Loyola University Chicago
Loyola eCommons

Dissertations

Theses and Dissertations

1989

The Effects of Metacognitive Strategies and Social Interaction upon Mathematics Learning: A Reciprocal Teaching Study

Denis W. Rudy
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Education Commons

Recommended Citation

Rudy, Denis W., "The Effects of Metacognitive Strategies and Social Interaction upon Mathematics Learning: A Reciprocal Teaching Study" (1989). *Dissertations*. 2705.
https://ecommons.luc.edu/luc_diss/2705

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1989 Denis W. Rudy

THE EFFECTS OF METACOGNITIVE STRATEGIES AND SOCIAL INTERACTION UPON MATHEMATICS LEARNING: A RECIPROCAL TEACHING STUDY

by

Dennis W. Rudy

A Dissertation Submitted to the Faculty of the Graduate School
of Loyola University of Chicago in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

April

1989

ACKNOWLEDGEMENTS

The author would like to sincerely thank his dissertation committee chairperson, Dr. Todd Hoover, for all of his help in editing and proofreading this dissertation. In 1985, Dr. Hoover first introduced the author to the world of computers. Since that time, my professional life has benefited by both the word-processing and the statistical capabilities that Dr. Hoover has taught me.

A heart-felt appreciation is extended to Dr. Ron Morgan, whose interest in learning theories and reserach methods has altered my view of the teaching and learning process forever. His knowledge, advice, and interest in this research project and all other learning endeavors was most helpful and motivational.

My relationship with Dr. John Dossey began in the fall of 1977, when he exposed me to that first research article in mathematics education. The standards he set as Chairman of the Advisory Committee for my Master's thesis, and in his own professional life have influenced me since.

A special thank you to Dr. Kay Monroe Smith, who I first worked for as a graduate assistant within the Department of Curriculum and Instruction. Dr. Smith allowed me to teach in her elementary mathematics methods course my first year at Loyola

University of Chicago. Since then, I have always been involved in teaching one graduate level education course or another.

To all four of these people, I want to acknowledge your time and effort in reading and revising this dissertation. Thank you for the interest you have shown in my work.

Finally, I would like to thank my wife, Carol, for her support and encouragement during this research project. Two years of work required much encouragement.

"...the best that any individual scientist, especially any psychologist, can do seems to be follow his own gleam and his own bent, however inadequate they may be. In fact, I suppose that actually this is what we all do. In the end, the only sure criterion is to have fun. And I have had fun" (Tolman, 1959, p. 152).

VITA

The author, Dennis W. Rudy, is the son of Arthur and Rosemarie Rudy. He was born November 5, 1953, in Chicago Illinois.

His elementary and secondary education was obtained in the parochial schools within the Archdiocese of Chicago. He graduated from St. Laurence High School, Burbank, Illinois, in 1971.

In September, 1971, Mr. Rudy entered DePaul University of Chicago, receiving the degree of Bachelor of Science in elementary education in February, 1975. Mr. Rudy taught in the parochial elementary schools within the Archdiocese of Chicago from January, 1975, through August, 1977.

In September, 1977, Mr. Rudy formally began a graduate program at Illinois State University in Normal, Illinois. At this time he also taught in a junior high school in Bloomington, Illinois, for the Diocese of Peoria. In September of 1978, Mr. Rudy taught mathematics in a parochial high school in Aurora, Illinois, which was within the Diocese of Joliet.

Mr. Rudy began his career in the public schools in Septemebr, 1979, in Lisle, Illinois teaching mathematics at a junior high school. During this period of time, Mr. Rudy was formally awarded his Masters of Science degree in elementary from Illinois State University in December, 1982.

In July, 1985, Mr. Rudy received a sabbatical leave from the Lisle public schools and entered Loyola University of Chicago to pursue a doctoral program of study. While at Loyola University, Mr. Rudy served as a graduate assistant for the Department of Curriculum and Instruction within the School of Education, and also taught an undergraduate mathematics methods course for the university with the rank of instructor.

In August, 1986, Mr. Rudy again returned to the Lisle public schools for an additional year as a mathematics teacher in a junior high school. In August, 1987, he left the Lisle public schools to become part of the district office administration at Community Consolidated School District 15 in Palatine, Illinois. Mr. Rudy first began as Director of Elementary Programs for District 15, and is currently Director of Curriculum Implementation and Assessment for this public school district.

Mr. Rudy has presented at various educational conferences hosted by professional organizations of which he is a member. These include the National Staff Development Council (NSDC) and the Mid-Western Educational Research Associaltion (MWERA), an affiliate of the American Educational Research Association (AERA). Other organizations to which he belongs are the National Council of Teachers of Mathematics (NCTM) and the Association for Supervision and Curriculum Development (ASCD).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CONTENTS OF APPENDICES

Appendix

CHAPTER 1

INTRODUCTION

A number of studies from the domains of cognitive science and specific content areas of education have been conducted in an effort to train students in a laboratory setting to use various strategies that purport to aid in the recall of information and thus increase learning potential. These attempts to uncover discrete learning strategies have evolved into a systematic search for more generalizable strategies that would not only be applicable to a single content area, but would also transfer to other content areas and to the regular classroom setting to improve academic performance.

Attempts to discover these more generalizable strategies have led to the development of an entirely new field of research directed at the training of metacognitive and problem-solving skills. One generally accepted definition taken from the extant literature describes metacognition as: "one's knowledge concerning one's own cognitive processes and products or anything related to them. Metacognition refers, among other things, to the active monitoring and consequent regulation and orchestration of these processes..." (Flavell, 1976; p.232).

It is surmised that students' use of these metacognitve strategies would result in an increased academic performance over those students who did not use these strategies. That is why so many researchers and classroom practicioners find the topic of metacognition so worthy of investigation.

This interest in learning strategies and metacognition has resulted in numerous studies that include, but are not limited to research in mathematics and science problem-solving, reading comprehension, intelligence training, study skills, thinking skills, and instruction designed to enhance a student's ability to learn. These studies have produced a plethora of strategies, some with exotic names, that range from content specific to more general rehearsal and organizational strategies designed to increase one's memory and retention. It is unfortunate that for the classroom practicioner many of these strategies are indistinguishable from one another. Also, these strategies still lack the capability of transfer and generalizability problem that plagued the earlier laboratory research.

The Reciprocal Teaching Method

One model of strategy instruction that has reported particular success with respect to transfer and generalizability to other content areas and regular classroom settings is the reciprocal teaching method (Palinscar & Brown, 1984). Their model of interactive instruction termed reciprocal teaching, is based upon the Vygotskian notion of a "zone of proximal development" (Vygotsky, 1978). Vygotsky suggested that learning is a social situation in which a more experienced learner models activities and gradually leads a novice learner to a higher level of performance. With time and practice, the novice becomes capable of performing the learning activity on his or her own. In this way, the novice learner is lead to the furthest, comfortable reach of his or her potential through the use of expert scaffolding provided by the more experienced learner. The interaction and socialization process between the novice and expert learners is critical in guiding the novice learner to a desired level of performance. The novice learner participates in the group activity at a level of comfort, observing and modeling an expert learner, while finding support and encouragement through the social context of the situation (Palinscar & Brown, 1984).

Brown and Palinscar incorporated the Vygotskian notion of learning into their studies by creating a model of instruction that is based on the interactive principles of Vygotsky, but also includes instruction in specific strategies designed to improve academic performance. Brown's initial work dealt with the student's ability to learn from texts (Brown & Campione, 1981). From there, her work with Palinscar and others evolved into studying what commonalities could be found among the activities and strategies that successful readers employed while engaged in reading for various purposes (Brown, Palinscar, & Armbruster, 1984). From these studies, the authors selected four strategies that have the dual function of enhancing comprehension and at the same time providing a self-monitoring function for the learner. Brown and Palinscar reported that these strategies when utilized are both comprehension-fostering and comprehension-monitoring activities.

As a result of her review of successful training studies, Brown decided that the four reading strategies must be taught as part of an interactive procedure that allowed students to participate at a level at which they were capable, forced them to be active, provided feedback on their performance, and included instruction in applying the strategies (Brown, Day, & Jones, 1983). It is here that Brown linked her intervention of strategy training components with Vygotsky's notion of allowing students to participate at their level of comfort, with the more experienced learner providing guidance to assist the more novice learner to a greater level of performance.

Brown's attempt to design an instructional method that contained the best of other successful strategy training studies with the interactive instruction proposed by Vygotsky evolved into the reciprocal teaching method. Reciprocal teaching is comprised of the following three components:

1) instruction in specific strategies designed to enhance comprehension-fostering and comprehension-monitoring,
2) interactive instruction that employs the expert scaffolding of a more experienced learner guiding a novice learner, and
3) a cooperative learning environment in which learners support and help each other to reach their potential level of performance.

In particular, what the reciprocal teaching method offers that other strategy training attempts have lacked, is an instructional component based upon empirically established learning and developmental principles coupled with a shared responsibility for learning among the novice and more experienced learners. This unique combination of expert scaffolding, concrete strategies, and cooperative learning has allowed Brown and Palinscar a measure of success unknown to most cognitive training studies.

Statement of the Problem

It is suggested that the reciprocal teaching method and other cognitive-training strategies that have successfully been used to foster reading comprehension can be applied to other subject areas such as mathematics (Brophy, 1986). It is the modeling process and the social context of reciprocal teaching that is of particular interest to mathematics researchers who are interested in teaching cognitive skills. Here again the work of Vygotsky is seen as providing guidance to the study of group learning as an important variable for research within the field of mathematics teaching and learning (Silver, 1985).

The Study

Purpose

The intent of this study was to research the effects of the interactive teaching method, known as reciprocal teaching, upon learning when utilized in a mathematics classroom at a junior high school. More specifically the study was designed to focus upon the effects the comprehensive-fostering and comprehensive-monitoring components of reciprocal teaching have upon learning when used in the particular content domain of geometry.

It was assumed that the reciprocal teaching method would be effective in enhancing comprehension in other content areas that require reading, such as social studies and science. With this in mind, the topic of geometry was selected as appropriate for this study of reciprocal teaching within the field of mathematics for it depends heavily upon distinctive vocabulary, much terminology, and the ability to see relationships. Geometry would not be as limiting to reciprocal teaching as other topic areas in mathematics which rely heavily on a student's background in numerical understanding, computational skills, or use of various algorithms.

Definition of Terms

For purposes of this study, the following operational definitions were used to clarify the differing terminology employed by various researchers to describe similar concepts:

1) cooperative learning - the planned process of structuring learning processes cooperatively rather than competitively or individualistically (Johnson & Johnson, 1984; Slavin, 1983).

2) scaffolded instruction - an interactive teaching method in which a more experienced learner guides a novice learner into a higher level of academic performance through appropriate modeling characterized by mutual responsibility for the learning experience (Vygotsky, 1978; Palinscar & Brown, 1984; Palinscar, 1986).

3) strategies - specific activities or routines in which a learner may engage to foster retention, academic growth, acquisition of knowledge, or performance of a specific task.

4) reciprocal teaching - an interactive process of teaching that combines cooperative learning, scaffolded instruction, and the training of specific comprehension-fostering and comprehension-monitoring strategies as an intervention to improve understanding and performance when comprehending and studying texts (Baker & Brown, 1984; Palinscar & Brown, 1984).

Theoretical Framework

As much as possible, the intent of this study was to stay true to the format of reciprocal teaching when applying it to the teaching of mathematics in the classroom of a junior high school. For this reason, the same components used in previous studies of the effects of reciprocal teaching upon reading comprehension were utilized. The study centered on teaching mathematics through the process of scaffolded instruction, modeling and employing the four comprehension-fostering and comprehension-monitoring strategies used by Palinscar & Brown (1984), and included a cooperative learning component that created a social context similar to that of other studies.

It is argued that research in the field of metacognition and cognitive-training studies, in particular, have allowed us to confidently train students within the settings of a normal classroom to comprehend and learn from text (Baker & Brown, 1984). Further, the success of these studies within the field of reading research must now be extended to the training of cognitive skills found in domains other than reading. For example, studies are needed that attempt to apply strategies and methods found useful in processing information from texts to processing information received through a lecture (Garofalo & Lester, 1985). Thus, this study examined the effects of the reciprocal teaching method and its effect upon the processing of mathematical information found in texts and lectures when used in a normal classroom setting.

Research Questions

The actual null hypotheses used in the statistical analyses of this study are listed and further addressed in Chapter 3. In an effort to avoid any redundancy, the null hypotheses will not be repeated in this chapter. With this in mind, let it be known that qualitative and quantitative analyses employed to determine the results of this study focused on the following research questions:

1) Will instructing the students in the reciprocal teaching method enhance the students learning of geometric concepts as measured by various assessment instruments?

2) Does the reading-comprehension level of students become a factor when utilizing the reciprocal teaching method in the content domain of geometry?

3) What part, if any, does the training of metacognitive strategies have upon learning geometric content in this reciprocal teaching mode?

4) Is previous student achievement in mathematics a determiner of success when employing the reciprocal teaching method?

5) Will students grow in their use of cooperative behaviors when encouraged to do so as a component of the reciprocal teaching process?

Significance of the Study

This study was important for two distinct reasons. First, the study attempted to blend the success of cognitive training in the area of reading research into the domain of mathematics teaching and learning. If the positive effects of reciprocal teaching upon reading comprehension were found to transfer to mathematics learning, then the reciprocal teaching method would solidify a place in the field of cognitive-training research. It would seem that the potential of pragmatic benefits when applying the reciprocal teaching method to the normal classroom setting merited strong consideration for the continued study and use of this intervention method.

Secondly, and perhaps more importantly, if the use of reciprocal teaching in learning from a lecture or discussion in the field of mathematics yielded significant results, then a paradigm for studying the role of metacognition upon mathematics teaching and learning could be established. This could bridge the fields of reading and mathematics education by fostering further communication among studies and research currently conducted by various educators, psychologists, cognitive scientists, and classroom researchers.

Research Procedures

Selection of Subjects

The students participating in this study attend a public junior high school in a suburb of Chicago. The student body basically reflects the community, a predominantly white, middle-class suburb with less than ten percent of the population being minority students. The junior high school has a student body composed of sixth, seventh, and eighth grades.

The students are tracked into mathematics classes according to their ability level and past academic performance. Three distinct programs of instruction are available to students resulting in classes designed for remedial, average, and accelerated students. Approximately seventy-five students comprising three sections of average mathematics classes took part in the study. A further discussion and analyses of selection criteria for determination of students participating in the study is found in Chapter 3.

Procedures

The reciprocal teaching method was utilized in this study by employing the three distinct components of scaffolded instruction, concrete strategy-training, and cooperative learning as part of the daily instruction of the students in their regularly scheduled mathematics classes. The natural setting and parameters of the school day were not altered in terms of the time alloted for instruction, the length of a class period, or the number of class sessions addressing the topic of geometry.

Working under these conditions, the use of the reciprocal teaching method as intervention was utilized for twenty class days. Each class session lasted for fifty minutes as predetermined by the students normal schedule of classes. The class period was divided into two sections, with twenty-five minutes appropriated for teacher directed instruction, and twenty-five minutes remaining for students to engage in a cooperative learning situation with other students.

For purposes of this study, the twenty-five minutes alloted for teacher directed instruction combined the two, reciprocal-teaching components of scaffolded instruction and the training of concrete strategies. The teacher acted as a model to guide the students into an ever greater level of participation in the lesson, while giving explicit instruction in use of the strategies to comprehend and learn the geometric content presented.

Following the lesson each day, students engaged in a cooperative learning situation through working in small groups of three to four students. This cooperative learning component constituted the third and final facet of reciprocal teaching as used in this study.

Twenty lesson plans were developed and used with all students participating in the study. The lesson objectives and geometric content did not differ across treatment conditions. The same teacher, this researcher, instructed all three of the classes involved in the study.

A further discussion of treatment conditions, control and experimental groups, and a more complete design of the study is found in Chapter 3. Discussion of questions relating to the internal validity of the study are included in Chapter 4.

Analyses of Data

Both quantitative and qualitative dependent measures are reported in Chapter 4 of the study. Dependent measures of a quantitative nature include standardized achievement test scores, standardized criterion-referenced test scores, criterion-referenced scores, standardized ability test scores, and a recognized attitudinal measure of affect. Data produced by the administration of the quantitative dependent measures were analyzed through analysis of variance and analysis of covariance. This data was also reported in terms of frequencies and percentages.

Qualitative measures included observation of student behaviors in a coded protocol system. This measure was reported in frequencies and percentages with comments regarding any patterns that may emerge as part of the data analysis.

Limitations of the Study

The study occurred in the natural setting of an actual school day. As a research study, this field experiment fits the definition of a quasi-experimental design in which experiments have treatment conditions and outcome measures, but do not use randomization in the selection of subjects (Cook & Campbell, 1979). For this reason the threats to the internal validity of the study must be established and addressed before any results or implications of the study are explored or interpreted.

Since the study did use intact groups as experimental units it would have a strong transfer or generalizability component to other research if the issues of internal validity are first put to rest. The potential for pragmatic application and addition to our knowledge base through field experimentation still merits serious consideration for the continuation of quasi-experimental designs in the study of cognitive training methods.

Another concern that may limit the results of the study is the unintentional bias effect that may be present when utilizing the same teacher across differing treatment conditions. This issue is further discussed in Chapters 3 and 4 of the study.

Organization of the Dissertation

Chapter 1 has established an organizational framework that briefly outlines the intent of the study and some considerations that should be addressed. In Chapter 2 a more complete review of the literature and of the research studies that form the foundation of this study is provided.

The pàradigm of the study is detailed in Chapter 3, with considerable emphasis placed upon design components and methodology chosen for this experiment. Chapter 4 relates the data analyses of both quantitative and qualitative measures. The concerns and issues regarding internal and external validity, reliability and measurment are thoroughly articulated and evaluated in both Chapters 3 and 4.

Chapter 5 includes a brief summary of the experiment along with a commentary on the implication of the results of the study. Some suggestions for future studies in the areas of cognitive training, metacognition and mathematics teaching and learning are also found in the final chapter. The Table of Contents provides further information regarding figures, tables, appendices and references included in the study.

Summary

This was a field experiment that studied the relative effects of the reciprocal teaching method upon mathematics learning. This interactive teaching method, found useful in fostering reading comprehension and comprehension-monitoring, was applied to mathematics instruction in the content area of geometry at a junior high school. The study attempted to blend the studies and experimental methods used by cognitive scientists, developmental psychologists, and researchers from the fields of mathematics and reading, to reaffirm previous research findings and establish a model for further research.

CHAPTER II

RELATED STUDIES

In recent years educators from across the various disciplines have spent an ever increasing amount of time developing, implementing, and studying the effect that the use of learning strategies have upon instruction. Researchers interested in the field of teaching are no longer content with merely imparting knowledge, but are also interested in fostering the cognitive skills that enable students to increase their learning potential (McKeachie, Pintrich, & Lin, 1985). This learning potential or capacity issue has often resulted in instructional designs that address thinking skills, problem solving and learning to learn (Chipman & Segal, 1985; Weinstein & Underwood, 1985).

In this chapter, a brief look at learning theory from the perspective of cognitive psychology is presented. A systematic review of current learning strategy research follows, along with a selective discussion of possible applications and limitations for instruction. The search for metacognitive strategies is documented, as well as a discussion of the development of the interactive-teaching method known as reciprocal teaching. A particular emphasis is given to establishing the broad theoretical underpinnings of reciprocal teaching, some shared or similar perspectives from various domains of research, and the pragmatics of applying these studies to teaching.

A Cognitive Theory of Instruction

One recurring theme found in much of the literature from cognitive science is that people construct rather than receive information or knowledge (Resnick, 1984). Central to this notion of constructing knowledge is the idea of memory storage. Cognitive scientists believe that a learner takes in new information and must restructure it by relating this new information to prior information for memory storage. In this way the learner must actively take part in the learning situation if new knowledge is to be remembered or retained for future use. This active engagement on the part of the learner is generally accepted as an internal process, but may have some relation to overt behaviors a student performs in a particular learning situation (Wittrock, 1978).

Related to the notion of reconstruction described above, is the role that one's prior knowledge plays in a learning situation. Given that new information must be taken in, reconstructed from the learner's perspective, and related to prior information for memory storage; it becomes evident that prior knowledge becomes a foundation upon which further learning or retention is based. Several broad categories of prior knowledge are generally recognized in the literature including domain-specific knowledge, general world knowledge, and the knowledge necesary to interpret written symbols portrayed in texts (Campione & Armbruster, 1985; Resnick ,1984).

The focus of many of the studies that follow this research perspective has been on the processes or skills involved in coordinating prior knowledge and new information found in texts. The processes that a learner may choose to use when combining new information with prior knowledge have often been described as "strategies". The use of these "strategies" by a learner implies that one can control these processes, and that they may have some impact upon instruction worthy of further study by psychologists and educators (Resnick, 1984).

Learning Strategies

Weinstein and Mayer (1986) explain that learners are now seen as active participants in the learning and teaching process. This shared responsibility for learning acknowledges that learners must build upon their prior knowledge when processing information, and that student behaviors and thoughts influence the effectiveness of instruction. They go on to say that good teaching includes teaching students ways to think, remember, monitor their learning, motivate themselves, and, in the end, learn how to learn.

This objective of lifelong learning is attained by utilizing learning strategies to aid in information-processing. Weinstein and Mayer (1986) have produced a taxonomy for learning strategies, which capsulizes many of the previous studies. The authors propose eight categories of learning strategies, all with differing purposes, that vary to meet the context of a learning situation. Figure 1 contains the list of learning strategy categories used by Weinstein and Mayer, and includes an example of a learning strategy activity for each category.

Figure 1. A taxonomy of learning strategies (Weinstein & Mayer, 1986).

Strategy Category	Related Example
1) basic rehearsal	rotely repeating words to memory
2) complex rehearsal	underlining important terms while reading
3) basic elaboration	forming a mental image while listening
4) complex elaboration	paraphrasing a passage recently read
5) basic organizational	categorizing information
6) complex organizational	creating a hierarchy or flow chart
7) comprehension-monitoring	self-questioning for meaning
8) affective and motivational	using positive self-talk for relaxation

The above-mentioned taxonomy of learning strategies exemplifies some of the general techniques that learners may use to aid in the processing of information. The strategies range from mundane to sophisticated, and address memory storage, acquisition of knowledge, and even some of the constraints of learning. Use of this taxonomy provides a means for researchers from varying domains to compare learning strategies found in the extant literature. Dansereau (1985) agrees that individual capacity for acquiring and using information can be significantly enhanced by training learners to utilize information-processing strategies.

Weinstein and Mayer (1986) believe that the use of learning strategies in instructional programs will result in a useful data base for the continued study of classroom teaching and educational practice. The authors do caution, however, that the scissors have two blades; strategy instruction will not replace the role of subject matter knowledge in learning. Learning strategies are not a substitute for teaching domain specific knowledge, and at best are equallly important for effective instruction. This emphasizes the role of prior knowledge in the learning situation, even the concept of learning to learn is still anchored by a foundation of knowledge and information that functions as a prerequisite for further learning.

Learning Mathematics

Though much of the theory described above was generated by cognitive scientists or developmental psychologists, the practical applications for a teaching or learning situation are readily seen. Silver (1985) explains that devlopmental psychologists are no longer just interested in general learning and thinking, but are now concentrating their study and research to specific subject domains such as mathematics. Along with this has come a reciprocal interest in cognitive psychology among mathematics educators (Lester, 1982; Resnick and Ford, 1981).

In particular, Silver (1985) is concerned with the effect that cognitive psychology has upon the study of problem solving in mathematics. In his paper he relates that teachers are confronted with much literature about problem solving and learning strategies, but are given very little of the research base necessary to confirm or deny the research findings. He hopes that by sharing research findings from problem solving in mathematics with other domains an adequate research base for learning strategies will develop.

Learning to Read

It is evidenced that good readers differ from poor readers in their use and knowledge of learning strategies that aid in constructing meaning from text (Wittrock, 1978). Campione and Armbruster (1985) have identified three variables important to comprehending the meaning from a text: the material, the learner, and the chosen learning strategy. These variables are thought to interact as a function of the comprehension process and have been verified by other research (Brown, Campione, & Day, 1982; Brown, Palinscar, & Armbruster, 1984).

Although learning to read and learning from text can be thought of as difficult processes, two intervening variables have been cited as causes of poor comprehension and learning in students (Brown, Campione, & Day, 1982). First, prior knowledge, and second, inefficient use or selection of learning strategies are recognized as factors that effect the reading comprehension process of a learner. The authors recommend specific strategy training related to both the learner and the text to overcome these reading difficulties.

Summary of Learning Strategies Research

Although the last two decades of research have produced a plethora of learning strategies study, there are some commonalities that emerge from this abundant body of work. First, there are established and generally accepted studies that have generated strategies for use in various subject domains such as, but not limited to the following: reading comprehension, mathematics learning, problem solving, computer-assisted instruction, individual differences, special education, and thinking skills. These strategies are well-documented and easily accessed through many of the reviews of learning strategies (Pressley & Levin, 1983; Weinstein, 1986).

Secondly, though the research is voluminous, it has not decisively assured that learning strategies included in instruction will ever be used by students or transferred to other learning situations (Pressley, 1986). This dilemma of lack of transfer or generalizability has forced researchers to refocus their efforts from just simply identifying strategies to actually training students how and when to use strategies.

Metacognitive Strategies

As stated in the first chapter, attempts by researchers to develop more generalizable learning strategies have created the field of metacognition. Basically, the term "metacogniton" refers to the active monitoring and awareness a learner exhibits while engaged in a learning task (Brown, Armbruster, & Baker, 1986; Flavell, 1976; Silver, 1985). The authors note that though the term metacognition may be new, the field of metacognition is akin to much earlier research by Thorndike, Binet, and Dewey.

Though the content domains differ, it is clear that conscious attention to control and monitor one's learning occurs in all learning tasks. Learners engaged in any cognitive activity do have the potential to monitor their thinking. It is this potential to teach a learner to control his learning process that is so intriguing to teachers and researchers (Gerhard, 1987). The author also states that students can be taught how and when to use metacognitive strategies if they are appropriately modeled as part of the instructional process.

Reading and Metacognition

Brown, Armbruster, and Baker (1986) acknowledge that metacogniton plays a vital role in reading. They comment that successful readers have learned to monitor themselves while they are engaged in the reading process, and have developed specific strategies to aid in controlling their attention and focus as warranted. Variables previously cited as important in learning to read are again noted as important to the process of metacognition: the learner, the text, the task required by the learning situation, and the learning strategies.

In general, sucessful approaches to reading include metacognitive variables and vice versa (Baker & Brown, 1984). Self-regulatory behaviors normally associated with efficient readers meet the definiton of metacogntion. The current trend is to use the teacher as a model for learners in both self-regulation and task-specific stategies. Learners are no longer simply told what to do and left to complete a task on their own. According to the authors, this change reflects the Vygotskian notion of guided learning.

A recent quantitative synthesis of twenty studies purporting to assess the effects of metacognition on reading comprehension (Haller, Child, & Walberg, 1988) confirmed that metacognitive skills training does exhibit a positive effect upon reading comprehension. These studies contained various levels of training in the components of awareness, regulation, and monitoring of the reading process. In particular, the use of self-questioning was noted as being especially effective for readers in the seventh and eighth grades.

Metacognition and Mathematics Learning

Silver (1985) observed that psychologists and researchers from the field of mathematics share the same interest in mathematical problem solving. He states that purely cognitive explanations about successful problem solving are incomplete without a metacognitive component.

Lester and Garofalo (1985) in their article relating metacognition to mathematical performance note the success of the reading research as evidence that metacognitive strategy training is worthy of continued research in mathematics. In particular the Brown and Palinscar (1982) studies regarding strategy training in reading are of particular interest to the authors.

Numerous research models have been created to study the role of metacognition upon mathematics learning (Garofalo, 1987; Lester & Garofalo, 1985; Schoenfeld, 1983; Silver, 1985). It becomes apparent through closer analysis that many of these models are based upon two pivotal works. First, the classic problem solving model of Polya (1957) with the four steps of understanding, planning, carrying-out, and looking back has often been used to create the more recent models. Second, the more recent reading research by Brown et al. (1982, 1984a, 1984b, & 1986) was utilized to substantiate the research base and as a foundation for development of models to study mathematical performance.

The trend toward continued research in metacognition and mathematics is clear. It would benefit the field of mathematics to capitalize on research findings in this area from both the domains of psychology and reading.

The Vygotskian Perspective

A number of the metacognitive and learning strategy studies previously noted have been influenced by Vygotsky (1976) and his notion of a zone of proximal development. This concept is based upon a significant other leading a learner to a level of performance that would otherwise be unobtainable. This new level of performance is achieved as a result of modeling the desired performance. The learner and the teacher develop a relationship through the learning task that mutually binds them in responsibility for the desired learning or performance. Social interaction between the person serving as the model and the learner was cited by Vygotsky as an important factor that fosters cognitive growth in the learner. The learner is encouraged to gradually develop more responsibility and control over the learning situation, while the person modeling eventually relinquishes control.

This modeling process for learning based upon mutual responsibility of both learner and instructor is often described in learning strategy (Chipman & Segal, 1985) and reading research conducted by Brown et al. (1982, 1984a, 1984b, 1986). Thus the Vygotskian notion of proximal development is often the cornerstone for research in strategy training. The work of Vygotsky also relates to the following section of this chapter though it is not often noted by the particular researchers.

Cooperative Learning

Instructional methods and academic tasks that require students to work together are often grouped under the heading of cooperative learning (Slavin, 1987). Studies that have used cooperative learning report that students increase their achievement levels, express positive attitudes about learning, and exhibit competency in collaborative skills (Johnson & Johnson, 1984).

Two basic models of cooperative learning appear in the extant literature. First, Slavin (1983, 1987) proposes a model that is based upon group contingencies, rewards, motivation, and the teaming of four through six students. Second, Johnson and Johnson (1984) have created a model of cooperative learning which stresses the individual responsibility of students through a positive interdependence among group members. These two approaches to cooperative learning differ significantly in the areas of motivation, rewards and group contingencies.

Slavin (1983, 1987) believes that cooperative learning methods require group rewards for the individual in a learning situation to be motivated and successful. He believes that group competition is necessary to increase the instructional effectiveness of the group learning task. His model requires a reward structure based upon the total group performance in lieu of any rewards for individual efforts by a student within a group.

The Johnson and Johnson model of cooperative learning (1984) neither encourages nor requires group competition. Though it, too, concentrates on helping behaviors among the group members, constant competition is viewed as detrimental to the process of positive interdependence (Johnson, Skon, & Johnson, 1980; Johnson, Johnson, & Stanne, 1986; Lew, Mesch, Johnson, & Johnson, 1986). Four basic components of the Johnson model for cooperative learning are: positive interdependence, individual accountability, face to face interaction, and cooperative skills.

Some salient aspects of role of the teacher as defined by the Johnson and Johnson model (1984) include: 1) specifying the learning objective, 2) making sound decisions about grouping, 3) explaining the tasks to the students, 4) monitoring the cooperative groups, 5) increasing the collaborative skills of the students, and 6) evaluating the effect of the cooperative grouping. It should be noted that the Johnson brothers believe that students can be taught collaborative skills in conjunction with academic instruction of specific content material.

Social Interaction

Critics of the cooperative learning research often note a lack of evidence that students participating in a study actually did or did not act cooperatively (Webb, 1982). Proponents of this issue stress that students must be observed as exhibiting certain predetermined behaviors that reflect the varying treatment conditions if the outcome measures are to be believed.

Categories of behaviors normally associated with the collaborative skills of cooperative learning are often termed helping behaviors (Webb, 1980, 1982a, 1982b, & 1982c). Off-task and passive behaviors such as working alone form a dichotomous category to the helping behaviors. Though the topics of peer interaction and small versus large group learning have been extensively researched, Webb stresses the need for a thorough accounting of social interaction in all cooperative learning studies. Failure to verify the actual behaviors of students in varying treatment conditions of cooperative learning jeopardizes the internal validity of the research studies. This would nullify or limit the generalizability of the findings of the studies to other groups and settings.

Reciprocal Teaching

Palinscar and Brown (1984, 1986) describe three training studies they conducted to enhance a student's ability to learn from text. The studies utilized specific strategies designed to aid in the comprehension-fostering and comprehension-monitoring activities of the students. Their model of instructing students in the use of these comprehension-fostering and comprehension monitioring strategies is known as reciprocal teaching.

Reciprocal teaching has three basic components that are soundly based upon previous research in both the fields of cognitive psychology and reading. The model is unique in that it pragmatically combines the research findings from these two disciplines to enhance classroom learning. The three components of the reciprocal teaching model are: 1) expert modeling by the teacher of sound learning with ample time provided for student practice, 2) specific comprehension-fostering and comprehension-monitoring strategies to enhance learning, and 3) social interaction among the students and the teacher.

The first component of the reciprocal teaching model, expert modeling by the teacher, is based upon the Vygotskian notion of proximal development (1976). As previously described, a learner is guided to a higher level of performance in a learning situation through explicit modeling of strategies and techniques that foster academic growth. Gradually the learner is encouraged to take on a greater role in this interaction between learner and teacher through guidance and success while engaged in the learning situation. The responsibility for the learning situation is shared by the teacher and the learner.

Figure 2. Common traits of successful readers.

Studies have shown that successful readers:

1) understand the various purposes of reading,
2) use their prior background knowledge,
3) allocate their attention,
4) find a level of compatability between the reading material and their prior knowledge and experience,
5) monitor themselves while they read, and
6) draw and test inferences (Brown, Palinscar, & Armbruster, 1984).

Specific comprehension-fostering and comprehension-monitoring activities are taught through the reciprocal teaching process. These comprehension activities are learning strategies that comprise the second component of reciprocal teaching. In their previous work, Brown, Palinscar and Armbruster (1984) summarized six categories of practices that were found common to various studies of reading comprehension. These six commonalities pertaining to successful reading are found above in Figure 2.

These six common practices utilized by successful readers were synthesized by Palinscar and Brown (1984) into four specific and trainable learning strategies. The four strategies coined by the authors and comprising the second component of the reciprocal teaching are: 1) summarizing (self-review), 2) clarifying, 3) questioning, and 4) predicting. These four strategies were selected because they provide the dual function of comprehension-fostering and comprehension-monitoring when properly used.

The third component of reciprocal teaching evolved from successful training studies that forced the students to be active, provided feedback, and taught the students when and how to use specific learning strategies (Palinscar & Brown, 1984). Thus, an interactive training component was incorporated into the reciprocal teaching model that is similar to the Johnson and Johnson model of cooperative learning (1983) that demands a positive interdependence among the students and the teacher.

The basic procedure of reciprocal teaching begins with an adult modeling or instructing the use of a specific comprehension-fostering or comprehension-monitoring technique. Students are encouraged to particpate in the lesson by taking turns in modeling the activity within the group setting. Gradually, students become more adept at taking the leadership role in the learning activity.

Palinscar and Brown (1984, 1986) note that changes in the dialogue patterns of students engaged in the reciprocal teaching process are evidenced as the learner gradually assumes this leadership role. Palinscar (1986) states that these dialogue changes are also observed in the works of Vygotsky. This clearly establishes a link between the research of Vygotsky and reciprocal teaching.

Summary of Reciprocal Teaching Research

Palinscar and Brown (1984, 1986) confirmed that their interactive teaching process exhibited results that were significant and reliable. Problems that plagued other researchers conducting studies based on learning strategies were overcome by their unique combination of previous research findings with a level of specificity lacking in the works of others. Another distinguishable factor of their work was that their findings were generalizable to other research fields. Their efforts in the field of reading created a model of research that can be transferred to other content domains in actual field settings.

Mathematics Research Trends

Brophy (1986) authored an article that suggested areas to be further researched within the field of mathematics learning. The author noted that teaching cognitive skills was an area that should be explored. The works of Palinscar and Brown (1984) were cited by Brophy as studies of reading comprehension strategies that could be applied to mathematics teaching and learning. These findings are not only suggested as useful to the topic of mathematics problem solving, but to all subfields of mathematics that call for strategy training.

Brophy suggests that strategy training programs must include instruction in specific strategies and skills similar to the reading research (e.g. summarize, question, clarify and predict). Components such as metacognitive strategies, the use of prior knowledge, modeling by a teacher in strategy usage, and an increased level of student activity and social interaction while in the learning situation are all recommended areas for use and research within the field of mathematics.

Silver (1985) proposes that mathematics researchers should consider study in the areas of small group processes and cooperative learning. He notes both the Slavin (1980) and Johnson and Johnson (1983) models of cooperative learning to be of particular interest within the mathematics community.

The author agrees with Webb (1982) that although small group learning is used and often encouraged in the real world setting of the classroom, while little evidence exists that verifies or confirms the effects of these processes upon learning. Research in the area of social interaction may provide more information about which conditions of small group processes positively effect and promote learning. It is interesting to note that Silver encourages further knowledge about the works of Vygotsky (1976) as a source of some guiding principles that may be applied to learning and instruction in group settings. Noddings (1985) also recommends the theory of Vygotsky as a foundation for the creation of models designed to study mathematics learning in small group settings.

It is apparent that many of the recommendations for future research in the areas of mathematics teaching and learning are founded on recent studies in the areas of metacognition and reading comprehension. Both of these fields have provided specific strategies that promote learning, and a vehicle for instructing students in when and how to utilize these learning strategies. It seems that a bridge across the content domains of reading and mathematics research would foster both content domains. The area of cognitive psychology provides that bridge and is discussed in the final section of this chapter.

Chapter Summary

This chapter began with a brief description of a cognitive view of learning, followed by a section on learning strategies research. Specific attention was given to learning mathematics and learning to read as reported in the learning strategies research.

An analysis of the findings indicated that a cognitive view of instruction proposes that students learn best when they are: 1) actively engaged, 2) use their prior knowledge in linking new information to old information to aid in memory storage, 3) find personal meaning in their learning, 4) utilize learning strategies to aid in the processing and retention of information, and 5) use models to foster cognitive growth (e.g. concrete manipulatives, a more experienced learner, a teacher).

A definition of metacognition was provided, which again was contrasted to the fields of reading and mathematics learning. It was noted that successful readers routinely practiced and used metacognitive learning strategies while engaged in the reading process.

The Vygotskian perspective of instruction was compared to components normally associated with cooperative learning. A discussion regarding the limitations of cooperative learning resulted in pronounced need for verification of student behaviors while engaged in activities across group settings.

Finally, the reciprocal teaching approach was described as a model of instruction that combined the areas of strategy training, metacognition and social interaction to enhance learning from text. This model was proposed as a means to study mathematics teaching and learning in group settings.

CHAPTER III

PROCEDURAL METHODS

The three components of the reciprocal teaching intervention represent the independent variables of this study. As defined by Brown & Palinscar (1984), these three components have been operationally defined as scaffolded instruction, the training of concrete strategies, and cooperative learning. A discussion of concerns and related issues in adapting the reciprocal teaching method to the instruction of geometry in a mathematics classroom is warranted.

Independent Variables

As previously stated, this study applied the reciprocal teaching method to the teaching of mathematics in a normal classroom setting. To facilitate this process and yet provide clear assessment of the treatment, required the separation of the components of the reciprocal teaching method into two distinct independent variables. For this study, the scaffolded instruction and concrete strategy components were treated as one independent variable that was basically reflected in the teaching of a geometry lesson by the teacher. The teacher incorporated the same four comprehension-fostering and comprehension -monitoring strategies used in previous reciprocal teaching studies into the presentation format of the mathematics lesson.

The four strategies of summarizing, predicting, clarifying and questioning were modeled by the teacher through the process of scaffolded instruction, with students being encouraged to participate in the lesson and discussion at their level of comfort. This independent variable which combined two components from the reciprocal teaching model was termed "metacognitive strategies" for purposes of this study.

The remaining component of cooperative learning stood alone in the experiment as a second independent variable. Following each lesson, selected students engaged in cooperative group learning situations with other students as a part of the regularly scheduled mathematics class. The use of this small group learning in which students are encouraged to share and exchange information as a follow-up to strategy training represented the second independent variable of this study. This second independent variable was termed "social interaction."

Design of the Study

In order to determine the relative effects of the two independent variables selected for the study a design that reflects four treatment conditions was proposed. The four treatment conditions being: 1) an experimental group of students who receive the variable metacognitive strategies, 2) a control group of students who do not receive the metacognitive strategies variable, 3) an experimental group of students who engage in social interaction, and 4) a control group of students who do not engage in social interaction. Figure 3 shows the research model of the four treatment conditions establishing a design for the experiment based upon the two independent variables of metacognitive strategies and social interaction.

Figure 3. Design of this study.

		Metacognitive Strategies	
		YES	NO
Social Interaction	YES	I	II
	NO	III	IV

Selection of Groups

The study as proposed required the use of four groups to properly complete the experiment. As previously stated, this study employed a quasi-experimental approach in which treatment conditions and outcome measures exist without the random assignment of subjects for selection in the experiment (Cook & Campbell, 1979). Subjects were not randomly assigned in this study since the experiment occurred in the natural setting of a junior high school which had previously established intact groups.

Since randomization was not used to counteract the effects of any differences that may have existed among the groups, comparability of intact groups selected for participation had to be clearly established prior to any intervention taking place. This researcher noted that a lack of comparability among intact groups would seriously threaten the internal validity of the study, thereby nullify any chance of expressing any externally valid findings.

Correlated Group Design

Though the use of four intact groups would have provided a clear design for the study, only three comparable intact groups were found within the natural setting of the junior high school for use in the study. Chapter 4 contains a more thorough discussion of the statistical analyses of standardized achievement and standardized criterion-referenced test scores for six intact groups eligible for participation in the experiment. Not all six groups, however, were able to be used in the study due to performance differences that existed prior to any treatment.

Figure 4. Correlated group design.

Methods (treatment conditions)	Units (intact groups) A1	A2	A3
1			
2		Y means or	
3		dependent measures	
4			

An analysis of variance procedure performed on the above mentioned pretreatment measures verified that three intact groups did not differ significantly and were selected for inclusion in the study. The intact groups represented a correlated group design (Kerlinger, 1973) without randomization or matching. Conceptually, the use of three intact groups in this study met the criteria for quasi-experimentation (Cook & Campbell, 1979). The actual research design used in the study is shown above in Figure 4.

Sampling Discussion

Since only three intact groups were found to have sufficient similarites in measures to represent true correlated groups, then some adjustment was necessary to adapt the sample population to the proposed four-group design. The three intact groups yielded a sample size of N=73. When spread among the four treatment conditions the three intact groups resulted in unequal cell numbers, but did establish a sufficient sample to complete the study of the two independent variables as proposed.

Of the three intact groups selected for the study, two were seventh grade classess, while one was a sixth grade class. The groups were not found to significantly differ on the pretreatment measures, although the groups were from different grade levels.

In order to achieve four distinct groups for this experiment, some manipulation of the intact groups or treatment conditions was necessary. Through randomization, two of the three intact groups were selected to receive the metacognitive strategies treatment, while the third intact group did not receive the metacognitive strategies treatment. The metacognitive strategies group was comprised of one of the two seventh grade classes eligible for participation in the study, and also the sixth grade class.

Among groups, two intact groups (N = 49) received the metacognitive strategies treatment, while one intact group (N = 23) did not. Thus, two of the four treatment conditions necessary to conduct this study were established.

To achieve the remaining treatment conditions, a systematic matching technique was used to determine which respondents within a group would be selected for the social interaction treatment. For example, the 25 students comprising an intact group needed to be broken into yet two more distinctive groups within that treatment condition to establish the third and fourth treatment condition. It was decided to match the individuals within the groups on previous performance in the mathematics classroom by utilizing the most recent report card grades. These report card grades were based on criterion-referenced measures for this particular school district. Prior report card grades in mathematics were chosen as a predictor of future performance in the mathematics classroom. The natural setting of the junior high school classroom that existed prior to the introduction of any treatment variables was not disrupted or altered in the areas of the grouping of students, selection of a teacher, or required curriculum.

The systematic matching of students within each of the three intact groups began with the name of the student being assigned to one of the five various grade categories of A - E based on the most recent report card grade for mathematics. Once the names of the students were sorted into the five grade categories, a name of a student was randomly selected from within the A category of the intact group. The selection was placed in one of the two remaining treatment conditions of social interaction or no social interaction. Assignment to one of these two treatment conditions was alternated on every succeeding selection. If an odd number of students comprised the A category, the last remaining name of the student was placed in the B category with the names of other students for random selection. This process was continued until all grade categories A - E were randomly selected and placed in one of the two treatment conditions.

This random assignment of students was based upon a pretreatment variable, a previous report card grade in mathematics. Students from within the three intact groups were assigned to the two remaining treatment conditions through an established randomized-matching technique (Kerlinger, 1973). The selection of students from within the three intact groups into differing treatment conditions was consistent with the correlated group design and methodology previously employed to randomly assign entire intact groups to treatment conditions.

Figure 5. Population and cell numbers for this study.

		Metacognitive Strategies		
		YES	NO	TOTALS
Social	YES	24	10	34
Interaction	NO	25	14	39
	TOTALS	49	24	73

The three intact groups were randomly assigned to one of two differing treatment conditions. Students within the three intact groups were matched on a pretreatment variable and randomly assigned to yet one of two other differing treatment conditions. The assignment of students within and among the three intact groups produced cell numbers and the four treatment conditions as seen in Figure 5.

Dependent Measures

Various dependent measures were selected to measure the relative effects of the intervention. A 50-item criterion-referenced test that represented a typical final was used to evaluate student performance in meeting the predetermined criteria or objectives of the curriculum. In this case, the geometry content of the mathematics lesson constituted the criteria for assessment expressed through learner objectives.

Criterion-referenced Measures

Three criterion-referenced tests were developed for use in the experiment. These three tests are referred to as short quizzes (CRT1, CRT2, & CRT3) consisting of ten, open-ended questions regarding specific geometric content. The development of these tests closely followed established procedures suggested by Popham and Husek (1969) to ensure the reliabilty and validity of these measures. These tests were administered sequentially and occurred as part of the intervention after the students had been exposed to the appropriate content. Copies of these criterion-referenced tests actually used in this study can be found in Appendix C.

Norm-referenced Measures

A criterion-referenced test that is fifty items in length and represents a final examination or posttest measure of the geometry unit introduced in the classes. A further discussion of the development and analyses of this criterion-referenced test as a norm-referenced dependent measure appears in Chapter 4. Issues relating to the validity and reliability (r = .86) of the geometry final test (GEOFIN) developed for this study are addressed.

A standardized, criterion-referenced test appropriate for junior high school students in assessing knowledge of geometric concepts was also used in the study. The Sequential Assessment of Mathematics Inventories (Reisman & Hutchinson, 1985) is a twenty-one item assessment instrument that was normed and has established validity, reliabilty (r = .82) and internal consistency data. This measure was used twice in the study as a component to the geometry final test (GEOFIN). The Sequential Assessment of Mathematics Inventories (SAMI I) was first used as a pretreatment covariate serving as a predictor of student performance on a prior knowledge construct. Second, this measure was also used as a posttest (SAMI II) and as part of a gainscore covariate (SAMI II - SAMI I). More information regarding the reliability and validity of this measure is provided in Chapter 4.

Affective Measures

Another dependent measure used in the experiment was a measure of affect. "The Confidence in Learning Mathematics Scale" of the Fennema-Sherman Mathematics Attitude Scales (1976) is a twelve-item measure with established reliabilty (r = .93). The scale is Likert in design, providing for a format that is sensitive to degrees of measure for this particular facet of mathematics attitude. Twelve statements, six positive and six negative, followed by a choice of five possible answers comprise the scale. The literature regards these Fennema-Sherman scales as a respected example of attitude measurement in mathematics.

Ability Measures

Yet another dependent measure used in the study was the "Abstract Reasoning" portion of the Differential Aptitude Test (Psychological Corporation, 1982a). This portion of the test is designed to measure the ability to understand ideas not presented in words. Rather this instrument uses mathematical relationships, diagrams, or designs. Accepting the definition of ability as the capacity to learn, this instrument was included in the study as a measure of transfer. It, too has established reliability (r =.91) and established content validity provided in greater detail in the following chapter.

A final dependent measure used in the experiment was an observational checklist of behaviors identified by Webb (1982a, 1982b, 1982c). Four, distinct behavioral categories have been formed to reflect topics found to be important as reported by the current literature on cooperative learning (Johnson & Johnson, 1984). This observational technique will be further explained in the "Instrumentation" and "Procedures" sections of this chapter, and will include a sample of the checklist.

Summary of Dependent Measures

Figue 6 provides a summary of the eight dependent measures recognized as dependent variables for this experiment. An acronym is established for each measure for reference in future figures and tables that appear in this paper.

Figure 6. Dependent measures of this study.

Dependent Measure	Abbreviation
Criterion-referenced Test #1	CRT1
Criterion-referenced Test #2	CRT2
Criterion-referenced Test #3	CRT3
Criterion-referenced Geometry Final Test	GEOFIN
Sequential Assessment of Mathematics Inventories	SAMI
Confidence in Learning Mathematics portion of the Fennema-Sherman Scales	CLM
Abstract Reasoning portion of the Differential Aptitude Test	DAT
Observer's Checklist	OC

Study Components

The parameters of the study previously mentioned in this paper consisted of two independent variables, four treatment conditions, and several dependent variables used in the experiment. Figure 7 summarizes the above-mentioned components of the study.

Procedures

The study components were included in the planning and teaching of the twenty class sessions that make-up the geometry unit in the mathematics curriculum. Again, class sessions were fifty minutes in length and were broken into two, twenty-five minute portions to create the treatment conditions. At the beginning of each class session, the teacher conducted a lesson that either included the metacognitive strategies variable or did not include the metacognitive strategies variable. When the lesson was completed students then moved to a treatment condition that utilized the social interaction variable, or to a treatment condiion that did not utilize the social interaction variable.

Figure 7. Components of this study.

Independent Variables	Treatment Conditions
	1) Metacognitive Strategies with Social Interaction (A & B)
A) Metacognitive Strategies	Social Interaction (A & B)
	2) Metacognitive Strategies without Social Interaction (A only)
B) Social Interaction	3) Social Interaction without Metacognitive Strategies (B only)
	4) No Social Interaction and No Metacognitive Strategies (control)

Twenty lesson plans were developed to ensure that the geometric content of the lesson and all learner objectives were identical for the differing treatment conditions. Students in all twenty class sessions used the same texts and instructional materials for classwork, and received the same homework assignments.

The first portion of each class session was teacher directed lasting for twenty-five minutes. Lesson plans differed with respect to the behaviors that the teacher modeled to create the metacognitive strategies conditions consistent with treatment designs.

The second portion of each class session, again lasting twenty-five minutes focused on the students. Participants were given assignments relating to the geometric content of the lessons. Students in the social interaction condition worked cooperatively to complete the assigned tasks. Students not in the social interaction condition were required to work individually on the assigned tasks with only the teacher available for help as needed.

As previously mentioned, the three intact groups participating in the study were randomly matched within their respective groups to create the four treatment conditions. Only one of the three intact groups did not receive the metacognitive strategies component (N = 24). Instruction in this condition did not include guided practice in the use of the four comprehension-fostering and comprehension-monitoring strategies of summarizing, clarifying, predicting, and questioning.

All of the class sessions that occurred as part of the treatment conditions in the study were videotaped. The videotapes served two distinct functions. First, videotaping of the lesson portion of the class session allowed an outside observer to view the tapes and comment on any teacher behavioral differences that may appear across treatment conditions.

Secondly, the videotapes also provided a means for the "Observer's Checklist" to be employed to verify and establish behavioral patterns or differences among the students during the second portion of the lessson.

Instrumentation

The three criterion-referenced tests (CRT1, CRT2, & CRT3) used as dependent measures in the study were constructed to reflect the learner objectives associated with the geometry curriculum. An independent, outside evaluator who is also a trained mathematics educator has determined that the three, ten-item tests have content and criterion-referenced validity established through a process of matching the test items to corresponding learner objectives and teacher lesson plans. The fifty-item test, known as the geometry final test (GEOFIN) also underwent this same verification process and was also deemed to possess content and criterion-referenced validity for purposes of the study. The appendices contain more discrete information regarding the actual corresponding test items for these four measures.

A reliability score was not produced for the three, ten-item tests since too few items comprised these measures resulting in little variance and dispersion among the test scores. For this reason, it is accepted that the reliabilty of a criterion-referenced test can not be addressed in the same fashion that one might address the reliability of a normative test (Popham & Husek, 1969) due to a difference in data relative to the varying test formats. Therefore, a distinct criterion-referenced data analysis procedure was utilized and documented in the "Quantitative Analyses" section of this chapter, and further detailed in Chapter 4. The procedure used and issues associated with obtaining a reliability score for the criterion-referenced geometry final used as a norm-referenced dependent measure, are also discussed in the following chapter of this study.

All dependent measures cited have established validity and reliabilty levels that are generally regarded as sufficient documentation to merit their use in the study when used properly. No other instruments were created for use in the experiment by the researcher.

Null Hypotheses

The following null hypotheses have been formulated for use in the experiment:

HO1: There will be no significant differences in criterion-referenced measures (CRT1, CRT2, & CRT3) across methods of metacognitve strategies.

HO2: There will be no significant differences in criterion-referenced measures (CRT1, CRT2, & CRT3) across methods of social interaction.

HO3: There will be no significant interaction between methods of metacognitive strategies and social interaction upon criterion-referenced measures (CRT1, CRT2, and CRT3).

HO4: There will be no significant differences in norm-referenced measures (GEOFIN) across methods of metacognitive strategies.

HO5: There will be no significant differences in norm-referenced measures (GEOFIN) across methods of social interaction.

HO6: There will be no significant interaction between methods of metacognitive strategies and social interaction upon norm-referenced measures (GEOFIN).

HO7: There will be no significant differences in measures of affect (CLM) across methods of metacognitive strategies.

HO8: There will be no significant differences in measures of affect (CLM) across methods of social interaction.

HO9: There will be no significant interaction between methods of metacognitive strategies and social interaction upon measures of affect (CLM).

HO10: There will be no significant differences in measures of abstract reasoning ability (DAT) across methods of metacognitive strategies.

HO11: There will be no significant differences in measures of abstract reasoning ability (DAT) across methods of social interaction.

HO12: There will be no significant interaction between methods of metacognitive strategies and social interaction upon measures of abstract reasoning ability (DAT).

HO13: There will be no significant differences in measures of observed student behaviors (OC) across methods of social interaction.

Quantitative Data Analyses

Criterion-referenced Measures

An innovative approach, due to Popham & Husek (1969), was used to display and analyze the data produced by the criterion-referenced measures. The three, ten-item measures were expressed in terms of nominal data and reported as frequencies and percentages. For example, if fifty percent of the students met a criterion-level of ninety percent or higher on the test of dependent measure, then 50/90 was reported. The frequency and number of students who met the preset criterion levels of seventy, eighty, and ninety percent, but not a higher level were reported as percentages for each of the treatment conditions.

Figure 8. Displaying criterion-referenced data.

		Metacognitive Strategies			
		YES		NO	
		50/90		38/90	
	YES	10/80	Y=25	3/80	Y=25
		5/70	N=25	1/70	N=25
Social Interaction					
		21/90		5/90	
	NO	6/80	Y=25	4/80	Y=25
		8/70	N=25	1/70	N=25

Note. Y equals the number of dependent measures used in tabulating the percentages. In this example, a single Y measure was used by each of the N respondents. The number of Y dependent measures may exceed a given N population if multiple outcome measures are utilized (Popham & Husek, 1969).

Figure 8 provides an example of how this data was displayed. Interpretation of the results of the administration of the criterion-referenced measures hinge on the percentage of students who meet a preset criterion level of performance. For this study, the researcher has selected a 90% criterion level as the performance measure of choice.

Norm-referenced Measures

Data derived from the administration of the norm-referenced geometry final (GEOFIN) was analyzed through analysis of variance. Group means, F-ratios, and levels of significance for all four of the treatment conditions are reported in Chapter 4.

It was generally accepted that prior achievement in school is a fairly accurate predictor of future performance within the school setting. With this hypothetical construct in mind, measures of prior knowledge of the students participating in the study were established before any treatment intervention took place. These measures were subsequently used in conjunction with the norm-referenced geometry measure (GEOFIN).

It was decided that scores from a standardized achievement test in the area of reading comprehension, along with a score from a standardized criterion-referenced test of geometric concepts would serve as the pretreatment indicators of prior knowledge, the baseline measures. Student scores from the reading comprehension subtest of the Stanford Achievement Test (Psychological Corporation, 1982b) for each participant in the study were obtained prior to treatment intervention. Student scores produced from an administration of the geometric concepts portion of the Sequential Assessment of Mathematics Inventories

(Reisman & Hutchinson, 1985) prior to the intervention period were included as a baseline measure of geometry performance.

These baseline measures used as covariates were established prior to any experimental intervention. An analysis of covariance data was recommended by Cook and Campbell (1979) for data of quasi-experiments that used intact groups. The covariance technique controls for selection differences among students of intact groups that may have existed as exhibited by the covariate measures prior to treatment. The relative effect of these pretreatment measures upon the norm-referenced dependent measure (GEOFIN) of this experiment are also addressed as part of the data analyses found in Chapter 4.

It should be noted that the geometric concepts subtest of the Sequential Assessment of Mathematics Inventories (SAMI) was administered twice in a pretest/posttest fashion. Administration of this measure occurred prior to intervention and at the conclusion of the intervention. Gain scores were established and used as yet another covariate in an analysis of covariance technique across the four treatment conditions. Again, group means, F-ratios, and levels of significance are reported in the following chapter of this study.

Affective Measures

Both analysis of variance and covariance techniques are similarly used to display the group means, F-ratios, and levels of significance of the scores produced by the administration of the confidence in learning mathematics portion of the Fennema-Sherman Mathematics Attitude Scales across treatment conditions.

The same covariate measures used for the norm-referenced measure (GEOFIN) were utilized with the measure of affect (CLM). Therefore, analysis of covariance was also used to address any selection differences that resulted from the use of intact groups.

Ability Measures

The group means, F-ratios, and levels of significance helped to summarize the data resulting from the administration of the abstract reasoning portion of the Differential Aptitude Test (DAT). Both analysis of variance and covariance techniques were used to analyze the variance across treatment conditions. Covariates previously described for norm-referenced and affective measures were similarly used in the data analyses for the measure of ability (DAT).

Qualitative Data Analyses

Prior to the study, an observational checklist (OC) was developed to be used in the experiment as a means to document the behavioral patterns of students. The checklist was used to categorize the student behaviors during the twenty-five minute period that follows the daily instruction in the mathematics classroom. This checklist was used to verify:

1) any student interaction patterns that may have emerged during the intervention period,
2) whether the student behaviors observed in the social interaction condition differed from those behaviors observed in the treatment condition that did not utilize social interaction,

3) whether significant similarities or differences of observed behaviors validated the treatment conditions.

All of the twenty treatment sessions were videotaped for all three intact groups. This videotaping allowed the researcher to utilize an observational checklist to verify the student behaviors across treatment conditions by viewing the tapes. The coding of the observed student behaviors was proceduralized in a systematic format in which the researcher observes the behavior of a student on the videotape and then entered a tally mark in the column that best described the behavior.

Observations and subsequent tallies were recorded at one minute intervals. Twenty tallies were recorded during the viewing of a single class session for each of the two treatment conditions being observed. For example while viewing a videotape, the researcher tallied twenty marks for behaviors of the social interaction condition, and twenty marks for student behaviors of the non-social interaction condition. The researcher purposely used no student twice in tallying the observed behaviors, unless all other students from that particular treatment condition had been observed. In this way any emerging patterns of observed behavioral differences were reported in a true qualitative fashion (Miles & Huberman, 1984).

This observational format produced a total of twenty frequencies for both of the treatment conditions, and forty categorized tallies for each treatment session. This routine was followed daily for each of the three treatment sessions.

The categorized tallies of the observed behaviors were reported as frequencies and percentages. The data was recorded in matrices when viewing the daily class session, and was reported in a cumulative display as a summary of the observational analysis. The qualitative data produced by use of the observational checklist (OC) is analyzed and discussed further in Chapter 4. Figure 9 provides an example of the observerational checklist created for this experiment.

Also included in the following chapters are comments regarding the data analyses of the observed student behaviors when compared and contrasted with the results of the quantitative data analyses previously described. The issue of significant differences or confirmation of data across the treatment conditions is also noted.

Figure 9. Observational checklist (OC).

	Helping Behaviors		Non-helping Behaviors	
Treatment Condition	Giving or Sharing Information	Receiving Help or Listening	Working Alone	Off-task
Social Interaction				
No Social Interaction				

Bias and Error Concerns

Since all of the twenty days of intervention for all three class periods were videotaped it was possible for an outside observer to assess whether or not treatment conditions did vary as the study intended. This outside observer also confirmed whether or not the teacher/researcher exhibited any bias in choice of instructional methods for use across the differing treatment conditions.

First this outside observer assessed whether or not the same content and learner objectives were used across the three intact groups during the instructional phase of the lessons. This was accomplished by the outside observer examining the lesson plans and corresponding instructional objectives for the three daily lessons contained on a videotape. This person was then given three randomly selected videotapes that portrayed the treatment of a single day. The outside observer then viewed a portion of each of the three lessons contained on the videotape to assess the content and methods utilized for the differing treatment conditions.

Second, the observer was trained in use of the observational checklist (OC). The outside observer used this instrument to code the observed behaviors of students in the differing social interaction conditions. A comparison of observational data obtained by the outside observer and the researcher followed. In this manner any unintentional bias effects that the teacher projected through his behaviors or speech patterns would be detected. The issue of bias is further discussed in Chapter 4.

Viewing of the videotapes by both the researcher and an outside observer allowed an evaluation of treatment conditions to take place. The teaching methods, student behaviors, and the content of the lessons were observed and assessed. The viewing of the videotapes also established whether or not unintentional bias was exhibited by the teacher/researcher while engaged in a treatment condition or when coding the observed behaviors of a videotape.

It should be noted that the question of unintentional bias does occur whenever the same teacher is used across varying treatment conditions in an experiment. Though this concern is always a threat to the internal validity of a quasi-experimental or experimental study, the method outlined above addresses this concern adequately. If the same teacher was not used across the varying treatment conditions, a new issue of consistency and bias would warrant further discussion. It was decided that the benefits of conducting this study in an actual school setting, with intact groups of students, and the regularly-assigned classroom teacher merited the experiment to be conducted as proposed. Limitations of this study are further discussed in Chapter 4.

CHAPTER IV

ANALYSES OF DATA

This study was conducted to evaluate two instructional techniques readily available to a teacher in a regular school setting. First, the relative value of incorporating metacognitive learning strategies as part of the program of instruction to aid students in the acquisition and retention of new information was investigated. Second, the use of social interaction among students was studied to assess the effects of planned cooperative learning activities on student learning.

Field Setting Research

According to Cook and Campbell (1979), research done in a field setting typically involves treatments, outcome measures, group assignments and some comparison of group performance in which change is inferred. A distinction is made between randomized experiments characterized by random assignment of individuals to various treatment conditions, and quasi-experiments that do not use random assignment. The authors have termed a particular type of quasi-experiments a "non-equivalent group design" when intact groups are used and responses across treatment conditions are measured before and after treatment.

The study reported here is a quasi-experiment incorporating a non-equivalent group design since it meets the criteria established by Cook and Campbell (1979) for research conducted in a field setting. A discussion of internal validity as it relates to randomization and quasi-experiments is warranted and follows.

Threats to Internal Validity

Confirming the internal validity of a study requires an investigator to systematically address how variables not controlled through randomization or direct manipulation may effect the outcome measures of an experiment. Cook and Campbell (1979) list the following as potential threats to the internal validity of any study: history, maturation, testing, instrumentation, statistical regression, selection, mortality, interaction, ambiguity about causal influence, treatment imitation, compensatory equalization of treatments, group rivalry, and demoralization.

Randomization eliminates many of the threats to internal validity listed above but, the issue differs considerably with quasi-experiments. When respondents are not randomly assigned to treatment conditions, the researcher must make the threats to internal validity known and address each concern (Cook & Campbell, 1979). For this reason the following sections address each of the aforementioned threats to internal validity as applied to the present study.

History

This term is used to describe a variable that is introduced between the pretest and posttest that is not planned as part of the study. In this particular study, history does not threaten the internal validity since all respondents from all groups experience a similar pattern of history prior to and during the treatment phase of the experiment.

Maturation

Since the treatment phase of this study spans twenty class sessions over a four-week period, it was unlikely that a change in the maturation level of respondents occurred. This short period of time would not contribute to a substantial change in the cognitive levels of respondents due to age alone. For this reason, maturation was not deemed as a potential threat to the study.

Testing

Students from three intact groups participating in the study were assigned to one of four treatment conditions. Students from all treatment conditions used the same instructional and testing materials. The effect that the use of testing materials may have upon future testing was held constant across treatment conditions resulting in little chance of potential bias for a particular group of respondents.

Instrumentation

This threat to internal validity occurs when a pretest or posttest measure fails to distinguish differences among group performance due to many scores falling near the bottom or top of the measurement scale. This event did not take place during the course of the study, and will be more fully addresed later in this chapter when the null hypotheses are discussed.

Statistical Regression

If the mean of the pretest scores for any of the three intact groups were deemed high or low, there would exist the possibility that later scores would move toward the center and balance out. Since the group means for pretest measures were not found to statistically differ (see Table 1), statistical regression to the group mean was not considered a viable threat to the internal validity of the study.

Selection

Though three intact groups were utilized in this study, pretest measures were not found to statistically differ in group mean scores (see Table 1). These three intact groups were specifically chosen for participation in the study from six available groups based upon the pretest measures. Selection is always a problem in quasi-experimental research since group differences may exist prior to any treatment (Cook & Campbell, 1979).

Kerlinger (1973) notes that a major difficulty confounding most educational research is the inability to set up experimental groups at will. Since random assignment in school settings is sometimes impossible, intact groups must be used. An analysis of covariance technique can be utilized to adjust group differences prior to the treatment on one or more covariates. This researcher acknowledged that selection differences were a threat to the internal validity of this study though the recommended analysis of covariance analyses was utilized.

Mortality

Related to selection, differences may occur among the treatment conditions if a particular type of respondent dropped out of the study resulting in an unplanned effect upon the outcome measures. This study experienced no such mortality, thus a threat to the internal validity was not an issue.

Interaction

Selection differences among intact groups may interact with other threats such as history, maturation, or instrumentation resulting in a particular group to experience an unplanned change during the course of an experiment. Since history, maturation, and instrumentation were not found to threaten the study, any combination of these factors with selection differences posed no threat to the internal validity of the study.

Ambiguity About Causal Inference

Ambiguity is a constant threat in most correlational studies. For example, a researcher may address the question of higher achievement scores in mathematics resulting in a more positive attitude towards learning mathematics; or does a more positive attitude towards learning mathematics result in higher achievement scores in mathematics? This threat was not evidenced in the study conducted as an experiment.

Imitation

When experiments include information intended for only certain treatment conditions, it is possible that other treatment conditions may ascertain this information thus violating the treatment differences. In this study the three intact groups were further broken down into one of two treatment conditions (social interaction and non-social interaction). It is possible that in this study information intended for one of these two treatment conditions was communicated to a differing treatment condition. Though this information may or may not have been ascertained by various respondents it is highly unlikely that this information could be utilized to alter group outcome measures that are not dependent upon this information (e.g. helping behaviors). Imitation was not deemed a threat to internal validity.

Compensatory Equalization

If a treatment conditions provides a benefit to one group it may be necessary for administrative reasons to provide this benefit to all groups comprising treatment conditions in the experiment. All three intact groups used in the present study experienced similar treatment conditions. For this reason no compensatory equalization of treatment conditions was required.

Rivalry

If the assignment of respondents to treatment conditions is publicly known, a resulting rivalry between experimental and control groups may occur. This rivalry would damage the internal validity of the study since emotional or motivational forces may alter the treatment conditions. Since the respondents in this study participated in a normal classroom setting, little or no environmental changes were evident. It is assumed that the three intact groups experienced no rivalry, for only subtle differences distinguished the various treatment conditions from the normal school setting. None of the three intact groups could be viewed as a true control group or as a less desirable treatment condition, thus eliminating the potential rivalry between groups.

Demoralization

Respondents in a control group may feel that they are in a less than desirable treatment condition. This could result in the participants acting out their frustration concerning their placement in a particular treatment condition. Sometimes demoralization is evidenced by the respondents exhibiting anger during the experiment, or other times is confirmed by a group's withdrawal or indifference to the experiment. This type of behavior would confound the results of the experiment by adding a variable to the study not controlled by the researcher.

Though this seems unlikely to have occurred between the three intact groups, it could have played a part within the group treatment conditions (i.e. social interaction and non-social interaction). Unintended as it may be, demoralization of respondents in one or more treatment conditions did pose a threat to the internal validity of this study.

Internal Validity Summation

Cook and Campbell (1979) note that randomization does eliminate many threats to internal validity but not all. Concerns relating to history, maturation, testing, instrumentation, statistical regression, selection differences, interaction and mortality are put to rest through the random assignment of respondents to treatment conditions. Though randomized experiments are superior to quasi-experiments in controling potential threats to internal validity, they still share the same concerns that quasi-experiments do in regard to group rivalry, imitation, demoralization, and compensatory equalization. Forced inequities undoubtedly exist with all experiments in a field setting whether they are randomized or quasi-experiment.

Limitations of the Study

This researcher recognized and addressed twelve potential threats to the internal validity of this study. After a careful review of the twelve threats, only selection differences and demoralization potentially threatened the results of this study. These two concerns regarding internal validity are acknowledged by this researcher as limitations of this study.

It is recognized that the the threat of selection differences is a present and real danger in a study of this type as in all quasi-experiments (Cook & Campbell, 1979). Pretreatment differences among the intact groups would negatively effect the internal validity of this study. For this reason, only those groups found not to significantly differ on the pretreatment measures were chosen for participation in this study (see Table 1). The threat of pretreatment selection differences damaging the internal validity of this study was minimized by choosing to use analyses of covariance techniques when evaluating the outcome measures (Kerlinger, 1973).

Though random assignment of individual respondents to the four treatment conditions used in this study was not possible, the three intact groups were randomly assigned to one of the two treatment conditions related to metacognitive strategies (i.e. metacognitive strategies and non-metacognitive strategies). Also regarding assignment of respondents to the other treatment conditions of social interaction (i.e. social interaction and non-social interaction), a randomized matching procedure based on previous achievement in mathematics was used to ensure that selection differences within the groups would be minimized. This procedure provided an equitable distribution of respondents to differing treatment conditions within the three intact groups. Thus randomization was used in this study to determine between-group treatment assignments, and also used to establish parity for within-group treatment assignments.

Though demoralization may have persisted as a potential threat to the internal validity of this study, it was not directly related to the study being quasi-experimental rather than randomized. The within group treatment established by this experiment (social intearction or non-social interaction) allowed for demoralization to occur if respondents came to view that one treatment condition was truly more desirable than another. No amount of randomization could eliminate the demoralization threat to the internal validity of this study.

Pretreatment Comparison of Intact Groups

As noted above, a comparison of the three intact groups selected for this study was completed prior to any intervention taking place. Since this study centered around applying a reading comprehension model of instruction to geometry instruction in a mathematics classrom, it was decided that prior knowledge and performance in both the constructs of reading comprehension and geometry should be noted for all respondents. With this in mind, an analysis of variance was selected as the statistic of choice to ascertain any differences in group performance across the three intact groups.

Reading comprehension was measured for all respondents by scores produced from a previous administration of the Reading Comprehension subtest of the Stanford Achievement Test Form E (Psychological Corporation, 1982b). This test is widely accepted as a reliable and valid measure for norm-referenced achievement testing in a regular school setting ($r = .94$). Test results are reported in grade equivalents derived by raw scores being converted into scaled scores as part of the standard norming process.

Since geometry as a construct is not recognized as a subtest of the mathematics portion of the Stanford Achievement Test Form E, an alternate pretreatment measure was selected. The Sequential Assessment of Mathematics Inventory (Reisman & Hutchinson, 1985) is a standardized norm-referenced test designed to assess the mathematics performance of students through eighth grade. The test consists of eight subtests, including a separate subtest entitled Geometric Concepts. It is this subtest of the Sequential Assessment of Mathematics

Inventory that was selected as a pretreatment measure to compare group mean performance.

The Examiner's Manual of the Sequential Assessment of Mathematics Inventory (SAMI) lists evidence of content validity, construct validity, and criterion-related validity. These validity checks are provided to verify that items assessed during the administration of this test do in fact reflect the curriculum normally found in a regular classroom setting. Reliability estimates for the Geometric Concepts portion of this test is deemed moderately high (KR-20 = .82) for students taking this examination in grades sixth through eighth.

Raw scores are reported with tables provided in the examiner's manual for conversion to grade-level equivalents and percentiles. The Geometric Concepts portion of this test consists of twenty-one items each covering a different criterion objective. A copy of the Geometric Concepts portion of the Sequential Assessment of Mathematics Inventory along with directions for administration are found in Appendix A.

Results from the administration of the Stanford Reading Comprehension subtest and the Geometric Concepts subtest of the Sequential Assessment of Mathematics Inventory are reported in Table 1. Scores are reported as group means for each of the three intact groups for both the measures of reading comprehension and geometry.

The analysis of variance procedure for the Reading Comprehension subtest of the Stanford Achievement Test produced no significant differences in group mean scores across the treatment conditions of intact groups (SIGNIF OF F = .534). Similarly, no significant differences were found in group mean scores across the treatment conditions of intact groups for scores produced by the Geometric Concepts portion of the Sequential Assessment of Mathematics Inventories (SIGNIF OF F = .686).

Table 1

Pretreatment Comparison of Intact Groups (Class) Using Analysis of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Reading Comprehension Score (SAT) by Class	8.384	2	4.192	0.632	0.534
Geometry Pretest Score (SAMI) by Class	4.745	2	2.373	0.379	0.686

Note. Sample N = 73.

*p < .05. **p < .01.

Though the design of this study is a quasi-experiment of non-equivalent groups (Cook & Campbell, 1979), no significant differences in group mean performance were found prior to treatment on measures related to the constructs of reading comprehension and geometry. This finding is important for it has direct bearing upon some of the internal validity concerns of this study previously raised due to selection differences.

The analysis of variance produced no evidence that existing differences among the three intact groups would later confound or limit the inference of causality. It is still possible that the pretest measures of choice were not precise enough to accurately detect or measure differences in the constructs as they existed among the groups. In light of the pretreatment evidence, it is much harder to question the internal validity of this study based on selection differences alone.

Instrumentation Concerns

Norm-referenced Measure

For purposes of the study, it was necessary to develop a criterion-referenced test based upon the geometry content for the regularly scheduled mathematics classes. This instrument was specifically designed for use as a dependent measure at the conclusion of the treatment.

A prototype test was written and administered to an intact group of twenty-four, seventh grade students not participating in the study. The students had prior knowledge of similar geometry content. An item-analysis and reliability check of the results was performed and later used to construct the actual outcome measure.

The geometry outcome measure used to assess student learning of the geometry content in this study has an established reliability that is moderately high (alpha = .8574) as seen in Table 2. The test, entitled Geometry Final (GEOFIN) actually used in this study is fifty items in length and can be found in Appendix B.

Regarding content validity, a trained and experienced junior high mathematics teacher who was not part of the study reviewed the fifty-item Geometry Final with corresponding objectives. This impartial party confirmed the content of the test as representative of the concepts normally associated with geometry instruction at a junior high school level.

Table 2

Reliability Analysis of the Geometry Final

Statistics for Scale	Mean	Variance	Standard Deviation	Number of Variables
	65.2432	53.0359	7.826	50
Reliability Coefficients			50 Items	
	Alpha = .8574		Standardized Alpha = .8600	

Note. Alpha model computed is Cronbach's alpha (1951), is equivalent to KR-20 (Kuder-Richardson-20) for dichotomous data (SPSS-X, 1986).

This same party was also asked to assess the criterion-related validity of the Geometry Final. The results of this review indicated that the fifty test items do correspond to the criterion-related objectives as written, thus verifying the criterion-related validity of the test.

Criterion-referenced Measures

Three, ten-item tests, each designed to measure the geometry content for five class sessions, were developed as part of this study. More specifically, each group of five class sessions used the ten-item test. The instruments were based upon the lesson objectives for the corresponding class sessions, and were formative in nature. These three tests were not piloted, but were deemed appropriate measures possessing criterion-referenced validity by this researcher based upon a review of the individual test items and corresponding lesson objectives.

These criterion-referenced tests are known as Criterion-referenced Test 1, Criterion-referenced Test 2, and Criterion-referenced Test 3. Copies of the actual tests can be found in Appendix C.

In general, criterion-referenced assessment establishes a desired performance level for the respondents that is acceptable for the purposes of the testing. This type of assessment does not fit a normal curve distribution of scores, for it is not based on variance among the test scores. In criterion-referenced testing it is acceptable for all respondents to master the material producing a cluster of scores with little or no dispersion.

Variability and dispersion of test scores, both necessary to the reliability of a norm-referenced measure are not factors associated with verifying the reliability of criterion-referenced measures. Basically, the reliability of a criterion-referenced tests cannot be measured with methods appropriate for norm-referenced tests (Popham & Husek, 1969). The length of the tests, which contain only ten items, was a factor that severely limits the possibilty that enough variance or dispersion among test scores occurred to accurately assess the reliability of the measures using norm-referenced procedures.

With this in mind, no attempt was made by this researcher to assess the reliability of these three, ten-item measures. Instead, a criterion level of performance was selected based upon mastery of the geometry content contained in the instructional phase of the class sessions. This resulting nominal data for each of the three, criterion-referenced tests is reported in terms of frequencies and percentages. For example, if 50% of the students achieve 90% or higher mastery on the test measure, then 50-90 is reported.

Thus, the issues of reliability and validity as they pertain to the development and use of the three, criterion-referenced measures have been addressed in two ways. First, in lieu of a reliability score, the percentage of students who reach the preset criterion levels of 70, 80, and 90% for each of the three, ten-item tests are reported (see Table 6). Second, these measures have been deemed to have criterion-related validity for the purposes of this study.

Coding of Videotapes

To ensure that collaborative skills normally associated with social interaction were evidenced by the participants in this study, all treatment sessions were videotaped. The taping of the treatment sessions included both the instructional and student practice phases that occurred daily for all of the three intact groups.

As previously stated, an observational checklist (OC) was created that reflected those factors found to be important to social interaction as reported by current literature (Johnson & Johnson, 1984; Webb, 1982). Constructing this instrument based upon existing criteria used in previous research has insured that the measure possesses criterion-referenced validity. A copy of this instrument is found in Appendix D.

Regarding bias effects, a mathematics teacher not involved in this study viewed three selected videotaped sessions. This independent observer utilized the observational checklist to code the social interaction of the respondents according to the procedures outlined in Chapter 3. A preset criterion-level of 90% was selected to compare the results produced by the independent observer and those previously coded by this researcher.

The resulting analysis of videotaped sessions found no significant differences due to error or bias effects in the observational checklist of the independent observer and this researcher. It should be stated that the purpose and use of the observational checklist was limited to confirming whether those students in collaborative treatment conditions (social interaction) did differ in behaviors from those students in non-collaborative treatment conditions (non-social interaction). This data are qualitative in nature, and only one of several outcome measures used in answering the research questions of this study. Complete findings related to the observational checklist are found in Table 18 of this chapter.

<u>Summary</u>

This section of the chapter addressed the reliabilty and validity of instruments designed and developed for purposes of this study. These measures included the following: Geometry

Final, Criterion-referenced Test 1, Criterion-referenced Test 2, Criterion-referenced Test 3, and the Observer's Checklist.

The Geometry Final exhibited high validity and reliabilty when evaluated through procedures normally assosciated with norm-referenced testing. The Criterion-referenced Tests 1, 2, and 3 have been noted as valid measures, and a case was made that explained why a reliability measure was not given. A preset criterion level was used to compare results of the Observer's Checklist produced by an independent party and this researcher. This analysis confirmed that the instrument was reliable as used in this study.

External Validity Issues

External validity has been defined by Cook and Campbell (1979) as generalizations that can be made from a population of a one study to and across other populations. The authors stress that importance of generalizations across other populations over those generalizations that can be made to a specific population.

Related to external validity is the concept of construct validity formerly discussed as an internal validity issue. Kerlinger (1973) explains that a researcher verifying the construct validity of a test is concerned with more than what the test purports to measure, for one is also interested in the property being measured within the test.

Therefore, in the present study, factors normally associated with learning are addressed along with the measures used to assess the content of geometry. In this manner a respondent's attitude towards learning mathematics or spatial ability are hypothetical constructs that could effect performance when measured by a geometry test. For these reasons, a measure of affect and a measure of ability where included as post-treatment measures to address the external validity of the study.

In an attempt to generalize the results of this study across other target populations, an array of outcome measures have been utilized. Dependent measures related to cognitive performance (geometry), instructional strategies (metacognition), social interaction, affect and ability are discussed in the data analyses portion of this chapter. With this in mind, it should be noted that the primacy of internal validity is not being compromised by this researcher. Generalizations dependent upon the external validity of this study will only be made if concerns regarding the internal validity of the study have been fully satisfied.

Independent Variables

A review of the independent variables is provided to facilitate a discusssion of the research findings for the four treatment conditions. There were two independent variables included in this study. First, metacognitive strategies were introduced as a treatment variable. Second, social interaction was used as another treatment variable. These two independent variables produced a 2 x 2 matrix resulting in four treatment conditions. Table 3 shows the four differing treatment conditions of this study.

Table 3

Display of Differing Treatment Conditions

Treatment Conditions Created by Independent Variables
Metacognitive Strategies with Social Interaction (I)
Metacognitive Strategies Only/No Social Interaction (III)
Social Interaction Only/No Metacognitive Strategies (II)
Neither Metacognitive Strategies nor Social Interaction (IV)

Matrix Display of Differing Treatment Conditions

		Metacognitive Strategies	
		YES	NO
Social Interaction	YES	I	II
	NO	III	IV

Note. Two independent variables create a 2 x 2 research design.

Sample Size

As mentioned in a prior section of this chapter, within group treatment differences were created in order to utilize the three intact groups in four differing treatment conditions. This arrangement of respondents resulted in unequal cell numbers across treatment conditions as evidenced in Table 4.

The analysis of covariance procedure was selected to analyze the results of the data for the outcome measures benefited this study for this procedure is often used specifically to address unequal cell numbers in a research study. It appears that a sufficient cell number (N = 10) is found in each of the four treatment conditions, thus no concern regarding sample size is evidenced. The sample size is termed adequate for purposes of this study with no limitations related to population size noted.

Table 4

Sample Size of Differing Treatment Conditions

Metacognitive Strategies with Social Interaction (I)
Metacognitive Strategies Only/No Social Interaction (III)
Social Interaction Only/No Metacognitive Strategies (II)
Neither Metacognitive Strategies nor Social Interaction (IV)

Matrix Display of Sample Size

		Metacognitive Strategies: YES	Metacognitive Strategies: NO
Social Interaction	YES	I N = 24	II N = 10
	NO	III N = 25	IV N = 14

Note. Total population of sample N = 73.

Dependent Variables

A classification was made between the quantitative and qualitative dependent measures used in this study. Only one qualitative measure was utilized, that being the Observer's Checklist of coded videotaped episodes. There were four quantitative measures that were used as dependent variables in this study. The quantitative measures are as follows: the Geometry Final, the Confidence in Learning Mathematics portion of the Fennema-Sherman Mathematics Attitude Scales (Fennema & Sherman, 1976), the Abstract Reasoning portion of the

Differential Aptitude Test (Psychological Corporation, 1982a), and the Criterion-referenced Tests 1-3.

Table 5 lists the names of the instruments used as dependent outcomes in this study with a distinction made between qualitative and quantitative measures. An additional distinction was made between assessment measures that the researcher developed for use in this study and existing measures published by outside sources.

Table 5

List of Dependent Measures

Name of Outcome Measure	Brief Description
Observer's Checklist (OC)	Qualitative measure developed for use in this study.
Geometry Final (GeoFin)	Quantitative measure developed for use in this study.
Criterion-referenced Tests 1, 2, & 3 (CRT1, CRT2, CRT3)	Quantitative measure developed for use in this study.
Sequential Assessment of Mathematics Inventory (SAMI)	Extant quantitative measure.
Confidence in Learning Mathematics Attitude Scale (CLM)	Extant quantitative measure.
Differential Aptitude Test (DAT)	Extant quantitative measure.

Discussion of the Data

The design of this study incorporated two independent variables, four treatment conditions, three intact groups, and multiple dependent measures in an attempt to answer the research questions. This model produced thirteen null hypotheses that must be addressed in this section of the chapter.

To expedite this analyses and to facilitate a clear understanding of the data, acceptance or rejection of each of the null hypotheses will be discussed in two ways. First this section will systematically relate the data for each of the dependent outcome measures making references to the corresponding null hypotheses as appropriate. Second, at the conclusion of this chapter a summary table is provided to review the status of each of the thirteen null hypotheses as a result of the data analyses.

Data Analyses

Criterion-referenced Measures

Criterion-referenced Tests (CRT1, CRT2, CRT3)

Null hypotheses 1, 2, and 3 of this study relate to criterion-referenced measures. The three, ten-item tests constructed for use as criterion-referenced measures (CRT1, CRT2, & CRT3) produced data that was categorized in preset criterion levels in a manner suggested by Popham and Husek (1969).

The results of this categorization of the criterion-referenced data is found in Table 6. The data were not reported in group relative measures such as means, percentile ranks or other procedures common to norm-referenced data analyses. Instead the percentage of students meeting a certain criterion level is reported across the treatment conditions.

Through this analysis it is evident that the largest percentage of students scoring at the preset criterion level of

90% or higher was from the treatment condition that used both social interaction and metacognitive strategies. The next highest percentage of students scoring at the 90% or higher criterion level was the social interaction only treatment condition, followed closely by the metacognitive strategies only group of students. The treatment condition that utilized neither social interaction nor metacognitive strategies fared the lowest percentage of students scoring at the 90% or higher criterion level.

Differences in percentages of students scoring at the next highest criterion level at 80% or higher were less evident. It is quite clear that at the lowest criterion level of 70% or higher, the number of students from the treatment condition that received neither social interaction nor metacognitive strategies differed from other treatment conditions.

Conclusions Regarding the Criterion-referenced Data

In light of these findings, it can be said that the percentage of students who obtained scores at various preset criterion levels did differ across treatment conditions. This disparity was noticeable when comparing students from differing metacognitve strategies conditions, differing social interaction conditions, and in a combination of the metacognitive strategies and social interaction conditions. This analyses of the data confirm that null Hypotheses 1, 2, and 3 should be rejected due to noticeable differences in scores produced by criterion-referenced measures across the treatment conditions at the preset criterion level of 90% or higher.

Table 6

Display of Data for Criterion-referenced Measures (CRT1, CRT2, & CRT3)

Social Interaction		Metacognitive Strategies: YES	Metacognitive Strategies: NO	TOTALS
	YES	65/90 17/80 8/70 n = 24 Y = 72	43/90 13/80 7/70 n = 10 Y = 30	n = 34
	NO	40/90 17/80 11/70 n = 25 Y = 75	24/90 21/80 24/70 n = 14 Y = 42	n = 39
	TOTALS	n = 49	n = 24	N = 73

Note. Total population of sample N = 73. The symbol "n" is used to denote the population size, with the symbol "Y" denoting the total number of measures of a particular treatment condition. The notation 40/90 is interpreted as 40% of the respondents from the particular treatment condition met or exceeded the 90% criterion level (Popham & Husek, 1969).

Normed-referenced Measures

Geometry Final Data

The data produced by the administration of the normed-referenced test of geometry is related to testing null Hypotheses 4, 5, and 6 of this study. As seen in Table 7, an analysis of variance was performed to compare the performance of across the varying treatment conditions of metacognitve strategies and social interaction.

Table 7

Analysis of Variance for Geometry Final (GEOFIN) by Metacognitve Strategies (Meta) by Social Interaction (Social)

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	603.701	2	301.850	7.030	0.002**
Meta	66.366	1	66.366	1.546	0.218
Social Interaction	509.152	2	509.152	11.859	0.001**
2-Way Interactions					
Meta by Social	31.130	1	211.677	4.930	0.396

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

The results of this analysis confirmed that there was a significant statistical difference evidenced in the main effects category. Differences in performance were found to be significant ($p < 0.01$) across the social interaction treatment conditions. No significant differences were evidenced across metacognitive treatment conditions, nor was there any significant interaction found between the metacognitive and social interaction treatment conditions..

An analysis of covariance was used to factor out the possible effects that prior achievement in reading may have across treatment conditions. Table 8 portrays the data from the analysis of covariance for treatment conditions with a reading comprehension covariate.

Differences in the main effects category were found to be significant ($p < 0.05$) using the reading comprehension covariate. A further analysis of the main effects failed to show significant differences for either the metacognitive strategies or the social interaction treatment conditions. No significant differences were found in the interaction category for these treatment variables. Though an apparent difference occurred in the main effects category, a closer analysis of the data did not yield significant differences for the reading covariate across treatment conditions.

Table 8

Analysis of Covariance for Geometry Final (GEOFIN) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Reading Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Reading (SAT)	816.545	1	816.545	22.437	0.000**
Main Effects	295.260	2	147.630	4.057	0.022*
Meta	144.228	1	144.228	3.963	0.051
Social Interaction	110.541	1	110.541	3.037	0.086
2-Way Interactions					
Meta by Social	10.990	1	10.990	0.302	0.584

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

Using the geometry pretest as a covariate, an analysis of covariance was conducted to analyze group performance differences across the treatment conditions. Table 9 displays the data findings related to the geometry covariate effects upon performance at the end of the treatment.

Table 9

Analysis of Covariance for Geometry Final (GEOFIN) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Geometry Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Geometry Pretest (SAMI)	981.325	1	981.325	30.844	0.000**
Main Effects	442.533	2	221.267	6.955	0.002**
Meta	118.334	1	118.334	3.719	0.058
Social Interaction	289.579	1	289.579	9.102	0.004**
2-Way Interactions					
Meta by Social	10.159	1	10.159	0.319	0.574

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

Regarding Table 9, significant differences were found across treatment conditions in the main effects category ($p < 0.01$). The analysis confirms that though differences in group performance were not found to be significant in the metacognitive strategies treatment conditions, significant differences were found across the social interaction groups ($p < 0.01$). Once again, no interaction between the metacognitive strategies and social interaction conditions were found.

The geometry pretest was also used in this study as a posttest. A gain score covariate (SAMI II - SAMI I)was established for use in an analysis of covariance. The results of geometry gainscore effects upon group performance evaluated by the geometry final are found in Table 10.

Table 10

Analysis of Covariance for Geometry Final (GEOFIN) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Gainscore Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Gainscore (Gaingeo)	55.716	1	55.716	1.262	0.266
Main Effects	428.368	2	214.184	4.849	0.011*
Meta	12.907	1	12.907	0.292	0.591
Social Interaction	397.656	1	397.656	9.002	0.004**
2-Way Interactions					
Meta by Social	44.172	1	44.172	1.000	0.322

Note. Sample N = 73.
*p < .05. **p < .01.
Gainscore (Gaingeo) = (SAMI POST - SAMI PRE)

Main effect differences were found to be significant ($p < 0.05$) using the gainscore geometry covariate across treatment conditions. In this category, metacognitive strategy conditions did not differ significantly. The social interaction conditions differed significantly ($p < 0.01$) when analyzing group performance using this geometry gainscore covariate. No interaction was found to be significant across the treatment conditions of metacognitive strategies and social interaction.

Table 11 provides results of the data produced by an analysis of covariance using the intact group variable as a covariate. This procedure was utilized to assess any apparent differences in group performance that may be attributed to a particular intact group outperforming another. Again, the grometry final test was the norm-referenced measure used to compare performance across groups and treatment conditions.

Table 11

Analysis of Covariance for Geometry Final (GEOFIN) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Intact Group Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Intact Group (Class)	31.181	1	31.181	0.724	0.398
Main Effects	606.658	2	303.329	7.042	0.002**
Meta	65.141	1	65.141	1.512	0.223
Social Interaction	513.363	1	513.363	11.918	0.001**
2-Way Interactions					
Meta by Social	30.596	1	30.596	0.710	0.402

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

This analysis confirmed significant main effect differences ($p < 0.01$) across the treatment conditions. No significant differences were confirmed across the metacognitive treatment conditions, but social interaction conditions did differ significantly ($p < 0.01$). No interaction across the treatment conditions of social interaction and metacognitive strategies was confirmed to be significant.

A closer analysis of the cell means of those data revealed the source of variation across the social interaction treatment conditions. Scores of group means for the treatment conditions of social interaction with metacognitive strategies, and social interaction without metacognitive strategies were higher than those of treatment conditions that received no social interaction. These findings are consistent with those reported in Tables 7, 9 and 10.

Conclusions Regarding the Geometry Final Data

Null Hypotheses 4, 5, and 6 correspond to the normative data of the geometry final test. Null Hypothesis 4 was not rejected for significant differences were not found across treatment conditions of metacognitive strategies. Likewise, no significant interaction was detected across treatment conditions. Thus, the null Hypothesis 6 was not rejected. However, significant differences were confirmed ($p < 0.01$) across the social interaction treatment conditions resulting in a rejection of the null Hypothesis 5.

Measure of Affect Data

The Confidence in Learning Mathematics portion of the Fennema-Sherman Mathematics Attitude Scales (Fennema & Sherman, 1982) is the norm-referenced measure that is related to the testing of null Hypotheses 7, 8, and 9 of this study. This measure of affect was utilized as a posttreatment dependent measure. An analysis of variance was performed to compare group performances on this measure of affect across the treatment conditions. Results of this analysis are found in Table 12.

In the main effects category, significant differences were found ($p < 0.05$). The metacognitive strategies treatment conditions differed significantly ($p < 0.01$), while no differences were deemed significant across the social interaction conditions. Two-way interactions were not found to be significant for the differing treatment conditions.

Table 12

Analysis of Variance for Affect Measure (CLM) by Metacognitve Strategies (Meta) by Social Interaction (Social)

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	470.332	2	235.166	3.475	0.036*
Meta	446.495	1	446.495	6.598	0.012*
Social Interaction	11.666	1	11.666	0.172	0.679
2-Way Interactions					
Meta by Social	214.361	1	214.361	3.373	0.080

Note. Sample N = 73.
*p < .05. **p < .01.

A further assessment of the scores produced by this measure of affect was conducted using an analysis of covariance with covariates previously utilized in other norm-referenced data analysis of this study. Results for this analysis of covariance with the geometry pretest covariate are found in Table 13.

Main effect differences were significant ($p < 0.05$) for this measure of affect when using the geometry pretest covariate. The metacognitive strategies conditions differed significantly ($p < 0.01$). The social interaction conditions did not significantly differ. A significant difference was found in the interactions among the metacognitive strategies and social interaction treatment conditions ($p < 0.05$) for group performance.

Table 13

Analysis of Covariance for Affect Measure (CLM) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Geometry Pretest Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Geometry Pretest (SAMI)	395.711	1	395.711	6.485	0.013*
Main Effects	536.114	2	268.057	4.393	0.016*
Meta	534.030	1	534.030	8.752	0.004**
Social Interaction	0.274	1	0.274	0.004	0.947
2-Way Interactions					
Meta by Social	272.797	1	272.797	4.471	0.038*

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

A final analysis of covariance was performed using the geometry final test as a covariate. Table 14 displays the data for the measure of affect with the geometry posttreatment measure used as a covariate.

Using the geometry posttreatment covariate, significant differnces were evidenced in the main effects category of metacognitive strategy ($p < 0.05$). Social interaction conditions did not significantly differ in the main effects category. Interaction effects were found to be significant ($p < 0.05$) across the differing treatment conditions.

Table 14

Analysis of Covariance for Affect Measure (CLM) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Geometry Final Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Geometry Final (GeoFin)	680.585	1	680.585	11.445	0.001**
Main Effects	336.519	2	168.260	2.830	0.066
Meta	304.730	1	304.730	5.125	0.027*
Social Interaction	33.146	1	33.146	0.557	0.458
2-Way Interactions					
Meta by Social	293.196	1	293.196	4.931	0.030*

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

Conclusions Regarding the Affect Data

Null Hypothesis 7 relating to the measure of affect should be rejected since significant differences were found for the metacognitive treatment conditions (p < 0.01). Significant differences were also found at a lower level (p < 0.05) regarding interactions of the metacognitive strategies and social interaction conditions. Thus, null Hypothesis 9 should be rejected. No significant differences were found to be significant when comparing measures of affect across the social interaction treatment conditions. For this reason, null Hypothesis 8 was not rejected.

Measure of Abilty Data

The Abstract Reasoning Portion of the Differential Aptitude Test (Psychological Corporation, 1982) was used as a measure of ability in this study. The administration of this measure took place at the end of the treatment period and was used as a norm-referenced dependent outcome. Table 15 provides the data for an analysis of variance performed to assess group performance in this ability measure across treatment conditions.

Table 15

Analysis of Variance for Ability Measure (DAT) by Metacognitve Strategies (Meta) by Social Interaction (Social)

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	377.231	2	188.616	4.670	0.013*
Meta	309.580	1	309.580	7.665	0.007**
Social Interaction	48.913	1	48.913	1.211	0.275
2-Way Interactions					
Meta by Social	10.732	1	10.732	0.266	0.608

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

The main effects category of metacognitive strategies was found to differ significantly (p < 0.01) when this ability measure was used as the dependent variable. No significant differences were found in either the social interaction category of main effects or in the interactions of the two independent variables across the treatment conditions.

The same covariates previously established for other norm-referenced data analysis in this study were also used to further evaluate group performances using the ability measure. The geometry pretest was the covariate in an analysis as summarized in Table 16.

Table 16

Analysis of Covariance for Ability Measure (DAT) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Geometry Pretest Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Geometry Pretest (SAMI)	558.937	1	558.937	17.476	0.000**
Main Effects	412.235	2	206.117	6.445	0.003**
Meta	394.796	1	394.796	12.344	0.001**
Social Interaction	6.121	1	6.121	0.191	0.663
2-Way Interactions					
Meta by Social	28.684	1	28.684	0.897	0.347

Note. Sample N = 73.
*$p < .05$. **$p < .01$.

Main effects differed significantly (p < 0.01) with the treatment conditions of metacognitive strategies being noted as the source of the variance. No significant differences or interactions were found regarding the social interaction treatment conditions. The use of the geometry pretest covariate in this analysis did confirm significant differences for the metacognitive treatment conditions.

A second covariate, the geometry final test was also used in this study related to the ability measure. Table 17 portrays the data produced by an analysis of covariance procedure for this ability measure.

Again, main effects differed significantly ($p < 0.05$) for the treatment conditions of metacognitive strategies. No significant differences in group ability scores were evidenced across the social interaction treatment conditions. Interactions between the independent variables were not deemed significant.

Table 17

Analysis of Covariance for Ability Measure (DAT) by Metacognitve Strategies (Meta) by Social Interaction (Social) with Geometry Final Covariate

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariate					
Geometry Final (GeoFin)	661.747	1	661.747	19.751	0.000**
Main Effects	203.666	2	101.833	3.039	0.054
Meta	200.201	1	200.201	5.975	0.017*
Social Interaction	3.834	1	3.834	0.114	0.736
2-Way Interactions					
Meta by Social	30.974	1	30.974	0.924	0.340

Note. Sample N = 73.
*p < .05. **p < .01.

Conclusions Regarding the Ability Measure Data

Null Hypotheses 10, 11, and 12 correspond to the measure of ability data previously discussed. As a result of the data analyses, null Hypothesis 10 should be rejected for significant differences ($p < 0.01$) were found between the metacognitive strategies treatment conditions. Null Hypotheses 11 and 12 can not be rejected, for significant differences were not found in either the categories of social interaction or the interaction between the independent variables.

Measure of Social Interaction Data

Null Hypothesis 13 states that observed student behaviors will not differ between the social interaction treatment conditions. All seventy-three participants of this study were either grouped into a treatment condition that encouraged social interaction among students, or into a treatment condition in which students were asked to work independently.

Videotapes of the treatment sessions were used to assess the social interaction of the students through use of a protocol coding scheme previously described. The Observer's Checklist used as part of the coding scheme produced frequency data for two categories of helping behaviors, and two categories of non-helping behaviors. This frequency data for the differing treatment conditions was reported as percentages which are displayed in Table 18.

When analyzing the results of those data, it was apparent that percentages for the categories did differ for the treatment conditions. Students assigned to the social interaction condition consistently showed more helping behaviors than the students assigned to no social interaction condition. It should also be noted that off-task behaviors were less evident in the social interaction condition than in the no social interaction condition. The percentages for the category of receiving help did not exhibit pronounced differences.

Conclusions Regarding the Social Interaction Data

The Observer's Checklist and videotapings of the treatment sessions were utilized to assess whether planned differences in social interaction were evidenced between the treatment conditions. The testing of null Hypothesis 13 corresponds to the social interaction data discussed above.

Table 18

Summary Data for Observer's Checklists (Videotapes)

	Helping Behaviors		Non-helping Behaviors	
Treatment Condition	Giving or Sharing Information	Receiving Help or Listening	Working Alone	Off-task
Social Interaction	47.8%	14.7%	29%	8.5%
No Social Interaction	2.6%	9%	72.4%	16%

Note. Actual frequency data from observations represented here in percentages is found in Appendix E.

The data provided by the protocol coding scheme verified that the social interaction patterns differed between the treatment conditions as planned. For this reason, null Hypothesis 13 should be rejected.

Summary

Data, results, and the decisions regarding the thirteen null hypotheses used in this study were reported in this chapter. The research focused upon the effects that the two treatment variables (metacognitive strategies and social interaction) had upon various performance outcomes (i.e. quantitative and qualitative measures; criterion-referenced and norm-referenced tests; measures of achievement, affect and ability).

To benefit the reader, decisions regarding the null hypotheses are summarized in Table 19. Implications and generalizations about the findings of this study are discussed in Chapter 5.

Table 19

Summary of Decisions Regarding Null Hypotheses (HO)

HO	Decision	Rationale for Decision
1	Reject	Data based on preset criterion levels.
2	Reject	Data based on preset criterion levels.
3	Reject	Data based on preset criterion levels.
4	Fail to reject	No significant differences evidenced.
5	Reject	Significance evidenced ($p < 0.01$).
6	Fail to reject	No significant differences evidenced.
7	Reject	Significance evidenced ($p < 0.01$).
8	Fail to reject	No significant differences evidenced.
9	Fail to reject	No significant differences evidenced.
10	Reject	Significance evidenced ($p < 0.01$).
11	Fail to reject	No significant differences evidenced.
12	Fail to reject	No significant differences evidenced.
13	Reject	Pronounced differences in behaviors.

CHAPTER V

DISCUSSION

The results of the data analyses previously described in this study are now utilized in this chapter to discuss the original intent of this study. This discussion centers around how the research questions were addressed in this dissertation, and to what extent these questions were answered. Any implications found worthy of further study by this researcher are mentioned, along with any practical application of the research findings to a normal classroom setting.

The following components of this study are noted in this chapter: 1) the decision to reject or fail to reject the null hypotheses, 2) the effects of the independent variables upon the dependent measures, 3) the limitations of the research as addressed, 4) the generalizations that can be made from this study to and across other groups, and 5) recommendations for future research in this field.

The first section of this chapter contains a brief review of the study. A summary of the purpose, population, research design, limitations, and null hypotheses used in this dissertation is included. Finally, this summary serves as a general framework to address the five components deemed worthy of discussion.

Review of the Study

Purpose

The purpose of this research was to study the effects of an interactive teaching method, known as reciprocal teaching, upon mathematics learning in a junior high school classroom. More specifically the study focused upon the learning strategies that characterize reciprocal teaching (i.e. summarize, question, clarify and predict) when used within the particular content domain of geometry.

Population

Seventy-three students participated in this study. All of the students attended the same public school, a junior high school comprised of grades six through eight. The students basically reflected the ethnic make-up of the predominantly white, middle-class suburb of Chicago in which they all lived.

Classes at this junior high school are tracked into three academic levels consisting of developmental, average and accelerated. Selection of a student for a particular tracking was based upon prior school performance, results of achievement testing, and teacher recommendations.

For this study only intact groups were used. Six possible groups were initially considered for inclusion in the quasi-experiment. However, only three of the six intact groups of students were tracked into average mathematics classes. These three intact groups were selected for participation in the study.

Group mean scores for two, pretreatment measures recognized as predictors of future performance in mathematics were obtained for the three intact groups. A resulting data analysis found no significant differences in the group mean scores across the three, intact groups on either pretreatment measures. Two of the three intact groups were seventh grade mathematics classes, with the remaining group being a sixth grade class. The total student population of these three, intact groups comprised the total population of the study.

Design

The study was conducted in a field setting and kept the existing parameters of a natural, school environment. For this reason, the treatment condition occurred within the realm of a normal class period that lasted fifty minutes. The three intact groups selected to comprise the population of this study shared the same mathematics teacher, this researcher. No changes were made in the length of regularly scheduled classes, the time scheduling of classes, student assignment, or teacher assignment to conduct this research.

The design of the study required a treatment plan for the duration of twenty, successive class sessions. This treatment took place over a four-week period. Each treatment session scheduled reflected the parameters of the regularly scheduled mathematics class of fifty minutes in length.

Each class session was divided into two equal phases of twenty-five minutes in length. The first phase of the class session was the instructional component of the class session while the second phase of each class session was devoted to student follow-up activities.

Two independent variables of metacognitive strategies and social interaction were selected for research. These two variables constituted the components of reciprocal teaching, or in other words, the main treatment effects. It was the second half of each class period that utilized the treatment effects discussed in this study.

Based upon the independent variables, the resulting research design of the study required four treatment conditions. The four differing treatment conditons of this study were as follows: 1) social interaction only, 2) metacognitive strategies only, 3) social interaction with metacognitive strategies, and 4) neither social interaction nor metacognitive strategies.

Two of the three, intact groups were randomly assigned to receive either the metacognitive strategies treatment or not to receive this treatment. In addition, students within each of these two intial conditions were randomly matched and assigned to either a treatment condition that ecouraged social interaction or a condition that did not encourage social interaction. This process allowed for the three naturally occurring groups to be used in four varying treatment conditions required to address the research questions.

Dependent variables chosen for this study were criterion-referenced test measures, norm-referenced test measures, a measure of affect, a measure of ability, and an observational checklist of student behaviors. The use of these multiple dependent measures required various analysis of data techniques to be employed.

Some of the outcome measures were reported as frequency data and percentages. Another measure was reported in terms of a preset criterion level of performance. Two multivariate techniques, analysis of variance and analysis of covariance, were performed to analyze and interpret the data related to group mean scores. Videotaping was also incorporated into the study to assess whether any unintentional bias effects were evidenced, and to confirm student behaviors within differing treatment conditions.

Limitations

In Chapter IV of this dissertation, two limitations of this study were discussed. First, since this study was a quasi-experiment (Cook & Campbell, 1979) using intact groups in a field setting, selection differences were noted as a possible threat to the internal validity of the study and as a study limitation. Data analysis performed specifically to address this concern resulted in dismissing this threat.

Second, demoralization was cited as a possible limitation of this study. Demoralization may have occurred as a result of group rivalry between respondents within a particular intact group. This concern was not related to the use of intact groups or to randomization not being used to assign students to treatment conditions. The design of the study required four groups of students to comprise the four treatment conditions. Since only three groups were utilized, it was necessary to assign respondents within the three groups to differing treatment conditions that occurred simultaneously in the same classroom. The fact that this research took place in a field setting with the parameters described, raised the possibility that group demoralization be considered as a limitation of the study.

Discussion of Results

Decisions Regarding the Null Hypotheses

Thirteen null hypotheses were created to answer the research questions posed by this study. These thirteen hypotheses were related to five distinct categories of outcome measures. Four of the categories of outcome measures were classified as either one of the following: criterion-referenced, norm-referenced, affective and ability. A fifth outcome measure was descriptive in nature, relating to the observed behaviors of students. Decisions made regarding the null hypotheses are based upon these various outcome measures. A summary of the thirteen null hypotheses and the nature of their corresponding outcome measures is found in Figure 10.

Criterion-referenced Measures

A preset criterion level of ninety percent was selected for analysis and interpretation of the data from the criterion-referenced measures. Null hypotheses 1, 2, and 3 correspond to the data produced by these criterion-referenced measures. On the basis of the preset criterion level, a decision was made to reject all three of the null hypotheses based upon significant differences across all treatment conditions.

Figure 10. Nature of the outcome measures for the various null hypotheses (HO).

HO	Related Independent Variable	Related Outcome Measure
1	Metacognitive Strategies	Criterion-referenced
2	Social Interaction	Criterion-referenced
3	Interaction	Criterion-referenced
4	Metacognitive Strategies	Norm-referenced
5	Social Interaction	Norm-referenced
6	Interaction	Norm-referenced
7	Metacognitive Strategies	Affective Measure
8	Social Interaction	Affective Measure
9	Interaction	Affective Measure
10	Metacognitive Strategies	Ability Measure
11	Social Interaction	Ability Measure
12	Interaction	Ability Measure
13	Social Interaction	Observed Behavior

The group of students that received both the treatment variables of metacognitive strategies and social interaction scored the highest percentage consisting of the ninety percent criterion level (65/90). Groups of students that comprised the two treatment conditions of social interaction only (43/90) and metacognitive strategies only (40/90) produced data that differed slightly from each other. Scores from both of the groups of students receiving only one treatment variable still differed noticeably from the group that received both treatment variables. The group of students who served as the control group and received neither treatment variable ranked lowest in the percent of students who met the preset criterion level (24/90).

The analysis of data performed for these criterion-referenced measures may be questioned by researchers versed only in norm-referenced measures. For this reason, findings and conclusions resulting from analysis of the criterion-referenced measures are noted again in the section related to norm-referenced measures.

Norm-referenced Measures

Analysis of data produced by the norm-referenced measures resulting in both main effects and interaction effects, will be discussed for the corresponding null hypotheses. It should be noted that covariates were also used with the norm -referenced data analyzed through a multivariate technique. For these reasons, Figure 11 is provided to detail and interpret the findings related to null Hypotheses 4, 5, and 6, and to assist the reader.

Regarding the treatment condition that utilized metacognitive strategies, no significant main effects were evidenced. Though covariates were used along with the norm-referenced outcome measure, this data analysis failed to show any significant differences in group mean scores across the treatment conditions. Nor did the metacognitive strategies conditions produce enough variance to confirm any interaction effects. For these reasons, the data presented by this research failed to to reject null Hypotheses 3 and 6.

<u>Figure 11.</u> Decisions related to null Hypotheses 4, 5, and 6 regarding main effects, interaction effects, and the use of covariates for norm-referenced outcome measure (GEOFIN).

HO	Level of Effects	Covariate Utilized	Decision Regarding Null Hypothesis
4	Main (Meta)	None	Fail to Reject
4	Main (Meta)	Geometry Pretest	Fail to Reject
4	Main (Meta)	Reading	Fail to Reject
4	Main (Meta)	Geometry Gainscore	Fail to Reject
4	Main (Meta)	Class (intact group)	Fail to Reject
5	Main (Social)	None	Reject
5	Main (Social)	Reading	Fail to Reject
5	Main (Social)	Geometry Pretest	Reject
5	Main (Social)	Geometry Gainscore	Reject
5	Main (Social)	Class	Reject
6	Interaction	None	Fail to Reject
6	Interaction	Reading	Fail to Reject
6	Interaction	Geometry Pretest	Fail to Reject
6	Interaction	Geometry Gainscore	Fail to Reject
6	Interaction	Class	Fail to Reject

The social interaction treatment condition did vary significantly when group means were compared across the various treatment conditions. Main effects were evidenced at the 0.01 level of significance for four of the five measurements completed for scores produced by the groups of students who received the social interaction treatment variable. Only the covariate of reading comprehension failed to provide sufficient evidence to reject the null hypotheses. There was strong evidence to reject null Hypotheses 4 relating to social interaction, since the data significantly differed for group mean scores on the norm-referenced outcome measure.

Interaction effects were not noted above regarding the treatment variable related to the metacognitive strategies. This also had to be true for the treatment condition related to social interaction, since this researcher has already failed to reject null Hypotheses 6 related to the interaction of the treatment variables.

Affective Measure

Null Hypotheses 7, 8, and 9 are related to the affective measure used as a dependent variable in this study. Because this norm-referenced measure of affect was used in various multivariate analyses, Figure 12 is provided to relate the treatment conditions and findings associated with this affective measure.

Main effect differences were found to be significant at the 0.01 level for groups of students that received the metacognitive strategies treatment variable using the geometry pretest covariate. Significant differences at the 0.05 level were evidenced for the metacognitive strategies treatment condition when no covariate was utilized and, also, with the Geometry Final covariate. Based upon this evidence, null Hypothesis 7 was rejected.

Figure 12. Decisions related to null Hypotheses 7, 8, and 9 regarding main effects, interaction effects, and the use of covariates for the measure of affect (CLM).

HO	Level of Effects	Covariate Utilized	Decision Regarding Null Hypothesis
7	Main (Meta)	None	Reject
7	Main (Meta)	Geometry Pretest	Reject
7	Main (Meta)	Geometry Final	Reject
8	Main (Social)	None	Fail to Reject
8	Main (Social)	Geometry Pretest	Fail to Reject
8	Main (Social)	Geometry Final	Fail to Reject
9	Interaction	None	Fail to Reject
9	Interaction	Geometry Pretest	Reject
9	Interaction	Geometry Final	Reject

Regarding the treatment variable of social interaction, no main effect differences were found across the treatment conditions. Social interaction effects were evidenced, however, as part of the interaction effects created by introduction of both treatment variables. The researcher failed to reject null Hypothesis 8.

Interaction effects were found to differ significantly at the 0.05 level for the affective measure when using the covariates of the geometry pretest and the Geometry Final. Since this level of significance differed for the main effects, an analysis of cell means was performed. This process verified that the majority of variance evidenced in the interaction effects was produced by the metacognitive strategies treatment variable. Though there was significance reported at the 0.05 level in two of the three measures of interaction effects, the analysis of the data failed to reject null Hypothesis 9 at the 0.01 level previously established by this researcher.

Measure of Ability

Null Hypotheses 10, 11, and 12 correspond to the ability measure used in this study. As was the case with other norm-referenced measures used in this study, F-ratios were produced through use of multivariate data analyses. Decisions made regarding the measure of ability are found in Figure 13.

Figure 13. Decisions related to null Hypotheses 10, 11, and 12 regarding main effects, interaction effects, and the use of covariates for the measure of ability (DAT).

HO	Level of Effects	Covariate Utilized	Decisions Regarding Null Hypotheses
10	Main (Meta)	None	Reject
10	Main (Meta)	Geometry Pretest	Reject
10	Main (Meta)	Geometry Final	Reject
11	Main (Social)	None	Fail to Reject
11	Main (Social)	Geometry Pretest	Fail to Reject
11	Main (Social)	Geometry Final	Fail to Reject
12	Interaction	None	Fail to Reject
12	Interaction	Geometry Pretest	Fail to Reject
12	Interaction	Geometry Final	Fail to Reject

In the main effects category, the metacognitve strategies treatment variable produced significant differences in group mean scores at the 0.01 level of significance in two of three measured outcome categories. These were the measure of ability with no covariate, and the measure of ability with the geometry pretest covariate. Significant differenecs were found at the 0.05 level on the third measure which used the Geometry Final as a covariate. Null Hypothesis 10 was rejected due to this confirming evidence.

Significant differences were not found in either the interaction effects category, or in the main effects category of social interaction. Thus, null Hypotheses 11 and 12 were not rejected by the analysis of the data.

Measure of Observation

Null Hypothesis 13 related to the observational checklist (OC) developed for use in this study. Videotaping of all the treatment sessions provided a means to verify whether students in a certain social interaction treatment condition did or did not exhibit the expected behaviors.

The viewing of these videotapes required the researcher to use a protocol coding scheme to categorize various behaviors related to the treatment conditions. The use of the protocol coding scheme resulted in the observed student behaviors being categorized as one of the following: 1) giving or sharing information, 2) receiving help or listening, 3) working alone, and 4) off-task.

Two of these four categories (1 & 2) were termed helping behaviors and reflected behaviors that were considered desirable for the social interaction treatment conditions. The latter two categories of observed behaviors, working alone (3) and off-task

(4), are passive behaviors which normally occur when an individual is asked to work alone. These latter two behaviors are expected of students not working in the social interaction treatment conditions.

Observed behaviors were tallied and reported as frequencies and percentages. Students who were working in the social interaction condition exhibited behaviors characterized as Giving or Sharing Information 47.8% of the time, while students not in this treatment condition were observed in this behavior at a rate of only 2.6%. Off-task behavior also differed between the treatment conditions. Students in the social interaction treatment groups were viewed off-task 8.5% of the time, while students not in this treatment condition were observed off-task 16% of the viewing time.

Analysis of the data produced through use of the observational checklist provided two conclusions for the study. First, students assigned to the social interaction treatment condition did exhibit collaborative behaviors. This finding confirmed that the treatment condition of social interaction did in fact exist, and therefore, is an established variable in the interpretation of the data produced by this study. Second, based upon differences in observed student behaviors as evidenced by the videotapes, null Hypothesis 13 was rejected.

Summary of Decisions

Of the thirteen null hypotheses presented in this study, seven were rejected while six of the null hypotheses failed to be rejected. A discussion of conclusions reached regarding the independent variables and treatment conditions based upon these decisions follows.

Discussion of Independent Variables

Metacognitive Strategies

This study used the four strategies identified by Brown and Palinscar (1984) in their reciprocal teaching research to comprise the treatment condition of metacognitive strategies. Decisions reached regarding the null hypotheses which correspond to the use of these metacognitive strategies as a treatment variable are summarized in Figure 14.

Figure 14. Decisions regarding the null hypotheses related to the treatment variable of metacognitive strategies (Meta).

HO	Outcome Measure	Decision Regarding the Hypothesis
1	Criterion-referenced	Reject
4	Norm-referenced	Fail to Reject
7	Affective	Reject
10	Ability	Reject

The effect of metacognitive strategies upon student performance was confirmed in the outcome measures of: 1) criterion-referenced tests of geometry, 2) affect, and 3) ability. Those students who received the independent variable of metacognitive strategies outperformed students who did not receive this variable on three of four outcome measures. Only one outcome measure, the norm-referenced test of geometry, found no significant difference in student performance related to this treatment variable.

<u>Social Interaction</u>

The second independent variable introduced into this study was social interaction. This variable represents a learning activity in which students work together on assigned tasks. A positive interdependence among the group members was required similar to the cooperative learning model of Johnson and Johnson (1984). This variable is also present in both the Vygotskian notion of the zone of proximal development (Vygotsky, 1976) and reciprocal teaching (Brown & Palinscar, 1984). Decisions regarding the null hypotheses related to this treatment variable are listed in Figure 15.

<u>Figure 15.</u> Decisions regarding the null hypotheses related to the treatment variable of social interaction (Social).

HO	Outcome Measure	Decisions Regarding Null Hypothesis
2	Criterion-referenced	Reject
5	Norm-referenced	Reject
8	Affect	Fail to Reject
11	Ability	Fail to Reject
13	Observed Behaviors	Reject

Decisions made regarding the null hypotheses confirmed that the treatment variable of social interaction did make a difference in student performance in geometry as measured by both the criterion-referenced tests and the norm-referenced test. Students in the social interaction treatment condition outperformed students who were not engaged in social interaction.

No differences in measures of ability or effect were noted due to this treatment variable.

Reciprocal Teaching

For purposes of this study, the group of students that received both the metacognitive strategies and social interaction treatment variables comprised the reciprocal teaching condition. Decisions regarding the null hypotheses that correspond to the reciprocal teaching condition are noted in Figure 16.

Of the four null hypotheses related to reciprocal teaching condition of this study, only null Hypothesis 3 was rejected. Students in the reciprocal teaching condition outperformed all other treatment conditions as measured by student performance on the three, criterion-referenced tests of geometry. No differences were deemed significant for either the norm-referenced test of geometry, the measure of affect, or the ability measure.

Figure 16. Decisions regarding the null hypotheses related to the treatment variable of reciprocal teaching (Meta X Social).

HO	Outcome Measures	Decisions Regarding Null Hypothesis
3	Criterion-referenced	Reject
6	Norm-referenced	Fail to Reject
9	Affect	Fail to Reject
12	Ability	Fail to Reject

Results Related to the Treatment Variables

Social Interaction

The treatment variable of social interaction was the only condition which consistently and positively impacted the geometry scores of the students on both the criterion-referenced and norm-referenced measures. It should be noted that social interaction as defined in this study improved student performance in geometry within the regular classroom setting.

Metacognitive Strategies

The metacognitive strategies variable was deemed to significantly effect student performance on both the measure of affect and the measure of ability. When measured at the end of the treatment, students who received the metacognitive strategies variable expressed more confidence in their ability to learn mathematics than those students who did not receive this training. This was also found to be true for the test of abstract reasoning ability, in which students from the metacognitive strategies condition outperformed all other treatment conditions. The metacognitive strategies condition as used in this study contained teacher instruction and modeling in how and when to use these strategies (i.e. summarize, question, clarify, and predict).

Criterion-referenced measures of geometry for the metacognitive treatment condition differed significantly from the treatment condition which did not receive metacognitive strategies as a treatment variable. However, when using the norm-referenced measure, this finding was mixed and inconclusive as this pattern was not apparent when evaluating the effect of metacognitive strategies upon student performance in geometry when using the norm-referenced measure.

Reciprocal Teaching

As previously stated, the treatment condition which received both main effects treatment, contained all of the components normally associated with the interactive teaching method known as reciprocal teaching. The only interaction effect noted in this study between the two main effect treatments (consisting of metacognitive strategies and social interaction)was related to the criterion-referenced tests of geometry. Students from the reciprocal teaching condition outperformed students in all other treatment conditions as measured by the percentage of students who met or exceeded the preset criterion level of performance. No differences were deemed significant on the outcome measures of affect, ability, or the norm-referenced test of geometry for this treatment condition.

Summary of the Results

Social interaction had a positive effect on student performance on both the norm-referenced and the criterion-referenced geometry measures used in this study. Metacognitive strategies had a positive effect upon student performance on the measures of ability and affect that were also used in this study. Reciprocal teaching evidenced significant results on only the criterion-referenced geometry measures. Of these three treatment conditions, social interaction evidenced the most significance in both level and in the frequency of outcome measures.

<u>Limitations</u>

As previously noted, group rivalry effects may have threatened the internal validity of this dissertation. A thorough analysis of the data suggested that no group rivalry exhibited by demoralization of a particular group of students representing one of the four treatment conditions. Therefore, this threat to the internal validity of the study was not realized as a limitation to the study.

This study would still have been a stronger design if four intact groups were utilized, instead of the three groups used by this researcher. The four group design would have eliminated demoralization as a potential threat to the internal validity of the study from the onset of the study. Unlike the research design used, the four group design would have allowed all students from a particular intact group to comprise a single treatment condition.

Instrumentation was not deemed a limitation of the study, nor was sample size or length of treatment. For these reasons, no other issues regarding limitations of this study are reported.

<u>Generalizations</u>

The reciprocal teaching model is comprised of the following three components: 1) scaffolded instruction, 2) comprehension-fostering and comprehension-monitoring strategies, and 3) cooperative learning (Brown & Palinscar, 1984). The present study combined the first two components of reciprocal teaching to form the metacognitive strategies treatment condition. Social interaction as used in this study matches the third component of reciprocal teaching; namely, cooperative learning.

The social interaction variable had a significant impact upon student performance within a normal classroom setting. This unobtrusive variable should be targeted for both further research and immediate application in other school settings. The findings of this study confirm teaching collaborative skills to students can be accomplished by simply embedding these skills within instruction of a particular content lesson. Furthermore, small group activities are recommended for inclusion in all regular instructional programs regardless of the content areas.

Though the metacognitive strategies component was not evidenced by student performance in geometry, it is not without merit. Students who received training in when and how to use metacognitive strategies did exhibit a higher level of confidence in regards to learning mathematics, as evidenced by scores produced on the measure of affect used in this study. However, the metacognitive strategies variable did not significantly impact student performance in geometry but, did impact the abstract reasoning ability measure. Both these constructs of affect and ability are recognized as predictors of future performance in mathematics. It could be that the test items of the geometry measures did not require the same abilities or knowledge that was required to successfully complete the abstract reasoning test. For these reasons, the metacognitive strategies component is not ruled out for further study.

Recommendations

Three suggestions for further research within the realm of mathematics teaching and learning are proposed by this researcher. These areas of study are as follows: 1) metacogniton, 2) cooperative learning, and 3) the reciprocal teaching method.

First, metacognition occurs in a normal classroom environment even if teachers and students do not specifically address the concept. For this reason, metacognition is still recognized as an important variable for classroom research. Training in the four strategies used in this study (i.e. summarize, question, clarify, and predict) in conjunction with scaffolded instruction may or may not have been as specific and intensive as needed to effect student performance in mathematics. These four strategies may not have all similarly effected student learning. Future research that utilizes the same four strategies of reciprocal teaching probably requires a design that allows the relative effects of each of the four strategies upon performance to be studied separately and together. In this way, one or more of the four strategies may be found to be more useful in effecting student performance and learning. Other metacognitive strategies noted in the extant literature are also worthy of research in a regular classroom setting.

Second, since social interaction did significantly influence student performance in geometry; it merits future research. Studies relating cooperative learning behaviors across various content areas may discover a specific behavior, classroom activity, or subject that is particularly responsive to this treatment variable. Application of social interaction, and other cooperative learning models to the normal classroom environment is appropriate at this time. Further study at the junior high/middle school level is recommended.

Third, reciprocal teaching did provide a conceptual model in which to study both metacognition and social interaction in a field setting. Though the results of the use of the reciprocal teaching condition of this study were not found to be as effective as social interaction alone, it still should and can be applied to other field settings. Further studies which attempt to use the reciprocal teaching approach in a large group setting are needed. Results from similar studies in this area will aid in determining how the three components of the model can be applied to large groups. It should be noted that in the realm of a school setting, individual, small group, and large group models of instruction will always be necessary to meet the needs of the students and the learning situation. Continued research in this field will contribute to the knowledge base now available relating to learning in a group setting.

Final Remarks

This dissertation incorporated a blend of works from many different fields of study in an attempt to best answer the research questions. These questions also demanded a research design of quasi-experimentation in order to study the teaching and learning process as it naturally occurs in the school setting. Continued research within the domains of cognitive science, information-processing theory, mathematics education, and reading research can and should be conducted in the real world of the classroom. For it is where student learning and performance can be observed, verified, and ultimately improved.

REFERENCES

References

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. David Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), Handbook of Reading Research (pp. 353-394). New York: Longman.

Brophy, J. (1986). Teaching and learning mathematics: where research should be going. Journal for Research in Mathematics Education, 17(5), 323-346.

Brown, A. L., Armbruster, B. B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Orasanu (Ed.), Reading Comprehension: From research to practice (pp. 49-76). Hilllsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: on training students to learn from texts. Educational Researcher, 10(2), 14-21.

Brown, A. L., Day, J. D., & Jones, R. (1983). The development of plans for summarizing text. Child Development, 54, 968-979.

Brown, A. L., & Palinscar, A. S. (1982). Introducing strategic learning from texts by means of informed, self-control training. Topics in Learning and Learning Disabilities, 2, 1-17.

Brown, A. L., Palinscar, A. S., & Armbruster, B. B. (1984). Instructing comprehension-fostering activities in interactive situations. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), Learning and Comprehension of Text (pp. 255-286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Campione, J. C., & Armbruster, B. B. (1985). Analysis-acquiring information from texts: an analysis of four approaches. In J. W. Segal & S. F. Chipman (Eds.), Thinking and Learning Skills (Vol.1): Relating instruction to research (pp. 317-362). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chipman, S. F., & Segal, J. W. (1985). Higher cognitive goals for education: an introduction. In J. W. Segal & S. F. Chipman (Eds.), Thinking and Learning Skills (Vol. 1): Relating instruction to research (pp. 1-20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, T. D., & Campbell, D. T. (1979). Quasi-Experimentation Design & Analyses Issues for Field Settings. Boston: Houghton Mifflin.

Dansereau, D. F. (1985). Learning strategy research. In J. W. Segal & S. F. Chipman (Eds.), Thinking and Learning Skills (Vol. 1): Relating instruction to research (pp. 209-240). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fennema, E. & Sherman J. (1976). Fennema-Sherman Mathematics Attitude Scales. Washington D. C.: American Psychological Association.

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

Garofalo, J. (1987). Metacognition and school mathematics. Arithmetic Teacher, May 1987, 22-23.

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

Haller, E. P., Child, D., & Walberg, J. (1988). Can comprehension be taught? a quantitative synthesis of "metacognitive" studies. Educational Researcher, 17(9), 5-8.

Johnson, D. W., Johnson, R. T., Holubec, E. J., & Roy, P. (1984). Circles of Learning. Alexandria, VA: ASCD.

Johnson, D. W., Johnson, R., & Stanne M. (1986). Comparison of computer-assisted cooperative, competitive, and individualistic learning. American Educational Research Journal, 23(3), 382-392.

Johnson, D. W., Skon, L., & Johnson, R. (1980). Effects of cooperative, competitive, and individualistic condition's on children's problem-solving performance. American Educational Research Journal, 17(1), 83-93.

Kerlinger, F. N. (1973). Foundations of Behavioral Research. New York: Holt, Rinehart and Winston.

Lester, F. K. (1983). Trends and issues in mathematical problem-solving research. In R. Lesh & M. Landau (Eds.), Acquisition of Mathematical Concepts and Processes (pp. 229-258). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. (1986). Positive interdependence, academic and collaborative-skills group contingencies, and isolated students. American Educational Research Journal, 23(3), 476-488.

McKeachie, W. J., Pintrich, P. R., & Lin, Y. (1985). Teaching learning strategies. Educational Psychologist, 20(3), 153 -160.

Miles, M. B., & Huberman, A. M. (1984). Qualitative Data Analyses. Beverly Hills, CA: Sage.

Noddings, N. (1985). Small groups as a setting for research on mathematical problem solving. In E. A. Silver (Ed.), Teaching and Learning Mathematical Problem Solving: Multiple research perspectives (pp. 345-356). Hillsdale, NJ: Lawrence Erlbaum Associates.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. Educational Psychologist, 21(1 & 2), 73-98.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, 1(2), 117-175.

Palinscar, A. S., & Brown, A. L. (1986). Interactive teaching to promote reading in text. The Reading Teacher, 39(8), 771-777.

Polya, G. (1957). How to Solve it. Princeton, NJ: Princeton University.

Popham, J. W., & Husek, T. R. (1969). Implications of criterion-referenced research, Journal of Educational Measurement, 6(1), 1-9.

Pressley, M. (1986). The relevance of the good strategy user model to the teaching of mathematics. Educational Psychologist, 21(1 & 2), 139-161.

Pressley, M., & Levin, J. R. (Eds.). (1983). Cognitive Strategy Training: Educational applications. New York: Springer -Verlag.

Psychological Corporation (1982a). The Differential Aptitude Test. San Antonio, TX: Author.

Psychological Corporation (1982b). The Stanford Achievement Test Series. San Antonio, TX: Author.

Reisman, F. K., & Hutchinson, T. (1985). Sequential Assessment of Mathematics Inventories. Columbus, OH: Merrill.

Resnick, L. B. (1984). Comprehending and learning: implications for a cognitive theory of instruction. In H. Mandl (Ed.), Learning and Comprehension of Text (pp. 431-444). Hillsdale, NJ: Lawrence Erlbaum Associates.

Resnick, L. B., & Ford, W. W. (1981). The Psychology of Mathematics for Instruction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1985). Making sense of "out loud" problem-solving protocols. The Journal of Mathematical Behavior, 4, 171-195.

Silver, E. A. (1985). Research on teaching mathematical problem solving: some underrepresented themes and needed directions. In E. A. Silver (Ed.), Teaching and Learning Mathematical Problem Solving: Multiple research perspectives (pp. 247 -266). Hillsdale, NJ: Lawrence Erlbaum Associates.

Slavin, R. E. (1983). Cooperative Learning. New York: Longman.

Slavin, R. E. (1987). Cooperative learning: where behavioral and humanistic approaches to classroom motivation meet. The Elementary School Journal, 88(1), 29-37.

Tolman, E. C. (1959). Principles of purposive behavior. In S. Koch (Ed.), Psychology: A study of science (Vol.2) (p.152). New York: McGraw-Hill.

Vygotsky, L. (1978). Mind and Society: The development of higher psychological processes. Cambridge, MA: Harvard University.

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. Educational Psychologist, 15(2), 69-83.

Webb, N. M. (1982a). Group composition, group interaction, and achievement in cooperative small groups. Journal of Educational Psychology, 74(4), 475-484.

Webb, N. M. (1982b). Peer interaction and learning in cooperative small groups. Journal of Educational Psychology, 74(5), 642-655.

Webb, N. M. (1982c). Student interaction and learning in small groups. Review of Educational Research, 52(3), 421-445.

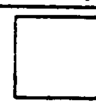
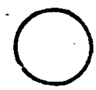
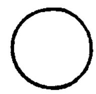
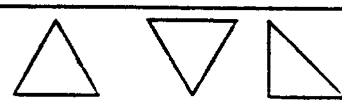
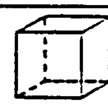
Weinstein, C. F., & Mayer, R. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), The Handbbok of Research on Teaching (pp. 315-327). New York: Macmillan, 1986.

Weinstein, C. F., & Underwood, V. L. (1985). Learning strategies: the how of learning. In J. W. Segal & S. F. Chipman (Eds.), Thinking and Learning Skills (Vol. 1): Relating instruction to research (pp. 241-258). Hillsdale, NJ: Lawrence Erlbaum Assosciates.

Wittrock, M. (1978). The cognitive movement in instruction. Educational Psychologist, 13(#), 15-29.

Wittrock, M. (1986). Handbook of Research on Teaching. New York: Macmillan.

APPENDIX A

Item	Directions	Correct Response	Score
G-1	Point to the inside of the figure.		+ – 0
G-2	Point to the shape that comes next.		+ – 0
G-3	Point to the circle.		+ – 0
G-4	Point to all of the triangles.		+ – 0
G-5	Point to the object that is like this shape.		+ – 0
G-6	Point to the figure that has corners.		+ – 0
G-7	Point to the closed figure.		+ – 0
G-8	Draw the same shape and size.	a square 4 dots by 4 dots, placed anywhere on grid	+ – 0
G-9	Tell if the name of the place where two lines intersect is a line, a plane, or a point.	a point	+ – 0
G-10	Tell if the two lines are parallel or perpendicular.	parallel	+ – 0
G-11	Tell if the name of the kind of angle shown is obtuse, acute, or straight.	acute	+ – 0
G-12	Tell if the two figures are congruent or not congruent.	not congruent	+ – 0
G-13	Tell if the two figures are congruent or not congruent.	congruent	+ – 0
G-14	Point to the equilateral triangle.		+ – 0
G-15	Say the letter or letters for the center of the circle.	P	+ – 0
G-16	Point to the circumference of the circle.	Student should point to some portion of the set of points that comprise the circle.	+ – 0
G-17	Say the letter or letters for the diameter of the circle.	CD or DC or CPD or DPC	+ – 0
G-18	Say the letter or letters for the radius of the circle.	PD or DP or CP or PC or AP or PA	+ – 0
G-19	Point to the parallelogram.		+ – 0
G-20	Point to the pentagon.		+ – 0
G-21	Say the letter or letters for the chord of the circle.	AB or BA	+ – 0

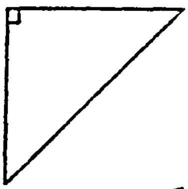
APPENDIX B

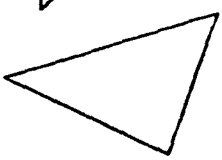
Name: ______________________________

Geometry Final Test 5/29/87

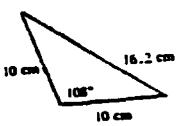
Tell whether each of these triangles is right, acute or obtuse.

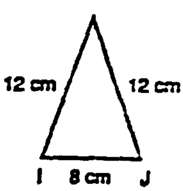
1. 

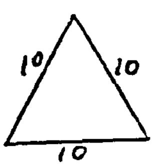
2. 

3. 

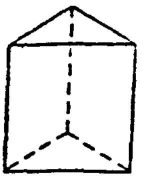
Tell whether each of these triangles is equilateral, scalene, or isosceles.

4. 

5. 

6. 

Use this triangular prism to list the number of:

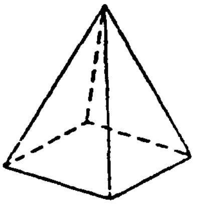


7. faces

8. edges

9. vertices

Use this square pyramid to list the number of:



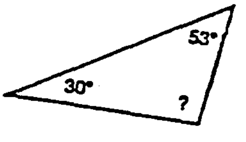
10. faces

11. edges

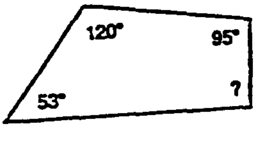
12. vertices

Find the measure of the missing angle in each of the following figures.

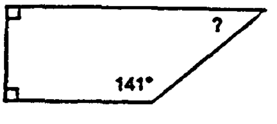
13.



14.

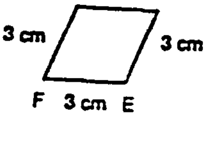


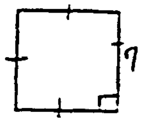
15.

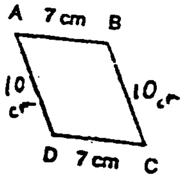


Use the terms, rhombus, rectangle, parallelogram, trapezoid and square to describe the quadrilaterals. List all of the above terms that apply to each figure.

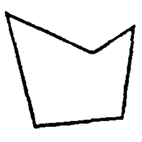
16.



17. 

18. 

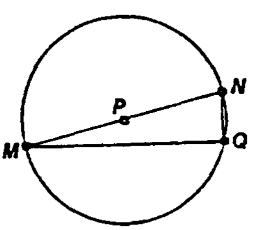
Write the name of the following figures.

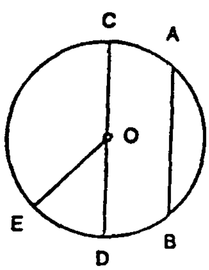
19. 

20. a seven-sided polygon

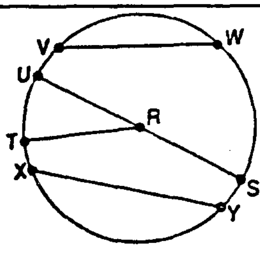
21. a twelve-sided polygon

List two radii for each of the following figures.

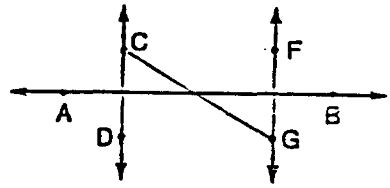
22. 

23. 

24. List a central angle from this figure.



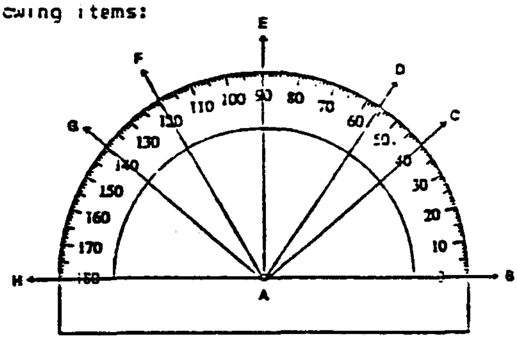
Use this figure to list the following:



25. A line perpendicular to line FG

26. A line parallel to line FG

Use the picture of the protractor for the following items:



27. What is the measure of angle CAE?

28. What is the measure of angle HAC?

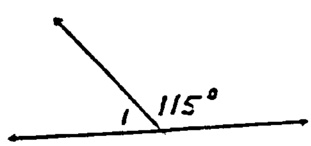
29. What kind of angle is HAB?

30. Name an angle that is complementary to angle FAH.

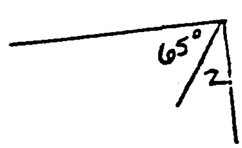
31. Name an angle that is supplementary to angle HAC.

Use the following figures to find the measures of the missing angles.

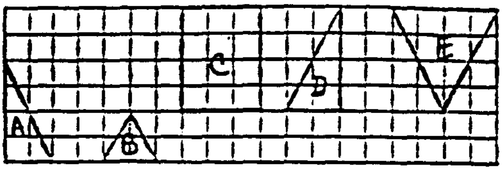
32. $\angle 1 =$ ________ $^{\circ}$



33. $\angle 2 =$ ________ $^{\circ}$



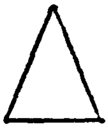
34. Use this figure to list a pair of similar figures.



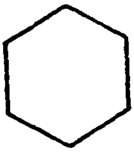
How many lines of symmetry can be found in the following figures. List all possible answers.

35. A square?

36. An isosceles triangle?

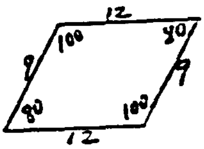


37. A regular hexagon?

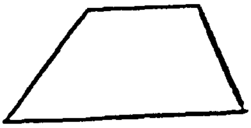


38. What's another name for a regular quadrilateral?

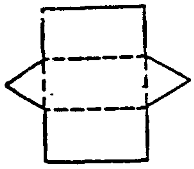
39. Is this figure a regular polygon?

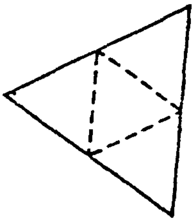


40. Use this figure to draw a diagonal.

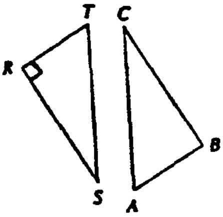


Fold these figures along the dotted lines. Will they fold into a prism or a pyramid?

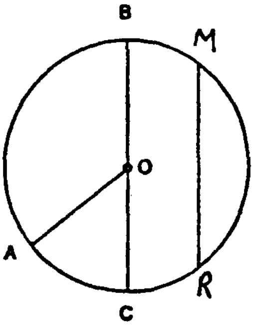
41. 

42. 

43. Name three pairs of congruent sides, and three pairs of congruent angles in this figure.



44. In this figure line segment MR is also known as a ____________________



Explain in words, the difference between:

45. A rectangle and a square:

46. A line and a line segment:

47. A rhombus and a square:

48. A trapezoid and a parallelogram:

49. An arc and a circle:

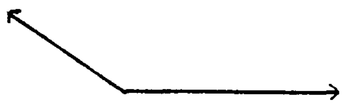
50. The terms similar figures and congruent figures:

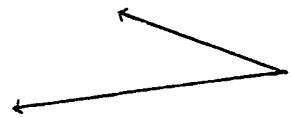
APPENDIX C

CRT1

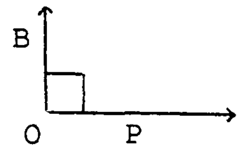
Name: ______________________________

Label these angles as: right, acute, obtuse or straight.

1. 

2. 

3. Name the vertex of this angle on the right.



4. Name the angle on the right 3 different ways.

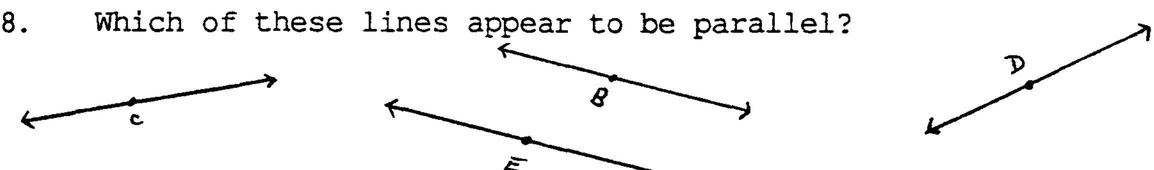
__________ __________ __________

5. What's the name of this figure?

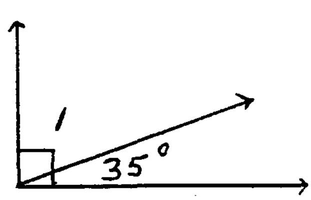
6. Write line segment AB using symbols.

7. Give an example of a ray.

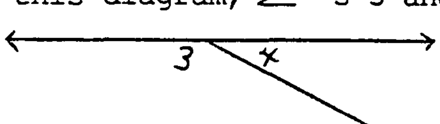
8. Which of these lines appear to be parallel?



9. Find the measure of ∠ 1.



10. In this diagram, ∠ 's 3 and 4 are called by what term?



CRT2

Name: ____

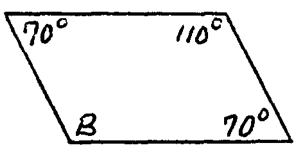
In words, describe the difference(s) between the following polygons:

1. Trapezoid/Parallelogram

2. Rectangle/Square

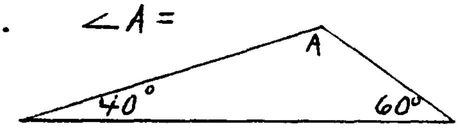
3. Rhombus/Square

Find the measures of the missing angles.

4. 

$\angle B =$

5. $\angle A =$



Answer right △, acute △, or obtuse △.

6.

7. For the same △, answer scalene, isosceles, or equilateral.

8. A polygon is a closed figure made up of line segments. Draw a figure made up of line segments that is <u>not closed</u>.

9. A pentagon has how many sides?

10. A nine-sided polygon is called what?

CRT3

Name: ____

Since △ ABC ≅ △ RST, then:

1. Name 3 pairs of ≅ ∠'s.

2. Name 3 pairs of ≅ sides.

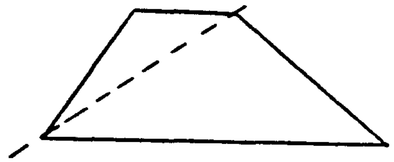
3. What's the length of $\overline{RS}$?

4. What's the measure of ∠RST?

5. What's the length of $\overline{ST}$?

6. What's the length of $\overline{RT}$?

7. Is this a line of symmetry?

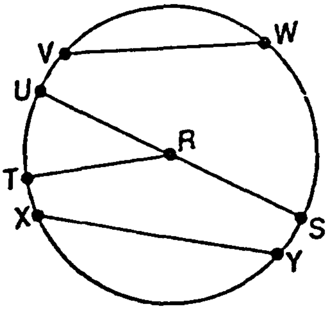


Name a:

8. Radius

9. Diameter

10. Central ∠



APPENDIX D

****

OBSERVER'S CHECKLIST

Session #: ____________ Date: ________________

Giving or Sharing Information	Receiving Help or Listening	Working Alone	Off-task

****

APPENDIX E

Frequency Data for Observer's Checklists (Videotapes)

	Helping Behaviors		Non-helping Behaviors	
Treatment Condition	Giving or Sharing Information	Receiving Help or Listening	Working Alone	Off-task
Social Interaction	344	106	209	61
No Social Interaction	19	65	521	115

APPENDIX F

Name: ____________________________

Confidence in Learning Mathematics Scale

1.	Generally I have felt secure about attempting mathematics.	SA	A	U	D	SD
2.	I am sure I could do advanced work in mathematics.	SA	A	U	D	SD
3.	I am sure that I can learn mathematics.	SA	A	U	D	SD
4.	I think I could handle more difficult mathematics.	SA	A	U	D	SD
5.	I can get good grades in mathematics.	SA	A	U	D	SD
6.	I have a lot of self-confidence when it comes to math.	SA	A	U	D	SD
7.	I'm no good in math.	SA	A	U	D	SD
8.	I don't think I could do advanced mathematics.	SA	A	U	D	SD
9.	I'm not the type to do well in math.	SA	A	U	D	SD
10.	For some reason even though I study, math seems unusually hard for me.	SA	A	U	D	SD
11.	Most subjects I can handle O.K., but I have a knack for flubbing up math.	SA	A	U	D	SD
12.	Math has been my worst subject.	SA	A	U	D	SD

APPROVAL SHEET

The dissertation submitted by Dennis W. Rudy has been read and approved by the following committee:

Dr. Todd J. Hoover, Director
Associate Professor, Curriculum and Human Resource Development, Loyola

Dr. John A. Dossey, External Reader
Professor, Mathematics, Illinois State University

Dr. Kay Monroe Smith
Assistant Professor, Curriculum and Human Resource Development, Loyola

Dr. Ronald R. Morgan
Associate Professor, Counseling and Educational Psychology, Loyola

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 5, 1989
Date

Todd Hoover
Director's Signature