school student. Schooling and hence learning had been a demoralizing experience for this young man, likely, for most of his life. It may be the case that these students would not recognize real motivation within themselves even if it were beginning to blossom. In turn, the inability to recognize it, would lead to an inability to communicate regarding its presence.

The responses to the same initial question, namely "What motivated you to try?" by each of the other three students who were interviewed in this class, amplify

this point. One said, "I just wanted a better understanding for math, cause, like I said, it's one of my weakest points." A second said, "Because I've been bad at math all my life, and I was trying to keep up with everybody else on the subject." Is this intrinsic motivation? Another responded, "Because, well, you gotta learn it sometime, and I just figured since I already knew a little bit of it I'd just sit here and work on all of it so that I could understand all of it. You know, not just sit there. I mean it is boring at times, but, you gotta figure it out. That's something you need to know." Again, the student seems to be struggling to admit, even to herself, that the task has value.

These somewhat shaky evidences of motivation were not unique to this third classroom. They were seen across all three classes. In addition, like in other classes, there was evidence in this class that motivation was intertwined with some of the other constructs. The student, who was cited in the previous paragraph as wanting "a better understanding for math," was demonstrating a tie between motivation and a mastery orientation to learning. This student continued in the interview to say, "I basically would feel better if I knew it. But, I mean, if I did have a job and math was involved in it, and if I didn't understand it a whole lot, I would probably get laid off or fired or something like that. . . . I would just get eliminated. Cause, I mean, my teacher last year posted a paper with all these different jobs that had something to do with math in it, and she said, don't ever say anything like 'Oh math is never going to be important in my job, that I'm never going to use it, why do I have to do this?' She proved me right, that math is everywhere. And no matter where I turn I'm gonna see it, and if I

hate it, I'm not going to do to well." This response demonstrated the complexity of the relationship between motivation and a mastery orientation to learning. The student was demonstrating a level of motivation. She started with "I basically would feel better if I knew it." However as she continued, the evidence pointed away from an intrinsic motivation and toward a more pragmatic desire to approach the task from a mastery orientation. Regardless of the categorization of this response to either motivation or a mastery orientation to learning, in either case the student was demonstrating some level of self-regulated learning.

To summarize, in looking for evidences of motivation, there was some, albeit limited support for the existence of the construct both in this classroom as well as the previous two. Again, like in the other two classrooms, the construct mixes with some of the other constructs of self-regulated learning. There was no evidence that motivation changed over time. Regarding the issue of cognitive difficulty of the task, in the other two experimental classrooms it appeared as if cognitive difficulty did play a role, however that role varied. Only one of the interviewees in this class gave a response that tied to the cognitive difficulty of the task. Referring to his effort on the worksheet from the third lesson he said, "it was a little bit harder, so I had to try harder." The key piece of this quote may be that it was a LITTLE bit harder. As has been noted in prior sections, for these low achieving students, a small increase in cognitive difficulty may have spurred motivation, however, larger increases in cognitive difficulty may have killed motivation completely.

Mastery Orientation to Learning

There were numerous evidences of a mastery orientation to learning in this classroom. Some have already been introduced in the prior section due to their tie to motivation. However, both through a review of the videotape of the classroom lessons as well as in the interviews of individual students, many additional pieces of supporting data were found for the argument that students in this classroom frequently approached the tasks from a mastery perspective.

Beginning with the very first lesson, during the activity where students were asked to collect data on the lengths of their pinky fingers, all students could be seen actively participating. Students whose specific activities were visible in the videotape of this lesson were seen erasing and cleaning up their work. Indeed, although the teacher offered to give students the solution to the first problem, no student requested that. They apparently were more interested in understanding how to do the material than simply getting the worksheet completed with the correct answers. This interest in understanding continued as the teacher lead a class discussion over the activity. Students volunteered to participate in the discussion both verbally and by showing their work on the board. Finally, while the teacher was still discussing the problems, the bell rang to end this first class period. The students were interested enough in understanding that they did not jump up and leave. In fact, the teacher requested that he be allowed a few more seconds to finish his explanation. The students obliged and remained seated and attentive until he finished.

Similar positive evidences were seen in the second, third and fourth lessons. For example, near the end of the second lesson, the teacher passed out a "practice quiz." Most students appeared to begin working on it. Even one student who was laying his head down on his desk during most of the hour sat up at this point and started working. Perhaps he felt he did not need to listen in class. However, he was clearly now interested in making sure he understood the material. About five minutes after he began working on the "practice quiz," this same student called the teacher over to check his work. As the teacher explained areas where he had made mistakes, the student listened and made changes on his practice quiz.

During the time the class was given to work on the practice quiz, a number of students asked the teacher to assist them. Indeed, the whole class continued to work with some asking for additional help right up until the bell rang to end the class period. A few students even stayed and worked with the teacher after the bell had rung. In addition another student came back early from the break (which was begun by the bell) and appeared to begin working again.

This dedication to understanding the material on the practice quiz was again a demonstration of a mastery orientation to learning. The practice quiz in this class was actually the same document that was given as a graded quiz in each of the other two classes. Interestingly, most of these students demonstrated as strong a desire to understand this material as their peers in the classrooms were grades of some kind were given on the quiz.

Examples of this mastery orientation were also evident in the video of the third and fourth lessons. For example, as the class was going over the work students had put on the board for the fourth lesson, one student rose his hand and said "I got something different for number three." The teacher asked for the equation he used and the student responded by stating his equation for the teacher to put on the board. While the teacher discussed the differences between the two equations now on the board for solving number three, most of the class appeared to be paying attention. Student engagement evidenced by the original students who volunteered to put material on the board, this student who asked questions seeking a deeper understanding of the material, and the remainder of the class who were for the most part attentive during this discussion were each a clear example of a mastery orientation to learning.

As in the other classrooms and constructs, there was again evidence of ties between this construct and other evidences of self-regulated learning. Specifically, students in this classroom could be seen demonstrating a mastery orientation to learning through the use of cognitive strategies. For example, during one class period while some students were putting their work on the chalk board, two other students could be seen sitting in the front of class carefully checking their work against what is being put up on the board. In another case, one girl was seen talking about how she did her work while another girl was leaning over close to her to listen. As the discussion within this small group continued, at least one student could be seen making

changes to her work. Eventually, the group called the teacher over for assistance. After the teacher had left, both of the two original girls were seen erasing and adjusting their work.

Data collected from the interviews with students demonstrated a similarly encouraging view of the students' orientation to learning. One student commented, "I really wanted to know how to do this. Since, what is the point for getting answers when it don't count for anything. I mean, no points or anything for them, so I might as well just learn it." Another student stated, "I was really hoping to understand. That way if that problem ever should happen to come up later on, I would know how to do it." The attitude of understanding for later use was expounded upon by other students. One commented, "The Pythagorean Theorem, I wanted to figure out. Because before I went to this remediation I didn't know how to do it at all, and that's what hurt me on the ISTEP." The ISTEP is the 8th grade equivalent of the Graduation Qualifying Exam. Hence, the upcoming Graduation Qualifying Exam was clearly a factor in this student's interest in mastering the material. This sentiment was echoed by other students. One said, "I was just trying to learn it (pause) for ISTEP."

It is important to remember the fact that the Graduation Qualifying Exam was not to be administered during this class. In this sense, the fact that it played a role in encouraging self-regulated learning was similar to the role college admissions policies or job requirements may play. In other words, teachers are always looking for motivators that are outside of the classroom to encourage student learning within the

classroom. The fact that these activities were outside the classroom distinguishes them from grades given within the classroom. In this particular classroom, the fact that no grades were given did not diminish the students' desire to master the material. One student summarized this well when during his interview he stated, "At the beginning I thought it was for a grade, but toward the end, I knew it was just for practice, but, I still kept on going."

The impact of an increase in cognitive difficulty on a mastery orientation to learning was again complex. During the interview, in discussing her efforts on a particular problem set one student stated, "The first two problems weren't really that bad, but the third one, it started to really be like OK, how do I do this? This is frustrating. . . . I'm gonna need help, and that's when [other students] and I decided to call you over." In this case, an increase in difficulty did not deter the student's desire to master the material. Another student, however, when commenting about a problem that he did not understand stated, "Well, I kinda wanted to understand it, but I was tired of doing a whole bunch of math before. I was kinda tired so I didn't really pay attention but (pause) it wasn't getting too hard." In this case the increase in cognitive difficulty did seem to hinder the student's effort at mastering the material. However, later in the interview this same student states, "but sometimes they got more complicated with all the parenthesis and stuff so, that's what I wanted to know." At this point he was suggesting an increase in cognitive difficulty increased his desire to master the material.

As has been noted in earlier sections, it appeared as if there was some relatively modest level of increase in cognitive difficulty that may have promoted a mastery orientation to learning, however, increases beyond that level were apt to hinder a mastery orientation. As an example of exactly this point, one student when asked about her desire to truly understand the third and fourth worksheets (the fourth was conceptualized to be the most challenging) stated that she just wanted to finish them, she had no interest in understanding them. However, when asked whether she wanted to understand or simply finish the first worksheet she said, "A little of both, finishing it, but yet I still wanted to understand it." When probed as to why she wanted to understand this one, but not the third and fourth worksheets she stated, "It was easy once you get to understand it a little bit. But then it got harder and I wanted to figure it out so that I would know how to do it, so I just kept doing it." Again, as long as the increase in cognitive difficulty was from easy to not too much harder, it may have encouraged self-regulated learning. If, however, the increase in cognitive difficulty was too large, it may have hindered the process.

There was no evidence found in this classroom that the student's orientation to learning changed over time.

Use of Cognitive Strategies

Similar to the first two experimental classrooms, the primary cognitive strategy that was evidenced in this class was that of seeking assistance. This was seen in a

variety of ways. In some cases students sought assistance from one another. In other cases from the teacher, and in yet other cases via attention given to a whole class discussion.

There were numerous examples of students seeking assistance from one another. As early as the first lesson, when students were working primarily as individuals, the videotape of the classroom showed one student turning around and observing the methodology another student was using. Throughout this first lesson students could be seen working together and comparing answers. In videotapes of the third and fourth lessons a similar use of peer assistance was shown. For example, during the fourth lesson the group at the front of the class was seen discussing the work. The audio on the videotape provided evidence that one girl in the group was trying to explain how she thought the problem should be calculated.

Students' responses in the interviews also showed that they had sought the assistance of one another. In discussing one of the algebra problems one student said, "That's when we all starting talking about, ok, maybe we should put both on the chart or maybe just one or the other." A little later in this same interview the researcher probed by asking, "What kind of discussion did you have about putting the numbers on the axis?" The student responded with, "Well, I wanted to go by 100's and then they're all like 100's and 50's. You know, just 100, 150, all like that. And I was like, well, that would work." When another student was asked about his methods for solving a particular set of problems he responded, "We did it together . . . it was

helpful because sometimes I understood it and sometimes when I didn't they would help me and they would tell me which answer it was and how they got it." In discussing the development of the graph in lesson three, yet another student stated "We just talked about it . . . so we talked about it and got we should go by 50 and on the bottom go by one."

There were also many examples of students seeking assistance from the teacher. Sometimes this was combined with seeking assistance from peers. For example, during the fourth lesson the videotape showed a group calling the teacher over and than one girl in the group explaining how she is doing it. When the teacher began responding, the other two students in the group moved in closer and adjusted their view so they could hear and see. Again, data from the interview confirmed the fact that students were seeking assistance from the teacher. When one student was asked what she did when she didn't understand a problem she responded, "We had to ask you to come over and that's when you helped us." Another student responded simply "I asked [the teacher] for some help." A third interviewee made the comment during his interview, "I think I got stuck on some of them, but, I raised my hand and asked for help."

Finally, in the area of seeking assistance, there was evidence that the students were actively participating in the class discussions. During her interview one student talked about her confidence with the work on one activity by saying, "It went pretty smoothly, BUT I wasn't sure when we got the decimals. But once we started going

over it in class . . . I was like, oh, ok, I did it right." In another interview, the researcher asked, "It looks like you had 32,000 and you erased it and changed it to 31,000. How did you know it was wrong?" The student responded with, "I think it's cause when you explained it on the board." Yet another student told the researcher, "I waited until everybody went to the board and saw what they did. And then I looked at my paper to see if I got the same answer, and if not then I would change it and correct it."

Besides the numerous examples and types of seeking assistance, there were other evidences of the use of cognitive strategies in this class. For example, in the videotape a student could be seen asking the teacher for definitions of median and mode and then writing those definitions down in a set of notes. Another type of cognitive strategy implementation was evidenced when during the interview a student was explaining how she actually calculated the median. She said, "I was just marking them off and I remember she [her prior teacher] said just circle the ones that were in the middle." In a similar manner, another student explained how his group had used a "guess and test" strategy to solve some of the problems. In explaining how they determined what scale to use on the axis of a graph he said, "We tried a whole bunch of things." Another student explains his guesses by saying, "I had to keep erasing . . . I came up short, so I had to make the numbers bigger."

As was the case in each of the first two classrooms, the use of cognitive strategies was spurred by the difficulty of the material. One student summarizes this

fact by saying "I needed help with this one because I was kinda confused." Also, as in the other two classes, there was no evidence that the use of cognitive strategies changed over time.

CHAPTER 5

DISCUSSION

Summary of Findings

The study was designed to investigate three questions around the topic of summative student evaluations. Those questions were:

1. How does the method of summative student evaluation impact self-regulated learning in low-achieving high school mathematics students?

2. How does the cognitive level of the task mediate the relationship between method of summative student evaluation and self-regulated learning?

3. How does the length of time students are exposed to a method of summative student evaluation impact self-regulated learning?

The methods of summative student evaluation that were considered included: (a) traditional grading (A, B, C, D, or F), (b) credit/no-credit, and (c) student-developed summative evaluation.

The dependent variable, self-regulated learning, was conceptualized to consist of four factors. They were: (1) self-efficacy, (2) motivation, (3) mastery orientation to learning, and (4) use of cognitive strategies.

Regarding the first question, there were clear evidences of self-regulated learning found in all three experimental classrooms. Those evidences were seen in the

analysis of each of the four constructs as well as through behaviors and statements that reflected a combination of some of the four constructs. A summary of the existence of the evidences in each area is shown in table 1.

Table 1

Summary of Evidences Found

Constructs of Self-Regulated Learning	Classrooms		
	Traditionally Graded	Credit/No-Credit	Student-Developed
Self-Efficacy	✓	✓	✓
Motivation	✓	✓	✓
Mastery Orientation	✓	✓	✓
Use of Cognitive Strategies	✓	✓	✓

The work of some of the authors whose articles were reviewed for this study suggested that the classrooms in which grading was lessened or eliminated may yield greater evidences of self-regulated learning (Butler & Nisan, 1986; Dylan & Wiliams, 1998; Kirschenbaum, 1973). Indeed, the work of Dylan and Wiliams offered particularly promising hope for low-achieving students. Nevertheless, because of the fact that evidences of self-regulated learning were found in all three classrooms, this study was unable to demonstrate a clear connection between the elimination of grading and an increase in self-regulated learning strategies. On the other hand, other literature argued that grades are *necessary* to maintain the motivation of students

(Ebel, 1974). This was clearly not the case either in this study. Across all three classrooms, there were indicators of motivation, as well as each of the other three constructs that were used to operationalize self-regulated learning. Therefore, this study clearly did not support the hypothesis that grades are needed as a motivational tool. In summary, while the lessening or elimination of grading could not be shown to increase self-regulated learning, it clearly did not lead to a decrease in the construct either.

Regarding the second question, "How does the cognitive level of the task mediate the relationship between method of summative student evaluation and self-regulated learning?" significant evidence was found to suggest the cognitive level of the task does play a role in this relationship. However, the role does not appear to be a mediating role. In other words, the cognitive difficulty of the task clearly impacts self-regulated learning, but, it does not appear to have a varying impact which is dependent on the nature of the summative student evaluation. Indeed, across all three classrooms it was clear that increasing the cognitive difficulty of the task raised the potential for lessening self-regulated student learning. Whether or not self-regulated learning was negatively impacted appeared to be based on the degree to which the cognitive difficulty of the task was increased. In cases where the difficulty of the task was significantly increased, self-regulated learning was diminished. However, in cases where the increase in difficulty was minimal, self-regulated learning was often actually increased.

As discussed briefly in chapter four, this is consistent with the work of Lev Vygotsky (1978) who suggested each student possessed both an actual developmental level and a zone of proximal development. Vygotsky explains the zone of proximal development when he writes:

> It is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. . . . The zone of proximal development defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are currently in an embryonic state. These functions could be termed the "buds" or "flowers" of development rather than the "fruits" of development. (p. 86)

In the current study, when students were confronted with tasks in which the cognitive difficulty had been increased slightly, it is likely these tasks remained within the student's zone of proximal development. Under those circumstances, the impact on self-regulated learning was positive. On the other hand, when the increases in cognitive difficulty were too large, the impact on self-regulated learning was negative. It is probable that these were cases where the increase in cognitive difficulty pushed the task beyond the upper boundary of the student's zone of proximal development.

Regarding the third question, "How does the length of time students are exposed to a method of summative student evaluation impact self-regulated learning?"

little evidence was found. In almost all cases, across all three classrooms, and in regards to all four constructs that comprise self-regulated learning, there was no evidence found to support the theory that the relationship between method of summative student evaluation and self-regulated learning may change over time. The one potential exception to this rule was in an analysis of self-efficacy in the student-developed summative assessment classroom. In this one case, there was a small amount of credible evidence that self-efficacy may have improved over time.

This finding provides support for Brookhart's (1997) model of assessment in which she suggests that classroom assessment directly impacts self-efficacy. In a later study, Brookhart and DeVoge (1999) also show that the reception of low marks leads to decreased self-efficacy. Given the focus of the current study on low-achieving students, it is reasonable to conclude that they entered the study with relatively low self-efficacy. It is also reasonable to assume that it may take time to change the perception they possess of their mathematical ability. The student-developed summative assessment classroom was the only classroom in which marks were completely eliminated. It was in this classroom, during the last lesson, that the strongest demonstrations of positive self-efficacy were seen. In future research, it would be wise to expand the length of the study to verify the current observations as well as determine if self-efficacy may flourish to an even greater degree given longer exposures to the elimination of marks.

Recommendations for Practice

There are two areas in which the results of the current study may offer insight to educational practitioners. The two areas are summative student evaluation or the more general topic of classroom assessments and classroom pedagogy. In the area of classroom assessment the current study confirms the work of Cameron and Pierce (1994) who suggest that rewards (in this case grades) do not necessarily decrease the motivational level of those receiving the rewards. The current study also provides support for Brookhart's (1997) model of classroom assessment which suggests that the assessments impact self-efficacy which in turn impacts effort. The strongest support for this can be seen in the very clear examples of self-efficacy demonstrated by the students in the student-developed summative assessment classroom at the end of the study. Recall the student who was doing a dance in celebration of his work on the board.

If Brookhart's model is accurate, classroom assessments which lead to increases in self-efficacy are to be desired. Certainly the removal of grades offers that potential, however, this study also suggested that the move to a credit/no-credit approach to grading may serve the purpose, at least for low achieving students. Brookhart and DeVoge (1999) have shown that students who have previous experience with receiving low marks tend to have lower self-efficacy and hence decreased effort. Perhaps for these students, a move simply to a credit/no-credit approach to grading would allow for the desired increases in self-efficacy. In the

current study, it was in the credit/no-credit classroom that students demonstrated a genuine excitement for earning credit on a particular quiz. Recall the student who reacted with the comment, "My name goes in the box! My name up in the box this time, I got it RIGHT!"

As a transition from the area of classroom assessment to the area of pedagogy, consider the work of Kluger and DeNisi (1996) who argue that feedback must support learning as opposed to directing attention to self. Clearly the traditionally graded classroom as well as the credit/no-credit classroom included feedback (namely summative evaluations) that directed attention to the student. Potentially due to the low-achieving students in this study, that fact did not appear to have a strong detrimental impact on self-regulated learning. However, through the constructivist approach to teaching that was employed in all three classrooms, students did receive consistent and continual feedback that supported learning. It is thus not surprising that all three classrooms demonstrated some level of self-regulated learning.

While this study was not specifically designed to investigate the constructivist approach to teaching mathematics, the qualitative nature of the study's design allowed for the gathering of a fair amount of evidence that seemed to relate this pedagogy to self-regulated learning. Indeed, the fact that evidences of self-regulated learning were found in all three classrooms suggests that this consistent factor may be an even more promising tool for improving self-regulated learning than the variable under study, namely, summative student assessment. This would support the argument of Silver,

Kilpatrick, and Schlesinger (1990) who argue that math classrooms should function as a community of learners. It would also support the positions of Bracey (2000), Geist (2000), and Stigler and Hiebert (1997) all of whom argue for a move toward a more constructivist approach to the teaching of mathematics.

Recommendations for Future Research

The current study was limited in a number of ways. First and foremost, the qualitative nature of the study limits its generalizability. For example, this study found evidence that self-efficacy may improve over time in a classroom that employs student-developed summative assessments. However, it is quite possible that this evidence was a factor of the particular students in the student-developed summative assessment classroom rather than the method of summative assessment. The finding would need to be verified in other studies, preferably in some quantitative studies. For example, a study that was designed to gather evidence specifically pertaining to self-efficacy over the course of an entire school year within a classroom that employs student-developed summative assessments as opposed to grading would be appropriate. To conduct such a study, the researcher could select a quantitative instrument designed to measure self-efficacy. After the selection of such an instrument, the data could be collected at multiple points throughout the school year. A statistical test could then be performed to determine the statistical significance of any differences that are found in self-efficacy over the course of the school year.

While the generalizability was limited by the qualitative nature of the study, the degree to which concepts were allowed to grow out of the qualitative data was also limited by the relatively rigorous framework that was developed for the analysis of the data. As was discussed in chapter three, the method that was used for coding the data in this study lies somewhere between a grounded theory approach to data analysis (where all analysis grows out of the data) and the creation of a predefined list of codes that limit what is seen. In this study, the types of things that were being looked for during the initial coding were predefined (namely self-efficacy, motivation, mastery orientation to learning, and use of cognitive strategies). The degree to which coding was allowed to grow out of the data was limited to coding *within* each of these constructs.

Another limitation of this current study centers around the fact that the study was unable to conclude that a removal or lessening of grades would lead to increased levels of self-regulated learning. The design of the study may have hindered that capability. For example, in the student-developed summative assessment classroom students were given no training or skill development in how to develop evidences of their own learning. It was clear in this study that these low-achieving students needed such training in order to develop that capability to generate their own summative evaluations. Based on the nearly complete lack of students developing evidences of their learning, it is quite possible even average or high-achieving students may need such training. Future research should incorporate student training in student

developed summative assessments as a component of the research study.

In addition, because the current study was limited to low-achieving secondary school mathematics students, much of what was found may well be unique to low-achieving students. For example, the fact that self-regulated learning did not appear to be either hindered or enhanced by changes in methods of summative student evaluation may be related to the fact that low-achieving students have developed a callousness to any forms of summative evaluation. The results may be very different for average or high achieving students. If a callousness has indeed developed in low-achieving students, the results may also be different for younger low-achieving students. Future research should investigate the relationship between self-regulated learning and summative student assessment with high-achieving students, students who fall into the average achievement range, and elementary level low-achieving students.

A further limitation of the current study was the focus on the area of mathematics. It is certainly possible that students' capabilities to generate their own evidences of learning may differ in other disciplines. In future research it would be wise to investigate the relationship between self-regulated learning and summative student assessment in disciplines other than mathematics.

Finally, to the degree to which quantitative studies could be developed to investigate any of the specific components of self-regulated learning within any of the three summative student evaluation conditions, it would certainly be helpful in further the understanding of these relationships. While it is unlikely a quantitative study could

investigate each of the constructs of self-regulated learning within each of the summative evaluation conditions in the broad manner that was incorporated in the current study, a focus on one of the areas of self-regulated learning within one type of summative student evaluation condition may be possible. Indeed an example of one such study has already been suggested regarding the construct of self-efficacy within the student-developed summative evaluation classroom and how self-efficacy may change over time.

Appendix A

Lesson Descriptions

APPENDIX A

LESSON DESCRIPTIONS

Initial Instructions

At the beginning of the first researcher led lesson, the following instructions were given to each group.

Group 1 - Traditional Grades

"I will be leading you through four activities over a period of three days. In each activity, you will have the opportunity to use a calculator if you would like it. You may work alone or with other people. If you get stuck, you may request a hint from me. I will also give you additional assistance if you need it. Finally, if you would like a copy of the answers you may request that. Once each activity has been completed, we will review the material together. After the review, you will be given a brief quiz. After our final class together, I will hold a drawing to give away a $20 gift certificate to Applebee's. For each A you earn on the quizzes, your name will be put in the drawing four times. For each B, your name will be put in the drawing three times. Your name will be put in twice for each C and once for each D."

Group 2 - Student-developed Summative Assessment

"I will be leading you through four activities over a period of three days. In

each activity, you will have the opportunity to use a calculator if you would like it. You may work alone or with other people. If you get stuck, you may request a hint from me. I will also give you additional assistance if you need it. Finally, if you would like a copy of the answers you may request that. Once each activity has been completed, we will review the material together. As you may know, students who are unable to pass the Graduation Qualifying Exam, may earn a diploma by demonstrating their mastery of the content in an alternative way. While you are working on this task, please consider how you might demonstrate your mastery of the concept. You may request my assistance in developing a demonstration of your mastery, or you may develop something on your own. You may do this during our class time or outside of class."

Group 3 - Credit/No-credit

"I will be leading you through four activities over a period of three days. In each activity, you will have the opportunity to use a calculator if you would like it. You may work alone or with other people. If you get stuck, you may request a hint from me. I will also give you additional assistance if you need it. Finally, if you would like a copy of the answers you may request that. Once each activity has been completed, we will review the material together. After the review, you will be given a brief quiz. On each quiz you will be given a score of either 'credit' or 'no credit'. After our final class together, I will hold a drawing to give away a $20 gift certificate

to Applebee's. For each quiz that you earn 'credit' on, your name will be placed in the drawing for the gift certificate."

Lesson 1

Topic

Indiana Academic Standard 2 - Geometry

Subtopic

Apply the Pythagorean Theorem to solve real-world problems in two dimensions.

Task

Each student was given two pieces of 8.5 by 11 inch paper, a ruler, and a calculator. The students were asked to:

1. Start at the bottom left corner of the paper and measure 3 inches to the right, mark it.

2. Start at the bottom left corner of the paper and measure 4 inches up, mark it.

3. Connect the two marks and measure the new line.

4. Start at the bottom left corner of the paper and measure 6 inches to the right, mark it.

5. Start at the bottom left corner of the paper and measure 8 inches up, mark it.

6. Connect the two marks and measure the new line.

Discussion of Task

Students were asked to share their results on the board. Then, the researcher led the class as a whole in a discussion aimed at finding a pattern that would allow students to calculate the hypotenuse for any triangle if they knew the lengths of both legs (Pythagorean theorem). Once the theorem had been summarized on the board, worksheet one was given to each student. They were told that they may work alone or with a partner.

The hint students were given if they requested one was the first row of the table completed with the algebraic work shown.

Discussion of Worksheet

After students had been given an opportunity to attempt all or most of the worksheet, the researcher led a class discussion over the calculations and answers to worksheet one.

Assessment

For groups one and three, assessment one was administered. The open-ended

problem was taken from the fall 2000 Indiana Graduation Qualifying Exam.

Group two students were asked at the beginning of the lesson to think about how they might demonstrate mastery of this topic. Because a formal summative assessment was not a part of the lesson for this group, additional time was available both for completing and discussing both the task and the worksheet.

Lesson 2

Topic

Indiana Academic Standard 3 - Statistics

Subtopic

Find the mean, median, and mode of a set of numbers.

Task

Students were given a ruler and asked to measure the length of each of their two pinky fingers to the nearest millimeter. They were also allowed to use calculators if they so desired. Students were told if they have questions during this activity, they should write the question down. As students completed the activity, the data was collected on the board.

Discussion of Task

The class engaged in a discussion regarding any questions that came up during the data collection process or that arose from looking at the data. The researcher then led a discussion in statistical measures of central tendency. The discussion concluded with the development of definitions for mean, median and mode. Students were then given worksheet two. They were told that they may work alone or with a partner.

The hint given to students who requested one was the definitions of mean, median, and mode.

Discussion of Worksheet

The instructor will go over the worksheet reviewing the concepts of mean, median and mode.

Assessment

For groups one and three, assessment two was administered. All items on this assessment were taken from the spring 2001 Indiana Graduation Qualifying Exam.

Group two students were asked at the beginning of the lesson to think about how they might demonstrate mastery of this topic. Because a formal summative assessment was not a part of the lesson for this group, additional time was available both for completing and discussing both the task and the worksheet.

Lesson 3

Topic

Indiana Academic Standard 6 - Functions

Subtopic

Convert among tables, graphs, equations, and expressions in words.

Task

Students were placed in groups of three and asked to complete worksheet three. While they were completing this activity, the researcher monitored the progress of each group and offered assistance if it was requested or if it was appropriate. For students who requested a hint, the hint was the first line of the table completed and the first point of the graph plotted.

Discussion of Worksheet

After most groups had been given enough time to throughly attempt each piece of the activity, the researcher led a discussion covering the solutions and explained the process for arriving at the solutions.

Assessment

For groups one and three, assessment three was administered. The open-ended problem was taken from the fall 2000 Indiana Graduation Qualifying Exam.

Group two students were asked at the beginning of the lesson to think about how they might demonstrate mastery of this topic. Because a formal summative assessment was not a part of the lesson for this group, additional time was available both for completing and discussing both the task and the worksheet.

Lesson 4

Topic

Indiana Academic Standard 5 - Algebra

Subtopics

1. Solve equations in one variable by algebra or by trial and error.
2. Write an equation (in one variable) to solve a real-life problem.

Task

Students were again placed in groups of three and asked to complete worksheet four. The researcher again monitored the progress of each group and offered assistance if it was requested or if it was appropriate. For the first two

problems, the hint given to students who requested one was the first step of the solution. For the third problem the hint was "let x stand for the price Toni paid for each video."

Discussion of Worksheet

After most groups had been given enough time to throughly attempt each piece of the activity, the researcher led a discussion covering the solutions and explained the process for arriving at the solutions.

Assessment

For groups one and three, assessment four was administered.

Group two students were asked at the beginning of the lesson to think about how they might demonstrate mastery of this topic. Because a formal summative assessment was not a part of the lesson for this group, additional time was available both for completing and discussing both the task and the worksheet.

Appendix B

Interview Protocol

APPENDIX B

INTERVIEW PROTOCOL

Depending on the student's response, each of the questions were followed up with deeper probes. To assist the student in recalling exactly what they were doing, students were shown a copy of the worksheet he or she was completing during each lesson. For each question, the researcher asked the student to respond for each worksheet/lesson (thus getting a measure of how their thought processes may have changed over time).

Question 1

This question was aimed at identifying the student's use of cognitive strategies. The question was, "How did you go about thinking through the problems, or trying to figure them out?"

Question 2

This question was aimed at identifying any changes in the self-efficacy of the student. The question was, "Thinking only about the three lessons I worked with you, have these made you any more or less confident in your ability to pass the math portion of the Graduation Qualifying Exam?"

Question 3

This question was aimed at identifying the motivational level of the student as evidenced by the task value identified by the student. The question was, "Did you find the activities you were involved in valuable? Why or why not?"

Question 4

This question was aimed at identifying the motivational level of the student as evidenced by the level of intrinsic motivation of the student. The question was, "When you did try to solve the problems and learn the material during the lessons with me, what motivated you; why did you try?"

Question 5

This question was aimed at discovering the degree to which the student demonstrated a mastery orientation to the tasks. The question was, "During each of the lessons I taught, did you feel you were concerned with understanding the problems in depth or were you only concerned about being able to get the answers correct? Why?"

Question 6

This question was only given to students in the Student-developed Summative Assessment group. The question was, "Did you think about or develop ways of

showing your ability to complete these tasks? Why or why not?"

Appendix C

Protocol Clearance from the Human Subjects
Institutional Review Board

WESTERN MICHIGAN UNIVER

Huma

Centennial
1903·2003 Celebration

Date: February 5, 2003

To: Jainping Shen, Principal Investigator
John Ritzler, Student Investigator for dissertation

From: Mary Lagerwey, Chair Mary Lagerwey

Re: HSIRB Project Number 03-01-12

This letter will serve as confirmation that your research project entitled "An Investigation on the Impact of Structure of Summative Student Evaluation on Self-Regulated Learning" has been **approved** under the **exempt** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may **only** conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: February 5, 2004

Walwood Hall, Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276

BIBLIOGRAPHY

Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high-achieving students: Relations to advanced reasoning, achievement goals, and gender. Journal of Educational Psychology, 90(1), 94-101.

Ames, C. (1984). Achievement attributions and self-instructions under competitive and individualistic goal structures. Journal of Educational Psychology, 76(3), 478-487.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. Journal of Educational Psychology, 84(3), 261-271.

Argyris, C. (1964). Integrating the individual and the organization. New York: Wiley.

Arnstine, D., & McDowell, J. A. (1993). Unfair rewards: Merit pay, grades, and a flawed system of evaluation. Teacher Education Quarterly, 20(2), 5-21.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. Journal of Educational Research, 85(2), 89-99.

Bennis, W., & Nanus, B. (1985). Leaders: The strategies for taking charge. New York: HarperPerennial.

Berliner, D., & Casanova, U. (1988). Are grades undermining motivation? Instructor, 98(3), 18-19.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80(2), 139-148.

Block, P. (1993). Stewardship: Choosing service over self-interest. San Francisco: Berrett-Koehler.

Blount, H. P. (1997). The keepers of numbers: Teachers perspectives on grades. The Educational Forum, 61, 329-334.

Bong, M. (2002). Predictive utility of subject-, task-, and problem-specific self-efficacy judgments for immediate and delayed academic performances. The Journal of Experimental Education, 70(2), 133-162.

Bracey, G. W. (2000). Trying to understand math for understanding. Phi Delta Kappan, 81(6), 473-474.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. Applied Measurement in Education, 10(2), 161-180.

Brookhart, S. M. & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. Applied Measurement in Education, 12(4), 409-425.

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. Journal of Educational Psychology, 78(3), 210-216.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. Review of Educational Research, 64(3), 363-423.

Cizek, G. J. (1996). Grades: The final frontier in assessment reform. NASSP Bulletin 80(584), 103-110.

Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. Educational Measurement: Issues and Practice, 19(2), 16-23.

Conner, K., Hairston, J., Hill, I., Kopple, H., Marshall, J., Scholnick, K., & Schulman, M. (1985). Using formative testing at the classroom, school, and district levels. Educational Leadership, 43(2), 63-67.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. Educational Psychologist, 18(2), 88-108.

Covington, M. V., & Omelich, C. L. (1985). Ability and effort valuation among failure-avoiding and failure-accepting students. Journal of Educational Psychology, 77(4), 446-459.

Deming, W. E. (1986). Out of the Crisis. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Dweck, C. S. (1986). Motivational processes affecting learning. American Psychologist, 41, 1040-1048.

Ebel, R. L. (1974). Shall we get rid of grades?. Measurement in Education, 5(4), 1-5.

Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary academic subjects: Implications for instruction in measurement. Educational Measurement: Issues and Practice, 12(3), 23-30.

Friedman, S. J. (1998). Grading teachers' grading policies. NASSP Bulletin, 82(597), 77-83.

Friedman, S. J., & Manley, M. (1992). Improving high school grading practices: "Experts" vs. Practitioners. NASSP Bulletin, 76(544), 100-104.

Friedman, S. J., & Troug, A. L. (1999). Evaluation of high school teachers' written grading policies. ERS Spectrum: Journal of School Research and Information, 17(3), 34-42.

Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. Exceptional Children, 53(3), 199-208.

Garcia, T., McCann, E. J., Turner, J. E., & Roska, L. (1998). Modeling the mediating role of volition in the learning process. Contemporary Educational Psychology, 23, 392-418.

Geist, E. A. (2000). Lessons from the TIMSS videotape study. Teaching Children Mathematics, 7(3), 180-185.

Glasser, W. (1992). The quality school (2nd ed.). New York: HarperPerennial.

Glesne, C., & Peshkin, A. (1992). Becoming qualitative researchers. White Plains, NY: Longman.

Goodlad, J. I. (1984). A place called school. New York: McGraw-Hill.

Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of processing. Journal of Educational Psychology, 83(2), 187-194.

Gorman, W. J. (1989). Effective student evaluation. Education Canada, 29(3), 4-9, 15.

Herzberg, F. (1987). One more time: How do you motivate employees? Harvard Business Review, 65(5), 109-120.

Jones, E. D., Wilson, R., & Bhojwani, S. (1997). Mathematics instruction for secondary students with learning disabilities. Journal of Learning Disabilities, 30(2), 151-163.

Juarez, T. (1994). Mastery grading to serve student learning in the middle grades. Middle School Journal, 26(1), 37-41.

Juarez, T. (1996). Why any grades at all, Father? Phi Delta Kappan, 77(5), 374-377.

Kirschenbaum, H. (1973). Boxed in by grades. Teacher, 91(1), 46-48.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119(2), 254-284.

Langford, D. P. (1997). Quality learning training manual. Molt, MT: Langford International.

Lewis, C. (1992). Making sense of common sense. Critical Studies in Mass Communication, 9, 277-292.

Likert, R. (1967). The human organization: Its management and value. New York: McGraw-Hill.

Lindner, R. W., & Harris, B. (1992). Self-regulated learning: It's assessment and instructional implications. Educational Research Quarterly, 16(2), 29-37.

Malehorn, H. (1984). Ten better measures than giving grades. Clearing House, 57(6), 256-267.

Maslow, A. H. (1970). Motivation and personality (2nd ed.). New York: Harper & Row.

Malpass, J. R., O'Neil, H. F., Jr., & 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), 281-289.

Mathematical Sciences Education Board. (1991). Counting on you: Actions supporting mathematics teaching standards. Washington, DC: National Academy Press.

McCabe, J. S. (1978, March). Characteristics of classroom environments and their relationship to cognitive development. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada.

McCombs, B. L. (1984). Processes and skills underlying continuing intrinsic motivation to learn: Toward a definition of motivational skills training interventions. Educational Psychologist, 19(4), 199-218.

McGregor, D. (1960). The human side of enterprise. New York: McGraw-Hill.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis (2nd ed.). Thousand Oaks, CA: Sage.

Page, E.B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. Journal of Educational Psychology, 49, 173-181.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. Contemporary Educational Psychology, 21(4), 325-344.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. Contemporary Educational Psychology, 20, 426-443.

Pajares, F., & Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. The Journal of Educational Research, 90(6), 353-360.

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

Rothney, J. W. M. (1955). Evaluating and reporting pupil progress. Washington D.C.: National Education Association.

Schloss, P. J., Smith, M. A., & Posluzsny, M. (1990). The impact of formative and summative assessment upon test performance of special education majors. Teacher Education and Special Education, 13(1), 3-8.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. Journal of Educational Psychology, 73(1), 93-105.

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

Schunk, D. H. (1999). Social-self interaction and achievement behavior. Educational Psychologist, 34(4), 219-227.

Senge, P. M. (1990). The fifth discipline. New York: Doubleday.

Shen, C., & Pedulla, J. J. (2000). The relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: A cross national analysis. Assessment in Education, 7(2), 237-253.

Silver, E.A., Kilpatrick, J., & Schlesinger, B. (1990). Thinking through Mathematics. New York: College Entrance Examination Board.

Smith, D. L. (1992). Validity of faculty judgements of student performance: Relationship between grades and credits earned and external criterion measures. The Journal of Higher Education, 63(3), 329-340.

Starch, D., & Elliott, E. C. (1912). Reliability of grading work in mathematics. School Review, 21, 254-259.

Stewart, L. G., & White, M. A. (1976). Teacher comments, letter grades, and student performance. Journal of Educational Psychology, 68(4), 488-500.

Stiggins, R. J. (1999). Assessment, student confidence, and school success. Phi Delta Kappan, 81(3), 191-198.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. Educational Measurement: Issues and Practice, 8(2), 5-14.

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction. Phi Delta Kappan, 79(1), 14-21.

Stipek, D. J., & Kowalski, P. S. (1989). Learned helplessness in task-orienting versus performance-orienting testing conditions. Journal of Educational Psychology, 81(3), 384-391.

Tyler, R. W. (1942). General statement on evaluation. Journal of Educational Research, 35, 492-501.

VanderStoep, S. W., Pintrich, P. R., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. Contemporary Educational Psychology, 21, 345-362.

Vygotsky, L. S. (1978). Zone of proximal development: A new approach. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), Mind in society: The development of higher psychological processes (pp. 84-91). Cambridge, MA: Harvard University Press.

Werts, C., Linn, R. L., & Jöreskog, K. G. (1978). Reliability of college grades from longitudinal data. Educational and Psychological Measurement, 38, 89-95.

Williams, J. E. (1996). The relation between efficacy for self-regulated learning and domain-specific academic performance, controlling for test anxiety. Journal of Research and Development in Education, 29(2), 77-80.

Yung, B. H. W. (2002). Same assessment, different practice: Professional consciousness as a determinant of teachers' practice in a school-based assessment scheme. Assessment in Education, 9(1), 97-117.

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(4), 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.