An Investigation Of High School Geometry Students Proving And Logical Thinking Abilities And The Impact Of Dynamic Geometry Software On Student Performance

Lalitha Subramanian
University of Central Florida

AN INVESTIGATION OF HIGH SCHOOL GEOMETRY STUDENTS' PROVING AND LOGICAL THINKING ABILITIES AND THE IMPACT OF DYNAMIC GEOMETRY SOFTWARE ON STUDENT PERFORMANCE

by

LALITHA SUBRAMANIAN
M.Sc. University of Mumbai, 1991

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Teaching and Learning Principles
in the College of Education
at the University of Central Florida

Orlando, Florida

Summer Term
2005

Major Advisor: Michael C. Hynes

ABSTRACT

The purpose of this study was to investigate (a) the role of a yearlong geometry course on high school geometry students' logical thinking and proof construction abilities, (b) the linkage between students' logical thinking and proof construction abilities, and (c) the impact of dynamic geometry software on students' performance. In addition, this study also ventured to determine if the type of geometry course had any impact on students' logical thinking and proof construction achievement. The sample for the study consisted of 1,325 high school geometry students enrolled in regular, honors, and mastery courses in four high schools from the school district affiliated with the Local Education Agency (LEA) during the academic year 2004-2005. Geometer's Sketchpad™ (GSP) was assumed to represent the dynamic geometry software. Responses of students on two pre-tests and two post-tests, each with one on logical thinking and one on proof, were analyzed to address the research questions. Results of the analyses indicated no significant effect of the yearlong geometry course on the performance of students on proof tests but a fairly significant effect on the tests of logical thinking. Use of GSP was found to have some impact on honors and mastery students' performance on proof post-tests. Honors students showed a higher logical thinking level than their regular and mastery counterparts in both GSP and non-GSP groups. There was a significant positive correlation between students' performance on the tests of logical thinking and proof.

DEDICATION

I dedicate this dissertation to my wonderful husband Subu, whose continuous support and encouragement made it possible for me to complete my courses and this research project. I credit him the major portion of the doctoral degree I have earned by completing this dissertation. He has always been there for me and has extended a helping hand in every possible way. He has been patient with me through the highs and lows of the past three years, and has sacrificed his own career so that I could attain this degree.

ACKNOWLEDGMENTS

I sincerely thank my advisor and chair of my dissertation committee Dr. Michael Hynes without whose guidance, support, and encouragement, I would not have been able to realize my dream of attaining my doctoral degree. He has always been a great source of inspiration to me. I place on record my gratitude to the co-chair Dr. Enrique Ortiz for his spontaneous assistance and to the other committee members, Dr. Juli Dixon, Dr. Lea Witta, and Dr. Ram Mohapatra, for their support and guidance. I would also like to thank Dr. Aldrin Sweeny for helping me contact Dr. Kenneth Tobin, and Dr. Kenneth Tobin for making the use of their TOLT instruments possible. I am grateful to Dr. Venketsamy for willingly consenting to pilot test my proof test instruments. I record my appreciations and gratitude to Gayle Sitter for her goodwill and assistance in establishing contacts with the participating schools. I also thank the mathematics department chairs and geometry teachers of the participating schools for consenting to administer the tests to their classes and their support in making my project a success.

I thank Dr. Zalman Usiskin, Dr. Eric Knuth, Dr. Sharon Senk and Dr. Tami Martin for reviewing my test instruments and providing invaluable feedback that improved their validity and consistency. My heartfelt gratitude goes out to Dr. Douglas Brumbaugh for the innumerous and rewarding conversations, which boosted my self-confidence and paved the way for the successful completion of my study. I thank my fellow doctoral students Janet Andreasen, George Roy, other graduate students, and Dr.V.M. Kadam, who spent weekends and holidays grading my test papers. My special thanks to Bonnie Swan for her help in identifying appropriate technology literature.

I thank the faculty and staff of the Department of Teaching and Learning Principles for being so helpful and supportive throughout the period of my study. I record my appreciation to Carine Strebel Halpern for all the help she has extended during the three years of my study. She has always been ready and willing to offer a helping hand whenever I approached her with a problem. I also wish to place on record my appreciation to Dr. Nancy Lewis for her friendship. My numerous interactions with her had a positive influence on me. I have been fortunate to get the support and cooperation from all faculty and office staff in the College of Education, particularly the tech service people. I place on record my appreciation and gratitude to all.

Special thanks to Pallavoor Vaidyanathan who has most appropriately directed me to College of Education at the University of Central Florida. Last but not the least, I thank my family for their wonderful support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

GSP	GEOMETER'S SKETCHPAD
TOLT	TEST OF LOGICAL THINKING
LEA	LOCAL EDUCATION AGENCY
R	REGULAR GEOMETRY COURSE
M	MASTERY GEOMETRY COURSE
H	HONORS GEOMETRY COURSE

CHAPTER 1: INTRODUCTION

Overview

If mathematics can be referred to as both queen and servant of the sciences, proof can be considered both the queen and servant of mathematics (Mingus & Grassl, 1999). The concept of proof is of great importance to mathematics in general, and to geometry in particular. As Mingus and Grassl pointed out, geometry is a wonderful area for developing logical reasoning. Investigating, making conjectures, and developing logical arguments to justify conclusions are important aspects of studying geometry. Developing analytical thinking in students through the process of logical reasoning has been one of the major objectives of integrating proofs into the geometry curriculum for more than a century. Traditionally, high school geometry was considered the most appropriate vehicle to acquaint students with deductive reasoning through the methods of proving geometric facts and to familiarize them with the rigors of formal logical reasoning.

The Principles and Standards for School Mathematics (National Council of Teachers of Mathematics [NCTM], 2000) emphasizes the importance of technology in teaching and learning mathematics. Specifically, the Standards document advocates use of dynamic geometry technology to teach geometry at all levels of mathematical education, emphasizing that the dynamic interactive nature of this technology helps in engaging students in meaningful mathematical activities and promotes deeper understanding of concepts. A student who understands the meaning and significance of geometric properties will be ready for the formal deductive reasoning to prove these properties. Various dynamic geometry software packages like Geometer's Sketchpad™, Cabri™, and Geomaster™ are available on discs for installation on computers as well as online for download that could be used in classrooms. In addition, recent versions of calculators like TI 83 Plus of *Texas Instruments* come loaded with some of these packages. Thus dynamic geometry technology is made accessible to teachers for use either for demonstration or for planning individual/small group work.

In spite of recent reform initiatives advocating use of dynamic geometry software to teach high school geometry and the many studies confirming their effectiveness in mathematical teaching and learning, very few teachers use dynamic geometry technology in their instructional processes, as was revealed by the preliminary survey of high school geometry teachers conducted by the researcher. This survey found that the main obstacles teachers cite for non-use of dynamic geometry include: (a) Non-availability of necessary accessories for use of the software in classrooms; (b) Lack of sufficient training in the use of the software, resulting in reduced confidence level; and (c) Time constraint, due to the need for completing the course work within the allotted time frame. The survey and findings are discussed in detail in Chapter 2.

This study addressed the issue of teaching proofs in high school geometry courses. In particular, it investigated (a) the role of a yearlong geometry course in high school geometry students' logical thinking and proof construction abilities; (b) the relationship between students' logical thinking and proof construction abilities; and (c) the impact of dynamic geometry software in students' performance. Geometer's Sketchpad (GSP), a dynamic interactive geometry software commonly used by those geometry teachers who use the dynamic geometry technology in classroom instruction, was assumed to represent dynamic geometry software.

Background for the Study

The conceptual framework for this study was derived from the relevant research literature as well as from the researcher's tacit theory. The concept of proof is of great importance in the study of mathematics (Martin & Harel, 1989). Smith and Henderson (1989) stated, for example, "The idea of proof is one of the pivotal ideas in mathematics. It enables us to test the implications of ideas, thus establishing the relationship of the ideas and leading to the discovery of new knowledge" (p. 178). *The Principles and Standards of School Mathematics* (NCTM, 2000) states, "…by the end of secondary school, students should be able to fully understand and produce mathematical proofs" (p. 56). Geometry has long been regarded as a vehicle for developing the logical reasoning capabilities of students. Engagement in proving activities encourages students to think and reason logically and also to understand the axiomatic structure of mathematics. As Polya (1957) indicated, "… if general education intends to bestow on the student the ideas of intuitive evidence and logical reasoning, it must reserve a place for geometric proofs" (p. 217).

Proving is intimately connected to the building of mathematical ideas. So, proving should be a natural activity for students, just like the other mathematical activities of defining, conjecturing, reasoning, and generalizing (Battista & Clements, 1996). The various geometric facts and theorems may have little use to the student in later life, but the problem solving capability he/she develops through the rigor of mathematical problems and proofs remain inherent, which would certainly make him/her a better and more successful problem solver later in life. In spite of all these benefits, there is a prevalent and persistent question lingering in the minds of high school mathematics teachers and students about why we should learn or teach proofs. It is a common experience of mathematics teachers the world over that students perceive geometric proofs as one of the most difficult tasks in their high school geometry course.

Many educators are divided in their opinion about the extent to which formal proof should be emphasized in high school geometry (Battista & Clements, 1996; Hanna, 1996). While some support traditional axiomatic systems and rigorous proof, others believe that geometry should involve less formal investigation of geometric ideas and some educators believe in a gradual transition from informal investigation to a more formal proof (Battista & Clements). According to Battista and Clements, formally presenting the arguments in terms of proofs is a method of establishing the validity of ideas. As for Hanna, the main objective of proof in the mathematics classroom is to promote understanding rather than just to expose students to the rigors of mathematical reasoning.

A. Dennis (2000) conducted a survey of mathematics students' interactions with proof, using a mixed mode method. Her findings brought to light difficulties encountered by undergraduate students while learning proofs in college mathematics courses. The students who participated in this study had completed very few proof activities in their high school level or advanced level mathematics courses. A. Dennis's findings attest to the importance of teaching proof in high school mathematics courses. However, with regard to the effect of doing proofs in developing students' formal reasoning skills, conflicting results have been observed in studies conducted at different times. For example, Ireland's (1973) investigation of high school students' proof writing ability and formal deductive reasoning ability indicated a low correlation, while another study by Stover (1989) indicated a low correlation between inductive reasoning and proof writing abilities and greater correlation between deductive reasoning and proof writing abilities. Ireland (1973) also noted that a student who could write a formal proof very well might not have the skill in deductive reasoning. Likewise, a student with excellent formal deductive reasoning ability might not be able to write a formal proof well. The results of his study indicated three types of logical reasoning that could contribute to students' ability to construct proofs: conditional reasoning, class reasoning, and ordinal reasoning. He defined conditional reasoning as reasoning associated with arguments based on conditional statements, class reasoning as arguments based on quantifiers, and ordinal reasoning as arguments involving physical relationships such as length, area, weight, speed, etc. (Ireland, 1973).

Greeno (1980) gave a broad definition to the geometric knowledge necessary to construct geometric proofs in addition to formal reasoning ability. He described four types of geometrical knowledge: (a) General knowledge (definitions, meaning of terms, etc.); (b) Visual knowledge (image of corresponding angles, vertical angles, opposite vertices, etc.); (c) Inferential knowledge (theorems, axioms, etc.); and (d) Knowledge of strategies (tests of congruency, transitivity, reflexivity, etc.). McBride and Carifio (1995) categorized the first three types as declarative knowledge and the last type as procedural knowledge. According to Heinze and Kwak (2002), declarative knowledge is necessary but not sufficient for success in proof writing. In addition to geometric knowledge, students need attention, perception, memory-search, heuristic, and logical reasoning to achieve success in proof writing (McBride & Carifio, 1995).

The development of logical reasoning ability was initially studied by Piaget (1957), and has since been extensively researched (Tobin & Capie, 1980a). Most of the studies confirmed the importance of logical reasoning ability as a factor of cognitive achievement (Tobin & Capie). In particular, Tobin and Capie studied and developed valid instruments to test five logical reasoning components--proportional, combinatorial, probabilistic, correlational reasoning, and control of variables. Saouma BouJaoude (2004) termed these components as formal operational reasoning and investigated the relationships between these five components and students' ability to solve chemistry problems. BouJaodes's study revealed that formal operational reasoning was a major predictor of performance on conceptual chemistry problems. BouJaoude also cited studies conducted by several researchers, which indicated that formal operational reasoning was the main predictor of performance on different types of science problems. Researchers like Stover (1989) and Knuth (1999) have suggested a strong relation between logical reasoning and proof writing ability of students. The study conducted by Fawcett (1938) confirmed the dependence of the degree of transfer of the logical training brought about by the proof-oriented geometry coursework on the efficiency of instruction. The longitudinal proof project of McCrone, Martin, Dindyal, and Wallace (2002) confirmed the positive correlation between students' logical reasoning ability and proof writing ability.

The various studies discussed above support the researcher's tacit theory that by developing the formal logical reasoning, better understanding of the nature of deductive reasoning and proof could be fostered in students and vice versa. To the researcher's knowledge, research studies investigating the relationship between the formal operational reasoning modes and the proof construction ability of students are very few. Experience in teaching mathematics has convinced the researcher that focusing on the “why” aspect of geometric properties in addition to the “what” and “how” aspects helps orient students’ logical reasoning abilities and to develop their proof construction abilities. While informal proving and justifying can be fostered through the engagement of students in argumentations of the why of statements, developing the ability for proof-construction requires challenging students in deductive argumentation.

Dynamic geometry technology is a very useful tool that can be exploited in the classroom instructional processes to develop formal reasoning in students (Galindo, 1998). The highly visual nature of such interactive tools helps make proving activities more meaningful to students. Moreover, new approaches to proof and justification could be designed to successfully motivate students and develop an appreciation to the nature and functions of geometric proofs. The interactive nature of dynamic geometry software allows students to create simple geometric figures, explore relationships, make conjectures about their properties, and test those conjectures (Galindo). Rasmussen (1992) categorized proof tasks that could be done using sketchpad into two broad types: tasks where the insight is derived from the dynamics of the model and tasks where insight comes from the process of construction.

Roberts and Stephens (1999) investigated the effectiveness of Geometric Supposer™, a dynamic interactive computer software used in the 1980s. Their study showed that students who were exposed to the software developed positive attitudes toward geometry, were more willing to explore concepts, and obtained higher scores on higher level application problems (Roberts & Stephens). Explorations using dynamic geometry tools provide rich and varied mathematical experiences that would attract students of different talents and ability levels, resulting in enhanced learning outcome for all students in the class. The studies and views discussed above support the researcher's tacit theory that use of dynamic geometry technology in classroom instructional processes could help improve students' formal reasoning and proof constructing abilities of students.

Research Questions

The research presented here was an effort to investigate (a) the role of a year-long geometry course on high school geometry students' logical thinking ability and their proof-construction ability; (b) the linkage between students' formal logical reasoning ability and their ability to construct proofs; and (c) the effectiveness of the use of dynamic geometry technology on student performance. Because the abilities under investigation were quite complex, the researcher ventured to determine a measure of these abilities by means of students' performance on two pre-tests and two post-tests. Part I of the pre- and post- tests were tests of logical thinking and Part II of both tests were proof tests. Moreover, due to the nature of the variables involved, a detailed descriptive analysis was also done to supplement the results obtained through statistical analyses. The study was grounded in the following questions:

1. What is the effect of a yearlong geometry coursework on the performance of high school geometry students on a proof test?
2. Does the use of dynamic geometry software in the instructional process influence the performance of high school geometry students on a proof test?
3. Does the type of geometry coursework influence the performance of high school geometry students on a proof test?

4. What is the effect of a yearlong geometry coursework on the performance of high school geometry students on a test of logical thinking?
5. Does the use of dynamic geometry software in the instructional process influence the performance of high school geometry students on a test of logical thinking?
6. Does the type of geometry coursework influence the performance of high school geometry students on a test of logical thinking?
7. Is there a relationship between students' performance on tests of logical thinking and proof, based on a yearlong geometry coursework?
8. How well do high school geometry students learn to construct geometry proofs by the end of the high school geometry course?

Purpose and Significance of the Study

Proof constructing is a process that involves a chain of reasoning, using the given information, relevant diagrams, axioms, postulates, theorems, etc., logically leading to the conclusion (Stover, 1989). It is, therefore, natural to expect that students' engagement in proving activities would help in developing their formal reasoning abilities. The extent to which this result is achieved in the course is dependent upon various factors such as the curriculum, instructional process, and socio-mathematical norms set in the classroom to encourage and challenge students to engage in meaningful mathematical discourse and argumentation.

Current mathematics and science reform initiatives proposed as part of the No Child Left Behind Act (2001) focus on restructuring high school mathematics courses so as to prepare students for the challenging jobs and other vocational opportunities of 21st Century

America. Some of the proposals include requiring two more years of state-mandated mathematics tests in high school grades (CNN, 2004). The President's current reform proposal also includes testing 12^{th} graders in the National Assessment of Educational Progress (Whitehouse, 2004). Several states are currently considering revision of more rigorous curricula for their high school mathematics courses (Olson, 2005). The reform efforts are essentially geared towards improving high school standards and accelerating the mathematics achievements of high school students. Some states like Texas, Arkansas, and Indiana have already included geometry in the list of high school graduation requirements (Honavar, 2005). Many more states are considering making geometry mandatory at the secondary school level, in order to equip our students with the problem-solving and formal reasoning skills necessary for college-bound as well as for non college-bound students. In light of these, the purpose of the current research is to investigate the extent to which formal logical reasoning and proof-constructing skills should be incorporated and tested in the high school geometry course.

The President's current initiatives also focus on professional development projects to help mathematics teachers strengthen their skills (Whitehouse, 2004). In light of this initiative, the research reported here was an effort to replicate and extend the existing research regarding the use of dynamic geometry technology in enhancing reasoning and proving skills in high school students, thereby throwing some more light on the need for including training in the use of dynamic interactive technology as a mandatory program in professional development courses. The personal interviews with geometry teachers conducted by the researcher subsequent to the preliminary survey showed that few high school geometry teachers in the local school district in which the survey was conducted used dynamic geometry technology in their classrooms. Most of those using such technology used Geometer's Sketchpad (GSP) in their classes and then only for demonstration purposes. Very few teachers seemed to have access to a computer lab with GSP installed, where they could take their students occasionally for individual or paired work during geometry instruction. The preliminary survey is discussed in detail in Chapter 3. The findings of the research could inform decision makers about the need to apply for some of the federal grants from the Mathematics and Science Partnership Program (Whitehouse, 2004) to equip each of our local high schools with computer labs where students can engage in meaningful exploratory activities with dynamic geometry software that would enhance their problem solving and reasoning abilities.

Because many students experience difficulty in developing proof construction ability in the high school geometry course, the investigations of this study were expected to provide valuable information that could help improve instructional practices in the teaching-learning process of geometry. To summarize the significance of the current research venture, the findings of this study were expected to: (a) Enlighten educators regarding the extent to which students acquire the ability to construct proofs after a year-long geometry coursework; (b) Provide insight into the relationship between students' logical thinking and proof construction ability, thereby informing teachers about the need to focus on development of logical thinking at all levels of school mathematics; (c) Help geometry teachers address the problem of adjusting instructional process to incorporate appropriate use of dynamic geometry technology; (d) Inform education institutions and professional development organizations about the need to continue to revitalize mathematics education courses so that student proving becomes a fundamental goal of high school geometry course work; (e)

Inform education institutions and professional development organizations about the need to develop additional training programs and workshops to train geometry teachers on the successful use of dynamic geometry technology in their geometry classrooms; (f) Help lawmakers other concerned authorities understand the need for emphasizing logical reasoning skills in standardized tests; and (g) Inform curriculum developers about the need to emphasize inclusion of activities that develop logical thinking and activities that utilize the capabilities of the dynamic geometry technology.

Assumptions

For the purpose of this study, a major assumption made by the researcher was that high school students normally take geometry course at the $9^{th}/10^{th}$ grade level, and they would have attained a fundamental proficiency with basic geometric facts and properties. The study assumed uniformity in age and ability level of participating students at the beginning of the academic year so that their performance at the beginning and at the end could be compared. It was also assumed that all students were familiar with the concept of parallel lines, lengths of segments, congruency of segments, and elementary properties of triangles and quadrilaterals. It was also assumed that by the end of geometry course work, students should be familiar with different styles of writing proofs, inductive and deductive reasoning, and theorems based on properties of triangles and quadrilaterals specified by the *Principles and Standards of School Mathematics* (NCTM, 2000). A third assumption made was that the current study considered GSP to represent any typical dynamic geometry software. Whereas many interactive dynamic technology products like Cabri, TI92, Euklid, and Pythagoras were available for classroom use, almost all the participant teachers of Group B used GSP in their classroom instructional process.

Limitations

One of the main limitations of this study was the restricted and non-randomized sample for the study. The sample for this study was comprised of high school geometry students in the geometry classes of teachers selected from the respondents to a preliminary survey conducted by the researcher. Further detail is provided in Chapter 3. Although 10 public high schools were contacted initially, only 4 schools voluntarily participated in the study. A threat to the external validity due to selection of sample groups could not be completely eliminated. A more generalized study would include private and other schools located in rural and urban regions.

The test-taking abilities and difference in maturity levels of the students in the participating schools might also have affected the results of this study. The difference between the modes of administering the tests might be another confounding factor along with the teaching styles of the two groups in their geometry course work. Other threats to internal and external validity of the study might be due to instrumentation, attrition, and interactions among testing, selection, and other criteria.

Administering pre-test and post-test instruments to all groups of students at the same time reduced to some extent one threat to the internal validity of the tests, arising due to the history and maturation of the students (Campbell & Stanley, 1966). As the groups were selected from different schools, different teachers administered the tests to these groups within a span of two weeks. Providing guidelines for uniform modes of administration of the tests served to reduce this limitation to some extent. However, reduction of other threats for uniformity in test administrations, like differences in time of day, day of week, and topics taught until the time of administration, etc., could not be accomplished due to the difficulty the researcher encountered in getting responses from the participating teachers and scheduling frequent meetings with them. The pre- and post- tests were administered to intact classrooms rather than randomly selected students as this method was more convenient and reduced this limitation to a great extent. Critical reviewing and pilot testing the instruments developed by the researcher reduced the threat to the reliability and validity of the instrument to a great extent. Administering the same instruments to all students reduced, to some extent, the threat arising from instrumentation.

Only 25% of the surveyed high school geometry teachers were found to be using dynamic software in their geometry classrooms for instructional purposes. As a result, the GSP and non-GSP groups had unequal sizes. This posed a threat to the generalizability of the study. However, as the total number of students in the sample was quite large, a natural balance of the different aspects such as age, grade level, socio-economic background, and race of the students was assumed, thus controlling the possibility of this threat.

Taking into consideration only those students' scores that were consistently present throughout the academic year, and were present for both pre-test and post-test, reduced the threat due to mortality. The effect of pre-test on the groups' interactions during the geometry coursework and on the post-test could not be completely controlled. Nevertheless, because the tests were administered in the classroom setting, this effect could be controlled to some extent. The results of the pre- and post-tests were specific to the grade level of the students and to the curriculum and programs specific to the local school district. All these factors would restrict the generalization of the findings. Replication of the study with students of different grade levels, and in different school districts and at different times could check for the threat for external validity due to the history and maturation.

Delimitations

Because most of the respondents of the survey who confirmed using dynamic geometry software in their instruction have mentioned that they use GSP, the study focused on the impact of the use of GSP in instructional processes in developing logical thinking and proof constructing abilities of high school geometry students. A replication of this study for any computer interactive software would shed more light about the effective use of interactive technology in high school geometry classrooms.

The current study aimed to explore the relation between students' logical thinking and their proof construction ability in the context of high school geometry courses. However, the Standards (NCTM, 2000) document recommends incorporation of "Reasoning and Proving" in the mathematical curriculum at all levels. The findings of this study can be reinforced or supplemented by future experiments that investigate ways to continue to revitalize school mathematics instruction using technology, in order to incorporate suitable pedagogical considerations so that school mathematics becomes a place where students develop capacities like critical thinking, justifying, and proving.

CHAPTER 2: REVIEW OF THE LITERATURE

Overview

In this section, an overview of literature based on the following questions related to the current research is discussed:

1. What is the history of proofs in high school geometry curriculum?
2. What are some of the functions of proof in school mathematics?
3. Why should students be made to learn proofs in mathematics?
4. What are the different ways in which proof can be constructed and presented?
5. How do we determine students' ability to construct proofs?
6. What does research say about the use of dynamic geometry technology in geometry classrooms?
7. How do we determine the logical thinking ability of students?
8. What does the research say about the relation between students' ability to construct proofs and their logical thinking ability?
9. What is meant by some of the commonly used terms including proof, logical reasoning?

The literature review is supplemented by a brief narration of the researcher's lens on the abovementioned issues in order to form the conceptual framework for the study.

History of Proofs in High School Geometry Curricula

Geometry became a high school subject in the 1840s when it was made a pre-requisite for university admissions (DeVault & Weaver, 1970). Proof has been a part of the geometry course since then (Stover, 1989). Particularly, the two-column proof has endured many controversies for about a century and has a place in all high school geometry students' academic work even today (Herbst, 2002a). Commenting on the evolution of proof during the first decade of twentieth century, Herbst noted, "…the two-column proof format helped stabilize the geometry curriculum by melding together the proofs given by the text and the proofs expected from students" (p.285). Herbst characterized three periods in the study of proofs in geometry since its inception. He referred to the first phase, when students were asked to replicate the proofs given in texts, as the "Era of Text." During this period, proofs were written in paragraph form made up of long and complex sentences. Explicit reasoning was not done except when it was felt important. Students learned the art of reasoning by reading and reproducing these proofs.

Herbst (2002) referred to the second period, when students were taught to craft proofs for propositions in addition to replicating textbook proofs, as the "Era of Originals." Textbooks included exercises at the end of a chapter, which were some less important theorems and corollaries based on the theorems proved in the chapter. It was hoped that students would develop the capability of reasoning and proving by learning to prove these "originals." Stover (1989) cited National Education Association's (1893) definition of the term, originals, as proofs that are not included in the standard textbooks for geometry. During the ensuing period, the focus of proofs became developing the reasoning skills rather than learning geometry. Diagrams suggesting auxiliary constructions and hints for helping students craft the proofs became the norm in textbooks (Herbst). Moreover, teachers were provided with norms to monitor and guide students' work. Subsequently, the custom of explicitly indicating the logical reasoning for each statement, in order to identify the process of the argument properly became the norm of high school geometry proving activities. Recommendations for introducing informal geometry at the elementary grades were made with the hope that high school students would be equipped with the needed geometrical knowledge to craft proofs and so could focus more the development of formal logical reasoning skills (Herbst).

Herbst (2002) characterized the final stage, when students learn how to do proofs, as the "Era of Exercise." Textbooks that came after the recommendations of the Committee of Ten included specific instructions on methods and strategies to do proofs. According to Herbst, Schultze and Sevenoak first introduced the two-column format of writing proofs in 1913. This format emphasized the need to give reason justifying each statement, thereby helping the students form the logical chain, and saved time for the teacher to check students' work. This method of crafting proofs with explicit justifications for each statement came to be known as "formal proof."

The focus and approach towards proof in the history of secondary mathematics curriculum has been quite uneven with opposing trends for observance of the deductive proof rules (A. Dennis, 2003). These opposing forces greatly influenced the school mathematics curriculum during the first two decades of the 20th century. Curriculum changes that focused more on geometrical knowledge were recommended by the 1923 report of the National Committee on Mathematical Requirements of the International Commission on the Teaching of Mathematics, entitled "Reorganization of Mathematics in Secondary Education" (Hanna, 1983). Hanna noted that one of the major aims for teaching of mathematics recommended by the 1923 Report was to develop the powers of understanding. The revised objectives set forth by the report included ability to think clearly; ability to analyze complex situations and make generalizations; and the ability to appreciate the beauty of geometrical forms, the power of mathematics, its ideals of perfection, and its role on abstract thinking (Hanna, 1983). Nevertheless, during the first half of the 20th century, the content and approach of the secondary mathematics curriculum were revised in order to adapt it for the less mature students (Stover, 1989). Stover noted that classroom practice in the 1930s focused on the body of theorems to be learned rather than on the methods used to prove these theorems, which resulted in a reduced emphasis on the development of critical thinking in students enrolled in the geometry course.

Mid-20th century saw another trend towards increased emphasis on deductive method as a way of thinking. This was a result of widespread agreement that the development of deductive reasoning was a major goal of geometry courses at the high school level (Ireland, 1973). The New Math movement of the late 1950s and early 1960s advocated great emphasis on deductive proofs in high school geometry courses. This triggered a controversy among mathematics educators about the need and importance of teaching deductive proofs in high school mathematics classes. In 1962, 65 mathematics educators published a statement in the *American Mathematical Monthly* condemning the overemphasis of the formal approach to mathematics, mostly on pedagogical grounds, and questioning the suitability of emphasizing formality at the secondary level (Hanna, 1983). Projects such as School Mathematics Study Group (SMSG), Greater Cleveland Mathematics Program (GCMP), Minnesota School

Mathematics and Science Teaching Project (MINNEMAST), Newton Hawley's Geometry Project, which emphasized reading and construction in geometry, had varying direct influence upon the modernization of school mathematics programs during this period.

By the 1970s, reform in mathematics curricula had made good progress. Stover (1989) cited the comments of Ulrich who described geometry as having a nature of being so global and so important that it was capable of being interpreted in many different ways. Several researchers and mathematicians began voicing opposing views regarding the need for including geometry in the high school mathematics curriculum. Ireland (1973) confirmed this trend when he cited the report of Cambridge Conference on School Mathematics (1963) which suggested different routes other than the traditional Euclidean geometry to follow in teaching geometry, such as linear algebra, topology, coordinate geometry, and calculus. The result was disparities in points of view relating to both content and method. During the last quarter of the 20th century, there was considerable unrest in mathematics education (Dessart, 1981). The modern mathematics reform of the 1960s was followed by the back-to-basics movement in the 1970s, leaving the mathematics curriculum with many characteristics of two decades earlier (Dessart, 1981). As the 1980s began, new recommendations for school mathematics were being proposed.

During the later half of the 1970s, the SMSG approach dominated the secondary school geometry curriculum in the United States. The major contribution of the SMSG approach was to introduce postulates concerning abstract rulers and protractors into high school geometry (Dessart, 1981). These postulates, which capitalized on properties of the real number system, filled the logical gaps of the Euclidean geometry previously taught in high school (Dessart, 1981). Apart from SMSG approach, NCTM's 36th yearbook, "Geometry in the Mathematics Curriculum" (1973), identified seven approaches to formal geometry in the senior high school. Included were the conventional synthetic approach, an approach using coordinates, a transformational approach, and a vector approach. The relatively poor performance of the United States students on international assessments gave rise to a fresh bout of concerns and worries about the mathematics education standard in schools and resulted in curricular reform initiatives led by NCTM and publications recommending high expectations for all students and uniform content and process standards for school mathematics (Senk & Thompson, 2003). Between 1992 and 1998, NSF sponsored annual conferences of curriculum developers to discuss common issues and concerns in designing curricula based on the standards advocated by NCTM (Senk & Thompson).

Pursuant to the efforts of various educators, NCTM published *Principles and Standards for School Mathematics* in 2000. The goals for teaching geometry identified by NCTM in this publication include a strong focus on the development of reasoning and proof, using definitions and established facts, and using technology to improve the teaching and learning of geometry. The technology principle reiterates:

> Tools such as dynamic geometry software enable students to model, and have an interactive experience with, a large variety of 2-dimensional shapes. Using technology, students can generate many examples as a way of forming and exploring conjectures, but it is important for them to recognize that generating many examples of a particular phenomenon does not constitute a proof. Visualization and spatial reasoning are also improved by interaction with computer animations and in other technological settings. (p.41).

The current geometry curriculum emphasizes a transformational approach from early grades. The amount of demonstrative geometry is kept to a minimum and prescribed only for high school curricula. A review of different geometry textbooks by the researcher indicated that very few theorems and exercises are provided in many text materials.

Functions of Proof

For many high school mathematics teachers and students, mathematical proof often means the rigorous form of two-column proofs. Many mathematics educators have started reexamining this notion (Knuth & Elliot, 1998). Consequently, the emphasis has shifted away from a perception of mathematical proof as a rigorous two-column proof toward a conception of proof as convincing argument (Hanna, 1990). In this section the various roles and functions of proof perceived by educators and research findings about students' perceptions of about the nature and role of proof are discussed.

Educators' Perceptions of Roles and Functions of Proof

Many educators and mathematicians have identified various roles of proof in classroom and in mathematical community. Some of these roles are: (a) Verification (A. Dennis, 2000; Hanna, 1983; Knuth, 2002); (b) Explanation (A. Dennis, 2000; Hersh, 1993; Knuth, 2002, Lindquist & Clements, 2001); Mingus & Grassl, 1999; NCTM, 2000); (c) Communication (A. Dennis, 2000; Knuth, 2002, Mingus & Grassl, 1999; NCTM, 2000); (d) Discovery (De Villiers, 1999; Knuth, 2002); (e) Systematization (De Villiers, 1999; A.

Dennis, 2000; Hanna, 1983, 1990; Knuth, 2002); (f) Promotion of understanding (Hanna, 1995); and (g) Creation of new knowledge (Herbst, 2002b).

Hersh (1993) described the role of proof in the classroom as follows: “In the classroom, the purpose of proof is to explain. Enlightened use of proofs in the mathematics classroom aims to stimulate the students’ understanding” (p. 389). According to Hersh, mathematical proof can convince as well as explain. In mathematical research, its primary role is convincing. At the high school or undergraduate level, its primary role is explaining. Raman (2003) pointed out, and Hanna (1990) emphasized, that mathematicians routinely distinguish between proofs that demonstrate and proofs that explain.

Dreyfus (1999) identified a close relationship between proving and explaining. Both when proving a theorem and when explaining a state of things, the task is to answer “why?” However, there are a few important differences. According to Dreyfus, explanation may sometimes go beyond proof, while a proof may call for an explanation, which highlights the central idea of the proof. Thus, proof and explanation are interwoven in processes of understanding. Many mathematicians view proof as a form of discourse and a means of communication (Knuth, 2002; Mingus & Grassl, 1999).

It is a common notion that most of the geometric theorems were first discovered through intuition, then verified inductively and proved using deductive reasoning; however, there are many examples in the history of mathematics that suggest that new geometric facts were discovered while proving a theorem or a property. DeVilliers (1990) noted a few instances as a manifestation of this role of proof, when he contends that non-Euclidean geometries could never have been discovered merely by intuition or using quasi-experimental methods. He asserts that new results can be discovered a priori by analyzing the properties of given objects deductively. DeVilliers (1986) also emphasized the role of proof in systematization of mathematical structure. Further elaborating on this role, he stated, “Proof exposes the underlying logical relationships between statements in ways no amount of empirical testing and pure intuition can. Proof is therefore an indispensable tool for systematizing various known results into a deductive system” (DeVilliers, 1986).

Hanna (1995) and Hersh (1993) asserted that one of the main functions of proof in the classroom is promotion of understanding. A proof becomes legitimate and convincing only when it leads to real mathematical understanding (Hanna). According to Hersh, “the choice of whether to present a proof ‘as is,’ to elaborate it, or to abbreviate it, depends on which is more likely to increase the students’ understanding of concepts, methods, applications” (p. 10). For Raman (2003), proof involves both public and private arguments. Raman described private argument as “an argument, which engenders understanding,” and public argument as “an argument with sufficient rigor for a particular mathematical community” (p. 320). Hanna (1990) reiterated that proofs help to substantiate knowledge rather than to establish knowledge, as most proofs come after the knowledge of the things that they prove.

Hanna (1990) also described proof as a social process, stressing the relationship between proof and problem solving. Hanna (1983) also noted that a main role of proof in mathematics is to demonstrate the correctness of a result or the truth of a statement. According to Polya (1957), the role of proof is manifest in the relationship of proof to problem solving and conjecturing. In sum, an informed perception of proof, one that reflects the real meaning of proving in mathematical practice, must include a consideration of proof in each of these roles (Knuth, 2002). The role of proof in the classroom is different from its role in research. In research, its role is to convince. In the classroom, the role of proof is to provide insight to the students into why a theorem is true. According to Hersh (1993), there are two opposing views on the role of proof in teaching--absolutist and humanist. While absolutist view emphasizes complete and correct proof, humanist view accepts the role of proof in informal explanation.

Lastly, it is apparent from a brief review of some of the existing research that the main role of proof in classroom is that of justification and verification, which lead to promotion of understanding. “A good proof must not only be correct and explanatory, it must also take into account, especially in its level of detail, the classroom context and the experience of the students” (Hanna 1995, p. 48). While justification is the public/social aspect, for which the student presents deductive arguments in order to convince other students in the class, verification is the private aspect, whereby the student inductively verifies the truth of a statement and convinces himself.

High School Students’ Perceptions about Proof

Contemporary students rank doing proofs in geometry among the least important, most disliked, and most difficult of school mathematics topics (Ireland, 1973; Senk, 1982, 1989; Stover, 1989; Usiskin, 1982). Data collected by the Cognitive Development and Achievement in Secondary School Geometry (CDASSG) project confirmed that writing proofs was difficult for most students (Senk, 1985; Usiskin, 1982). Most students view a mathematical proof as a method to check and verify a particular case and tend to judge the validity of a proof by its appearance (Martin & Harel, 1989). The findings of Burger and Shaughnessy (1986) showed students’ poor understanding and appreciation of the nature and role of deductive proof in mathematics. Research results on students’ conceptions of proof show that most high school and college students have not been appropriately exposed to the process of proving and justifying the mathematical processes (Dreyfus, 1999). Dreyfus argued that the classroom discourse should provide opportunities for bringing out the tacit mathematical knowledge of our students in justifying and proving, enabling them to articulate their ideas and providing a suitable environment for developing their ability for logical and deductive argumentation.

Segal (1999) noted that most times students' perception of validity of a proof is influenced by the norm set by the teacher, which in turn is guided and reinforced by the examination system prevalent in the school district. The second year data analysis of the longitudinal proof project of McCrone, Martin, Dindyal, and Wallace (2002) revealed "Students' interactions with the teacher and peers lead to the development of taken-as-shared knowledge, or an understanding of social and socio-mathematical norms. These interactions also influence individual students' developing understanding of proof" (p. 3). They also found that students were uneasy with the open nature of local deduction items, and were much more comfortable proving a fact (McCrone, et al.). Although many students find proof construction a difficult process and tend to memorize or copy proofs, most appear to enjoy logical games like chess, crossword puzzles, and other computer games (Steen, 1990). Student-proving activities should engage them in sense making exploratory experiments, so that they come to view proof also as a logical game and learn to appreciate the nature and functions of proof in the curriculum.

In the researcher's view, most high school geometry students find proofs difficult. Some of the reasons for this may be lack of sufficient and appropriate experience and exposure to the proving and problem solving skills in the earlier grades, and lack of motivation resulting in students' inability to perceive the need for learning proofs. Consequently, they do not learn to appreciate the need and beauty of deductive reasoning and the inadequacy of inductive reasoning in generalization of statements and proving the truth of propositions and properties.

Importance of Proof in High School Geometry Coursework

Proof is a fundamental tool in mathematics used by mathematicians to verify and establish the validity of mathematical statements within a specified axiomatic system (Hanna, 1996; Malloy & Friel, 1999). Knuth and Elliot (1998) emphasized the importance of integrating the proving process in school mathematics as it is an essential component of doing, communicating, and recording mathematics that can be incorporated at all levels of school mathematics.

Recent reform initiatives have elevated the status of proof in school mathematics significantly (NCTM, 2000). The Standards document recommends that reasoning and proof should play a much more prominent role throughout the entire school mathematics curriculum and to be a part of the mathematics education to all students (Knuth & Elliot, 1998). A. Dennis (2000) also maintains that proof is an essential activity in doing and understanding mathematics. Hanna (1996) cited Greeno (1994), Schoenfeld (1994), Lakatos, 1976, David & Hersh (1981) when she discussed the controversial arguments among educators about the central role of proof in mathematics education. She concluded that proof deserves a prominent place in the school mathematics curriculum, as it is an ideal method of verification, justification, and a very valuable tool to promote understanding. Investigating, making and testing conjectures, and presenting logical arguments to justify conclusions are essential aspects of the study of geometry (Lindquist & Clements 2001). Hirsh (1985) emphasized development of logical processes, concepts, and language as some of the benefits of proofs in geometry. For Pandiscio (2002), “Proof offers powerful ways of developing and expressing insights” (p. 56).

Callahan and Stanley (n.d.), in their content statement for secondary geometry, chose proof as one of the ten pillars on which the foundation of school geometry is erected. They contended that proof needs to be a visible part of all students’ mathematical school experiences and the concept of proof should run through mathematics at all grade levels. For Waring (2001), proof plays a central role in highlighting interrelationships between different areas of mathematics curricula. He suggested that educators should nurture students’ ability to ask questions by exposing them to situations, which foster a questioning attitude. For Hoyles and Healy (2004), one of the basic goals of mathematics education is to engage students in the process of mathematical argumentation and proving. The view of the researcher also supports the importance of including proving and justifying activities at all grade levels in school mathematics. The power of a proof lies in its ability to explain, convince, communicate, and promote understanding.

Different Ways of Presenting a Geometric Proof

In this section, a brief review of literature about (a) the ways in which a proof can be presented, (b) the different types of arguments that can be used to construct a proof, and (c) the various levels of proof schemes or strategies students use as they progress to the formal deductive level of reasoning, are discussed.

Different Ways of Presenting a Proof

In the first two decades of the 20th century, the two-column format became the national standard. Its acceptance flowed from its use as both a tool to help memorize proofs, and as a scaffolding tool to move students towards creating their own original proofs (Callaghan & Stanley, 2004; Herbst, 2002a). In addition, high school students now learn paragraph format and flow-chart proofs. Hanna (1996) discussed new types of proofs like zero-knowledge proofs and holographic proofs, made possible by computers. Emphasizing the central role of two-column proof in the geometry curriculum Prince (1998) stated:

> The two-column proof is a somewhat rigid form, yet it demonstrates to the students that they may not just give statements or draw conclusions without sound mathematical reasons….Finding hundreds or thousands of examples that support a conclusion does not eliminate the possibility of one hidden, untried example that will contradict that conclusion. Giving students the idea that something can be proved by example leaves the door open for sloppy, half-hearted attempts and oversimplifications of the problem (p. 726).

Paragraph style of writing proof enables the student to describe his/her logical chain of reasoning in a more informal style. A flow-chart of the reasoning process helps the student understand the appropriate use of axioms, propositions, and definitions in the proving process. An important type of proof frequently used in school mathematics is the indirect method of proof. In this type, the student starts by hypothesizing the opposite of the conclusion as true, and proceeds to prove the absurdness of the hypothesis using a logical chain of reasoning based on previously known properties, axioms, propositions, and definitions, thereby confirming the truth of the conclusion (Hanna, 1996). NCTM (1989; 2000) encourages less emphasis on two-column proofs, and recommends that deductive arguments be expressed orally and in sentence or paragraph form.

Different Types of Arguments Used to Construct a Proof

Stover (1989) discussed three types of arguments that are appropriate for secondary mathematics students in the construction of proofs. She noted that these three types were termed as experimental, intuitional, and scientific by Bradford (1908) and as inductive arguments, local deductions, and global deductions by Senk (1982). She also listed five methods of proofs found in secondary geometry texts: (a) direct proof; (b) proof by use of contrapositive; (c) reductio ad absurdum; (d) proof by enumeration; and (e) proof by existence. She stated that some textbooks use informal proofs at the beginning of the course, with some direct proofs towards the end of the course. Informal proofs involve inductive reasoning whereas direct proofs involve deductive reasoning (Stover). According to Stover, while inductive arguments are appropriate for junior high mathematics, high school students should be able to do global deduction. Izen (1998) notes that inductive and deductive arguments can be used together as complementary processes.

Various Levels of Proof Schemes or Strategies Used by Students

Educators use several proof schemes to characterize students' mathematical behavior with regard to proofs (Knuth, 1999). A. Dennis (1977) used the three proof schemes--analysis, synthesis, and a combination of analysis and synthesis--in her study investigating their relative effectiveness. She briefly described synthetic process as the process that begins with the hypothesis and yields deductions until the conclusion is deduced, analysis as the process that begins with the conclusion and argues that it must be the conclusion of some proposition. She also refers to the use of combination of the analytic and synthetic strategies in proofs, until the two chains are connected.

Sowder and Harel (1998) organized students' proof schemes into three categories - externally based proof schemes, empirical proof schemes, and analytic proof schemes - with subcategories for each. They defined externally based proof schemes as schemes comprising arguments based on some outside sources like teacher, textbook, etc., empirical schemes as justifications made solely based on examples, and analytical proof schemes as the ultimate types of justifications. Knuth (1999) modified the proof schemes based on Balacheff's (1988) schemes. He categorized empirical proof schemes as naïve empiricism schemes, crucial experiment schemes and generic example schemes, and introduces verification schemes and illumination schemes as new subcategories for analytic proof schemes (Knuth). Later, based on his findings, he revised the subcategories of the three proof schemes. The revised proof scheme framework is: (a) External conviction proof scheme--authoritarian proof scheme, ritualistic proof scheme, symbolic proof scheme, and experiential proof scheme; (b) Empirical proof schemes--naïve empiricism proof scheme, crucial experiment proof scheme, generic example proof scheme, and visual-empirical proof scheme; (c) Analytic proof schemes--verification proof scheme, illumination proof scheme, and visual-analytic proof scheme (Knuth).

Apart from deductive and inductive reasoning, Callahan & Stanley (n.d.) described a third type called abductive reasoning. This reasoning uses insight drawn from a deep analysis of a single example to identify key features of situations leading to insights on how to form a proof (Callahan & Stanley), and is based on insights drawn from a penetrating analysis of a single example. In the researcher's view, proofs are like language, having rich range of styles and forms.

High School Students' Abilities to Construct Proofs

For the purpose of clarity of ideas, the review of literature for this section was guided by the following sub questions:

1. How do we determine high school students' abilities to construct geometric proofs?
2. What are some of the major problems encountered by students in the study of proofs?
3. What are some of the prerequisites for achieving success in writing proofs?

Determining Students' Abilities to Construct Proofs

Many educators have tried to gain more insight into the cognitive processes that enable students to understand, appreciate, and construct proofs. Piaget (1958) and van Hiele (1957) contributed significantly in this area (Stover, 1989). While Piaget's description did not use any reference to the curricula, van Hiele based categorizations on the geometry curricula (Battista & Clements, 1996). According to Piaget, children pass through three stages in the development of proof-construction ability: (a) Non-reflective, unsystematic, and illogical; (b) Use of empirical results to make and justify predictions; and (c) Formal logical deductions based on assumptions. At the third level, a student is capable of formal logical deduction based on the mathematical system (Battista & Clements).

The van Hieles developed a model to categorize the cognitive levels of students' geometric thinking as: (a) Level 1--visual, (b) Level 2--descriptive/analytic, (c) Level 3--abstract/relational, (d) Level 4--formal deduction, and (e) Level 5--rigor/mathematical (Battista & Clements, 1996). DeVilliers (1986) suggested that deductive reasoning in geometry first occurs at level 3. Studies by Senk (1989) and Battista and Clements (1996) support this notion. Hoffer (1981) classified the van Hiele levels from 1 – 5 as recognition, analysis, ordering, deduction, and rigor. Coxford (1991) modified Hoffer (1981)'s classification as visualization, analysis, abstraction, deduction, and rigor. Malloy and Friel (1999) adapted the definition of van Hiele levels as concrete, analysis, informal deduction, deduction, and rigor. Shaughnessy & Burger (1985) classified the levels as visualization, analysis; informal deduction, formal deduction, and rigor. Their study indicated that very few secondary school students reason at van Hiele level 3 (Shaughnessy & Burger). A. Dennis (2001) used the following descriptions of levels of proving defined by Balacheff (1988) in her investigation of pupils' proof potential:

> Proof by naive empiricism: the verification of a statement on the basis of the weight of evidence from a number of cases
>
> Proof by crucial experiment: the verification of a statement by showing its validity in a typical case (in effect a vain search for counter-examples).
>
> Proof by a generic example: the verification of a statement by appealing to the structural properties of mathematics with reference to a generic example.

Proof by thought experiment: the verification of a statement by appealing to the structural properties of mathematics independent of examples, of the person, and of time. (p. 3).

Chazan (1993) describes the first two proof schemes as realistic or practical and the third and fourth as theoretical proofs. Knuth and Elliot (1998) cited Simon & Blume (1966, p. 8) in describing the four levels:

> At the first level the student concludes that an assertion is valid from a small number of cases. At the second level, the student deals more explicitly with the question of generalization by examining a case that is not very particular [e.g., choosing an extreme case]. At the third level, the student develops argument based on a 'generic example' [e.g., an example representative of a class of objects]. At the fourth level, students begin to detach their explanations from particular examples and begin to move from practical to intellectual proofs. (p. 714)

According to McBride and Carifio (1995), some of the possible categories of student performance on proof writing might include: "unschooled, novices, intermediates, competents, and experts" (McBride & Carifio, p. 7). Waring (2001) illustrated the following framework for students' ability to construct proofs:

> Proof level 0: Pupils are ignorant of the necessity for, or even existence of, proof.

Proof level 1: Pupils are aware of the notion of proof but consider that checking a few special cases is sufficient as proof.

Proof level 2: Pupils are aware that checking a few cases is not tantamount to proof but are satisfied that either (a) checking for more varied and/or randomly selected examples is proof; (or) (b) using a generic example forms a proof for a class.

Proof level 3: Pupils are aware of the need for a generalized proof and, although unable to construct a valid proof unaided, are likely to be able to understand the creation of a proof at an appropriate level of difficulty.

Proof level 4: Pupils are aware of the need for a generalized proof and are also able to construct such proofs in a limited number, probably familiar contexts.

Proof level 5: Pupils are aware of the need for a generalized proof and are able to construct such proofs in a variety of contexts. (P. 5)

In spite of the efforts of many researchers and educators to provide a framework for assessing students' ability to construct proofs and the fact the reasoning and proving are emphasized at all levels of school mathematics, no standardized tests are available for assessing proof construction achievement of high school students. Stover (1989) cited many resources including National Assessment of Educational Progress (NAEP) and National Longitudinal Study of Mathematical Abilities (NSLMA) to conclude that proof writing in geometry is not tested in any nationwide mathematics achievement tests. Nevertheless, researchers like Ireland (1973), Senk (1983), and McCrone & Martin (n.d.) have provided very reliable and valid instruments for assessment of high school students' proof writing ability. Stover (1989) noted that large-scale assessments of students' knowledge of proof have been conducted in Europe. Citing the reports of such assessments conducted by

Krutetskii (1976) and Reynolds (1967), she observed that "Krutetskii's study revealed 'differences between proofs written by capable, average, and incapable students, evidenced in the thought processes involving generalizations, flexibility, and reversibility,' while Reynolds reported 'gradual improving in proof-writing ability from 1st through 6th form' [9th grade through 2nd year junior college] contrary to Inhelder and Piaget's theory." (p. 33-34).

The existing research about valid and reliable methods to assess students' proof construction ability point basically to three stages: Stage 1: A student considers a statement valid or true after testing it on a few arbitrary examples. This type of student has no clue of what generalization means and no idea of deductive reasoning; Stage 2: A student tests a greater number of cases, includes some extreme or typical cases before convincing himself/herself about the truth and generality of a proposition or conjecture; and Stage 3: A student realizes the inadequacy of testing examples and resorts to abstract deductive reasoning, based on definitions, axioms, and propositions in order to prove a conjecture. (Chazan, 1989).

Problems Faced by Students in the Study of Proofs

One of the major problems causing students' difficulty in proof construction is the problem they have with perceiving a need for proof (De Villiers, 1999). According to De Villiers, students often resist having to prove obvious facts such as the sum of the angles of a triangle equaling two right angles. De Villiers noted that this problem might be due to their failure to see the meaning, purpose and usefulness of proofs rather than their inability or low level of ability to reason logically. He also asserted that lack of motivation may be one of the major reasons for this problem. His reinforced his observation by citing several studies, "which have shown that very young children are quite capable of logical reasoning in situations that are real and meaningful to them." (p. 17).

Raman (2003) suggested two reasons for students' difficulty in generating a proof: "lack of knowledge (they don't have the key idea of the proof)", and "an insufficient epistemology (they do not see the essential connection between their privately held idea and what they expect to produce as a formal, public proof)" (p. 3). Stover (1989) cited Greeno (1976) and Kelanic, (1978) to suggest that "the general system of mathematics, instructional practices, curricular materials, and attained levels of cognitive development" (p. 1-2) may be some of the reasons for students' difficulty in constructing proofs. Stover also referred to Driscoll (1987) and Galbraith (1981) and listed some of the obstacles many students face when learning formal proofs as (a) inability to "think hypothetically or express their reasoning in writing," (b) lack of clarity about the "equivalence of a mathematical statement and its contrapositive," (c) unwillingness to " accept the conclusive evidence presented by a counter example," (d) tendency to "generalize too quickly from recognized patterns," (e) "lack of pre-requisite skills and conceptual understanding," (f) tendency to "focus on only part of the statement of a proposition," and (g) tendency to "change the conditions of a proposition to suit their own way of thinking." (p. 21-22)

One of the major reasons for students' difficulty with proof construction may be that they do not perceive proving activities meaningful and connected to their real world. By engaging students in sense-making proving activities and establishing connections of the various properties to the real world situations may help student appreciate the need for deductive reasoning in life, and make proof-writing more meaningful to them.

Prerequisites for Success in Proof Construction

Mingus & Grassl (1989) suggested that an early and broad introduction to proofs that suits students' cognitive development from concrete to formal reasoning may be the best way to foster their understanding of the process of proof and develop their ability to read and write mathematical proofs. For Fawcett (1938), a student is able to understand and construct geometric proofs when the student understands the significance and effects of undefined term, clearly defined terms, assumptions, and unproved propositions. According to Ireland (1973), for success in proof construction, a student should have an understanding of the deductive process in addition to the understanding of the axiomatic system as suggested by Fawcett (1938). Explaining the importance of understanding the deductive process, Ireland suggested that knowledge of logical inference rules, as part of deductive processes is one of the pre-requisites for success in doing proofs.

Liu and Cummings (2001) described two thinking processes that are essential for geometry learning that would help students progress through van Hiele's hierarchy: (a) concrete-abstract process (CA) and (b) abstract-concrete process (AC). CA refers to the process of inductive thinking, from particular facts to a general conclusion about concepts, ideas, etc., whereas AC is a form of deductive thinking which depends on advanced abstract and logical reasoning abilities (Liu & Cummings).

Dreyfus (1999) suggested three kinds of ideas that are essentially involved in the proof construction achievement: (a) A heuristic idea which gives a sense of understanding that ought to be true; (b) A procedural idea, which is based on logic and formal manipulations leading to a formal proof without any connection to informal understandings;

and (c) A key idea, which is an idea that leads to the production of proof. He emphasized that while heuristic ideas give a sense of understanding but not conviction and procedural ideas give a sense of conviction but not understanding, a key idea gives students a sense of both understanding and conviction. According to him, “thinking about how to make key ideas a more central part of both the high school and college curriculum then, seems to be an important step towards helping students develop a mature view of mathematical proof.” (p. 6). Dreyfus also pointed out that students’ linguistic ability may also affect their ability to construct proofs.

Waring (2001) noted five areas of basic skills that should be stressed in high school geometry courses: visual skills, verbal skills, drawing skills, logical skills, and applied skills. According to him, “For students to develop logical skills, many need to work informally with verbal and pictorial ideas before being rushed into rules of logic. They should be aware of ambiguities in language, of the uses of quantifiers, and so forth.” (p. 12-13). Senk (1982) listed some commonly accepted pre-requisites as: (a) knowledge of statement facts of the geometric system; (b) verbal and visual skills; (c) knowledge of inference rules; (d) strategies for constructing proofs, and (e) proof-writing formats. In order to understand, appreciate, and use the concepts and properties in geometry, students should actively participate in the development of the concepts so that they learn to work with the properties and rules instead of merely memorizing the facts (Battista 2002). Indeed, classroom practices which emphasize the social nature of proof such as debating, presenting arguments, and distinguishing between correct and incorrect arguments would help them gain more insight into the formal reasoning skills needed to achieve success in proof construction (Knuth, 2002).

McBride and Carifio (1995) observed that doing a proof requires attention, perception, memory-search, heuristics, and logic. Greeno (1978) listed four types of knowledge needed to succeed in proof construction: (a) general knowledge; (b) visual knowledge; (c) knowledge of inferential propositions; and (d) knowledge of strategies. While the first three are declarative, the fourth is procedural (McBride & Carifio). Bloom (1956) proposed a taxonomy of cognitive behaviors that arranges thought processes into a hierarchy: knowledge, comprehension, application, analysis, synthesis, and evaluation (McBride & Carifio). At the lower end are the recall processes and at the upper end are the reflective processes (McBride & Carifio). A student needs to progress through the first three levels in order to be able to acquire the skills of analysis, synthesis, and evaluation which are essential for success in proof construction.

Based on the above review of literature and her own experience, the researcher is inclined to believe that in order to be able to construct proofs, a student should have: geometry content knowledge of the definitions, terms, axioms, etc.; visual knowledge or knowledge of spatial relationships between different geometrical shapes; inferential knowledge, the ability to understand and use propositions, postulates, and theorems; and strategic knowledge, the knowledge of different strategies such as use of auxiliary constructions, counter-examples, quantifiers, conditional statements, etc.

Logical Thinking Ability

John Dewey has defined logical thinking as "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends…" (Fawcett, 1938, p. 6). Initially studied by Piaget and Inhelder in 1958, the development of formal logical reasoning ability has been widely researched in adolescents and adults (Tobin & Capie, 1980a). The findings of these investigations indicated that many adolescents and adults are limited in their ability to use formal modes of reasoning (Tobin & Capie, 1980a). According to Piaget (1958), a child begins to reason logically when he moves to the formal operational stage, at age 11 or 12, when he is able to judge an argument on the basis of its validity as opposed to judging it on the basis of his belief about the truth of the conclusion (Ireland, 1973; Sharpiro & O'Brien, 1970).

As per the van Hiele levels, the ability for formal logical reasoning is achieved when students reach level 5. Very few high school students reach this stage (Stover, 1989). Many research studies indicate that formal thought is required to learn many of the concepts taught in middle school, high school, and college science courses (Tobin & Capie, 1980b). Evidence also suggests that the majority of students in the middle and high schools are unable to utilize formal operations in problem solving. Concurrently, most of the researchers have urged that priority be given to the development of formal reasoning ability of middle and high school students (Tobin & Capie, 1980b).

Tobin and Capie (1980b) used five modes of formal reasoning identified by Lawson in their development of the Test of Logical Thinking (TOLT), a group test to assess logical thinking in high school students. Initial pilot test of the instrument established a reliability of .74. The instrument was modified so that multiple justifications as well as multiple solutions for each problem were provided. Data reported by Tobin and Capie (1980a) suggested that the 10 items were measuring a common underlying dimension evidenced by a high internal consistency and a one-factor solution obtained from factor analysis. The methodology used to develop the 10 TOLT items is applicable to other situations. TOLT provides a convenient means of obtaining valid and reliable measures of formal reasoning ability for researchers and teachers, particularly when data are required for groups of subjects. Tobin and Capie assert that it has many potential applications for research, “as independent variable in studies of teaching and learning, and as a dependent measure in studies that are concerned with variables that influence formal reasoning abilities of learners” (Tobin & Capie, p. 15). Group tests like TOLT have the advantage of ease of administration and objective scoring procedures.

When subjects were sorted by TOLT score at different educational levels, the results suggested very similar response patterns for subjects at each educational level. According to Capie and Newton (1981),

> The study has indicated a large proportion of pupils at each educational level who are not able to quantify the odds in simple problem situations. Reasoning patterns tended to be independent of problem setting…From an instructional perspective, reasoning patterns may be generalizable, allowing skills developed in one context to be applied to problem solving in other contexts. (p. 9-10)

Moreover, they observed significant interactions between instructional strategies and reasoning ability and suggested that this interaction might contribute to the transition from concrete to formal thought.

Stover (1989) referred to many studies in her discussion about the development of formal logical reasoning in children. Among them, some noteworthy studies are those conducted by Inhelder and Piaget (1958), Lovell (1971), Hill (1961), Roberge (1970. 1972), Platt (1968), Deer (1969), Wirzup (1976), and O’Brien (1973). Based on the views of these researchers, Stover argued that generally the study of formal logical reasoning in non-mathematical contexts has no significant influence on the proof construction ability of students. However, Stover also cited the research conducted by Senk and Usiskin (1983), which showed a substantial correlation between students' reasoning ability and their proof construction ability. Both Stover (1989) and Ireland (1973) cited conflicting findings revealed by researchers: while Inhelder and Piaget (1958) found that formal reasoning develops late and in fixed stages, Hill (1961) found that formal reasoning starts early and grows gradually. Stover also referred to the study conducted by Roberge (1970), which confirmed that hypothetical syllogism (transitivity property) is mastered by students by grade 8, and there is a marked improvement in the understanding of invalid reasoning forms by grade 10.

Ireland (1973) discussed the Cornell Critical Thinking Project (1965), headed by Robert Ennis, which was involved in research relating to the development of logical reasoning on several levels. Ennis studied development of class and conditional reasoning in grades 4-12. His research revealed that basic principles of conditional logic were not mastered by age 11 or 12, nor by age 17. This is in conflict with Piaget's model, which has students moving into the formal operational level by age 11 or 12.

Ireland (1973) cited Robert Kane's (1960) study, wherein he found that both high school and college students tend to accept invalid reasoning as valid if it leads to a conclusion with which they agree and to reject valid reasoning as invalid if it leads to a conclusion with which they disagree. This tendency suggested that to some degree, high school and college students had not reached the formal operational level described by Piaget (Ireland). The study of if-then reasoning conducted by Sharpiro and O'Brien (1970) indicated that school-age subjects treat inference patterns involving if-then sentences in a consistent and erroneous way. Hyram (1957) investigated the effectiveness of instruction on the development of logical thinking with 7th graders (Ireland, 1973). Citing Hyram's (1957) study, Ireland (1973) contended that upper-grade children could be taught logical thinking through appropriate instructional processes, if they have adequate reading ability. Ireland summarized his review of literature on development of logical reasoning as:

> (1) Contrary to Piagetian theory, reasoning abilities do not appear to develop in stages, but rather they seem to develop gradually as children get older; (2) Contrary to Piaget's "formal operational" model, children have mastered some logical principles prior to age 11-12 and have not mastered others even as young adults; (3) The ability to judge the validity of an argument is strongly influenced by the content of the argument at all ages; …. (6) Instruction on the use of the various logical principles does improve students' ability to apply those principles. (Ireland, 1973, p. 23)

Existing research presents conflicting ideas about the development of logical thinking in children. Nevertheless, the researcher tends to believe that students' ability to think logically can be nurtured and increased by careful and focused instruction, where age-appropriate proving and problem solving activities are infused in the instructional process, and suitable classroom socio-mathematical norms are set for creating a conducive atmosphere that encourages argumentation among students. At the same time, the researcher also believes that instruction focused on building formal reasoning modes suggested by Tobin and Capie would help students understand and appreciate the deductive reasoning process, thus help them achieve better proof-writing skills.

Relationship between Students' Proving and Logical Thinking Abilities

One can say that the importance of geometry in the mathematics curriculum is due to its logical development which fosters development of logical thinking in students (Ireland, 1973; Usiskin, 1982). Students' ability to construct proofs includes ability to combine hypotheses, inference rules, and other logical procedures (Susanna, 2003). Based on a review of relevant literature, Ireland illustrated three types of logical reasoning that could contribute to students' ability to construct proofs: conditional reasoning, class reasoning, and ordinal reasoning. He defined reasoning associated with arguments based on conditional statements as conditional reasoning; arguments in which the basic units are sentences containing quantifiers such as all, some, each, or synonyms of these as class reasoning; and arguments involving physical relationships such as length, area, weight, and speed, as ordinal reasoning.

Fawcett (1938) referred to the geometry with a focus on reasoning and proof as demonstrative geometry. Inclusion of demonstrative geometry in high school geometry courses has long been justified on the basis of its contribution in acquainting students with the nature of deductive reasoning (Fawcett). Fawcett commented that most high school geometry tests emphasize the importance of geometrical knowledge and provide very little opportunity to assess the development of students' logical thinking. He also stressed that "demonstrative geometry can be so taught that it will develop the power to reason logically more readily than the school subjects," and that "the degree of transfer of this logical training to situations outside geometry is a fair measure of the efficacy of the instruction" (p. 8). Fawcett advocated conscious efforts on the part of teachers "to bring in illustrations to show the place of logical thinking in life, to carry over geometry to life situations by asking questions on non-geometric material and attempting to get the pupils to apply their geometric types of reasoning to these problems" (p. 26).

Geometric proofs can be used effectively to improve logical thinking (Stover, 1989). While science verifies through observation, mathematics verifies through logical reasoning (Knuth, 1999). Knuth's study investigating secondary school mathematics teachers' perceptions about the nature of proof showed that majority of the teachers identified development of logical thinking as a primary role of proof in secondary school mathematics. It is the researcher's opinion that engaging in mathematical activities could help develop creative problem-solving skills. As Prince (1998) asserted, proof construction activities help students form the habits of ascertaining that no loose ends have been left, that all rules and logical implications have been considered thoroughly, and possibilities of contradictions and gaps have been addressed. Prince calls this type of rigorous logical thinking critical thinking. Students need the rigor of logic to compete in our rapidly changing, technological society.

Dynamic Geometry Technology in Geometry Classrooms

In recent years, several organizations have led efforts to help set standards and expectations for technology use by teachers. For instance, the International Society for Technology in Education (ISTE) and the National Council for the Accreditation of Teacher Education (NCATE) developed *National Standards for Technology in Teacher Preparation,* which provides guidance and expectations for teacher preparation programs in the area of technology (Edutopia, 2004). The recent reform initiatives assume that there is potential to take a positive step towards engaging students more actively in the process of mathematical thinking and mathematics learning through the use of technology (Manoucherhri, 1999). Although these important efforts contribute significantly to measuring the amount of technology that is available to teachers and students, they do not focus on measuring how effectively technology is being used in the teaching-learning process (Moore, 2001).

The concept of using technology to teach geometry is grounded on the developmentalist theories of Piaget and van Hiele (Roberts & Stephens, 1999). Winicki-Landman (2002) observed that a simple problem could create a learning environment in the classroom that fosters meaningful engagement of students in the proving process when appropriate learning activities with dynamic geometry technology are designed by teachers. Many educators recommend students' engagement with intuitive and empirical explorations and justifications initially so that they realize the significance and power of deductive argumentation (Battista & Clements, 1996; [Eves, 1972; Lakatos, 1976] Pandiscio, 2002). Pandiscio found that dynamic geometry software like Geometer's Sketchpad™ (GSP) and Cabri™ facilitate such exploratory activities, leading students to make and test conjectures.

In his study exploring secondary preservice teachers' learning of mathematics in a technology environment, Zhonghong (2002) observed that after conjectures were made based on the exploratory tasks and tested, the subjects continued GSP explorations to come up with insight for reasoning and proofs. Many teachers and educators used the dynamic geometry software to help discover properties and relationships, to make and test conjectures, and to construct geometric objects (Zhonghong). Zhonghong's findings support Battista (2002) claimed that GSP explorations can not only encourage students to make conjectures, but also can foster insight for constructing proofs.

Pandiscio (2002) studied secondary pre-service teachers' perceptions of need for and the benefits of formal proof when given geometric tasks in the context of dynamic geometry software. The subjects for this study were four pre-service secondary teachers. The most striking result was that all four participants saw dynamic software as a tool to make sense of proofs but not necessarily as a tool that is helpful to create proofs. All participants believed that using the software helped them understand the ideas embedded in the theorems and problems more fully than they would have understood without the aid of technology. In Pandiscio's words, "Perhaps such understanding, combined with an informal yet structured system of justification, could serve as a goal for students using dynamic geometry software" (p. 222).

Logo was developed by Seymour Papert as a conceptual framework for understanding children's construction and knowledge about mathematics and problem solving (Papert, 1980). Many studies investigating the effectiveness of Logo found that Logo can be effective in motivating children to learn through exploration and discovery (Clements & Meredith, 1993; Liu & Cummings, 1997, 2001; Papert, 1980). In Logo-based environments, students create geometric objects and apply constructions and transformations to those objects through a series of pseudo-programming commands (Glass & Deckert, 2001).

According to Battista & Clements (2002), GSP preserves the basic properties of a geometric construction, which enables the student to explore the generality of the consequences of the construction. Such activities foster insights for constructing proofs. DeVilliers (2002) describes the capability of GSP very lucidly using the words of Hofstadter (1997):

> The beauty of Geometer's Sketchpad is that it allows you to discover instantly whether a conjecture is right or wrong--if it's wrong, it will be immediately obvious when you play around with a construction dynamically on the screen. If it's right, things will stay 'in synch' right on the button no matter how you play with the figure. The degree of certainty and confidence that this gives is downright amazing. It's not a proof, of course, but in some sense, I would argue, this kind of direct contact with the phenomenon is even more convincing than a proof, because you really see it all happening right before your eyes.... (p. 5)

Rasmussen (1992) found working with geometric models constructed with the GSP to be engaging and insightful. Dynamic geometry software is characterized by the ability to use mouse operations to construct geometric objects. These objects "then retain their geometrical properties when manipulated further by the mouse" (Dye 2001, p. 157). Such behavior of diagrams created with dynamic geometry software is controlled by theory. This important property helps students to discover geometric relationships that remain invariant when certain elements of a construction are changed (Galindo, 1998). For example, the trace feature exploits the great advantage of dynamic constructions over static drawing on paper. Loci investigations afford many opportunities to look at students' means of justifications (Galindo). Students can be encouraged to formulate hypotheses and verify them using different strategies through sketchpad activities. Liu and Cummings (2001) posited that GSP is a useful tool for advancing children thinking through van Hiele's hierarchy.

As Allen, Channac, and Trillling (2001) observed, constructing dynamic figures involves "constructing programs for dynamic geometry systems." (p. 179). According to

Santos-Trigo and Espinosa-Perez (2002), routine tasks that appear in traditional geometry courses can be approached differently with the use of technology. They argued that:

> In some cases, students will identify conjectures that are not easy to observe in advance…. Following a surprise, many students may require a proof, may be not explicitly, but by demanding from others or from themselves an answer to their 'why' (or 'why not')….If possible, the proof, namely the answer to the 'why,' should arise from the observations and the revisions of the experimentation process itself (p. 38).

Santos-Trigo and Espinosa-Perez described two main ingredients that distinguish students' interaction with the task: (a) the identification of relationships or conjectures through the use of the software; and (b) the search for mathematical arguments to support those conjectures. DeVilliers (1999) contended that when students were exposed to GSP activities, verification served little or no motivation for doing proof. According to him, a good motivation is in challenging them to explain why the result is true and encouraging them to find deductive argument as an attempt at explanation.

There is a growing misconception among educators that powerful tools like GSP make proofs obsolete (DeVilliers, 2002; Hanna, 1995). DeVilliers illustrated the continued central role proof occupies in school geometry using the following example:

> Find the ratio of the area of quadrilateral IJKL to the area of quadrilateral ABCD – where I, J, K, and L are midpoints of the sides of the quadrilateral ABCD. – This example works well to sensitize students to the fact that although Sketchpad is very accurate and extremely useful for exploring the validity of conjectures, one could still make false conjectures with it if one is not very careful. Generally, even if one is measuring and calculating to 3 decimal accuracy, which the maximum capacity of

> Sketchpad 3, one cannot have absolute certainty that there are no changes to the fourth, fifth or sixth decimals (or the 100th decimal!) that are just not displayed when rounding off to three decimals. This is why a logical explanation/proof, even in such a convincing environment as Sketchpad, is necessary for absolute certainty. (p. 8-9)

According to Rasmussen (1992), Sketchpad is particularly effective in helping students explain and prove geometric theorems because of its (a) rich visual context (van Hiele level 1), (b) dynamic nature, (c) ease of use that encourages exploration, (d) ability to focus attention on geometric behavior of shapes constructed, (e) capacity to provide a diverse set of cases over which a conjecture can be tested, and (f) facility that lets students explain their thinking in their own language and communicate with others.

Mcgehee (1998) noted that dynamic geometry technology helped students connect visual justification and empirical thinking to higher levels of geometric thinking thus helping them develop their ability for logical justification in formal proof. According to Mcgehee, the teacher must plan activities that engage students in the interactive features of the software rather than just present the constructions and theorems on the screen, so that students gain proper insight into the validity of the theorems. He also suggested that instead of encouraging students to use the interactive feature at the end of the activity to verify geometric properties and theorems, they should be allowed to use these features at the beginning, in order to foster discovery. Students who are exposed to interactive geometry may be able to focus their attention on the more important aspects of a diagram rather than on the diagram as a whole (Glass & Deckert, 2001)

Roberts & Stephens (1999) compared students of average ability in three high school geometry classes that utilized computer software Geometry Inventor in varying amounts. The first class used it twice a week; the second class, once a week; and the third class did not use the software at all during a yearlong geometry course. They observed that using the software improved the interest and participation of students. The researcher's experience with GSP has convinced her that this software is useful in engaging students in interesting and meaningful learning activities. The software can also help students gain insight into the necessity for formal deductive reasoning and provide them with a sound base to learn proof-writing skills.

Definitions of Frequently Used Terminology

Proof: The term 'proof' can be broadly defined as a complete explanation and a convincing argument (Mingus & Grassl, 1999**)**. Various definitions can be found for proof, coined by educators, which reveal different purposes and functions of proof. However, the term proof commonly refers to 'either a chain of reasoning or to the problem which requires such a chain (Stover, 1989). Some of the definitions assigned to the term proof are presented in Table 1 along with the associated sources.

As the review of the various meanings reveal, there are different schools of thought that differ on the criteria for a valid proof (Hanna, 1983). All mathematicians and mathematics educators may not agree on a single frame of reference for the meaning of the term 'proof.' Nevertheless, as Hanna pointed out, rigorous proof should be viewed as a valuable mathematical tool rather than as the core of all mathematical thinking. For the purpose of this study, the term 'proof' is considered as an argument using a logical chain of reasoning, using known axioms, definitions, postulates, etc., assuming the truth of the antecedent, and proving the truth or falsity of the consequent of a statement.

Table 1: Definitions of Proof

Author (Date)	Definition
High School Mathematics: Teachers' Edition, 1964 (Cited in Hanna, 1983)	A proof of a statement is a derivation showing that the statement is a consequence of the postulates.
Hanna (1983)	A proof in mathematics or logic satisfies two conditions of explicitness. First every definition, assumption, and rule of inference appealed to in the proof has been or could be, explicitly stated; in other words, the proof is carried out within the frame of reference of a specific known axiomatic system. Second, every step in the chain of deductions which constitutes the proof is set out explicitly
Webster's Ninth New Collegiate Dictionary (1986, p.1982) (Cited in Stover, 1989)	"the cogency of evidence that compels acceptance by the mind of a truth or fact"
Hanna, 1996	a logical argument establishing the truth of a geometric statement in which all the information used and all the rules of reasoning are clearly displayed
Prince, A.A. (1998)	The word proof in my mathematics dictionary is defined as "the logical argument, which establishes the truth of a statement" (James & James 1968, p. 291).
Sowder, L., & Harel, G. (1998)	Proving, or justifying, a result involves ascertaining – that is, convincing oneself – and persuading, that is, convincing others
Principles and Standards for School Mathematics (NCTM 2000, p.56)	"arguments consisting of logically rigorous deductions of conclusions from hypotheses"
The Oxford American Desk Dictionary and Thesaurus (2nd Edition, 2001)	Facts, evidence, argument etc., establishing or helping to establish a fact.

Proof writing: In general, formally presenting a mathematical thought or argument in order to convince others about the truth or falsity of a statement is regarded as constructing or writing a proof. Proof is generally constructed as a step-by-step explanation based on prior knowledge and previously known axioms, postulates, definitions, and theorems, to draw a conclusion about a geometric statement, with supporting reasons.

Proof construction ability: Proof construction ability refers to the ability of a student to draw inferences or conclusions from known or assumed facts, to progress, by some method, from the given information to the desired conclusion. A student may use either synthesis or analysis or a combination of both to construct a proof. For the purpose of this study, students' proof construction ability is determined on the basis of skills purported by Ireland (1973). It is assumed that a student who is able to construct proof is able to:

1. Demonstrate an understanding of conditional statements.
2. Distinguish between true and false statements.
3. Recognize the use of counter-examples to disprove the generalization of a statement.
4. Analyze a given argument using valid inference patterns, and identify invalid inference patterns as invalid.
5. Recognize and use logical implications, quantifiers, etc., appropriately for proving the truth or falsity of a statement.
6. Understand and apply visual and spatial reasoning establish the validity of inferences.
7. Knowledge and understanding of definitions, postulates, etc.

Logical thinking ability: Logical thinking or reasoning ability refers to the ability of a student to use formal reasoning to solve problems. For Stover (1989), reasoning ability needed to construct proofs refers to the ability of an individual to draw inferences or conclusions from known or assumed facts, to progress, by some method, from hypothesis to conclusion. According to Ireland (1973), "Deductive reasoning and logical reasoning are used synonymously to represent the mental process of determining whether or not a statement is implied from the hypotheses" (p. 16). He identifies three of several modes of logic useful in proof construction: sentence logic, class logic, and ordinal logic:

> Sentence logic is concerned with arguments in which the basic units are sentences often connected by such logical connectives as if, only if, if and only if, then, and, or, not, and both. Reasoning associated with arguments based on if, only if, and if and only if are termed as conditional reasoning. Class logic is concerned with arguments in which the basic units are sentences containing quantifiers such as all, some, each, none, etc. Ordinal logic deals with physical relationships such as length, weight, area, speed, etc., and uses terms like longer than, equal to, faster than, etc. (p. 16)

In order to assess students' logical thinking and their cognitive level, Tobin & Capie (1980a; 1980b) use five modes of reasoning in their research: proportional reasoning, controlling variables, probabilistic reasoning, correlation reasoning, and combinatorial reasoning. The current research used the Test of Logical Thinking (TOLT) developed by Tobin & Capie (1980a; 1980b) in order to analyze the logical thinking abilities of high school geometry students based on these five reasoning abilities.

Dynamic geometry: Dynamic geometry refers to a family of computer programs (e.g., Geometer's Sketchpad, Cabri) that allows users to construct, measure, and manipulate shapes on screen. Objects' measurements are displayed on-screen and change dynamically as the user drags and resizes the object. These are tools, similar to word processors and spreadsheets, which can be used to create and support constructivist learning environments. The programs contain no information or instructional content. Tools available in dynamic geometry programs could be used just as easily by teachers with an objectivist philosophy of learning as by teachers with a more constructivist view (Hannafin, Burruss, & Little, 2001).

Dynamic geometry technology: Dynamic geometry technology refers to the technologies of computers and handheld calculators with the dynamic interactive software installed, which constructs geometric shapes and figures based on the basic properties defined for the specific shapes and figures.

Summary

The review of related literature and research was conducted to provide a sound rationale and foundation for the research undertaken in the present study. The following statements related to major sections of the literature review summarize the work of researchers and theorists and provide partial answers to questions raised in the study.

The brief review of literature on the history of proofs in the high school geometry curriculum prompted the researcher to observe that although the focus and emphasis of geometric proofs have undergone major shifts over the past two centuries, the basic intent of teaching proofs in geometry course still remains. The intent has been to develop in students logical thought processes and representations of these thought processes in precise and succinct terms, which could be effectively transferred to other real life contexts, and thus help them to become better problem solvers and face the ever-increasing challenges of the world.

In addition to verifying the truth of statements, proof serves as a means of explanation, communication, and systematization of mathematics into an organized axiomatic structure, tool for discovering and creating new knowledge, and a vehicle for promoting understanding. Geometric proofs could be integrated in the high school geometry classroom instruction with different focuses. One or more of the above mentioned functions of proof could be projected in the classroom by the teacher. Nevertheless, the fundamental objectives of teaching students to prove theorems verbally and to construct these proofs in writing should be to develop deeper insight into the formal thought processes and effective ways to articulate these thought processes in order to convince others.

Students go through various developmental stages before they are ready for formal deductive reasoning, which enables them to construct proofs. At each level, they acquire different components of logical thinking at varying degrees. For the purpose of this research, students' ability to construct proofs is determined by four factors: (a) ability to analyze conditional statements and translate general statements into conditional statements and vice versa (conditional reasoning); (b) ability to present arguments based on physical aspects such as length, area, etc. (ordinal reasoning); (c) ability to determine the truth or falsity of a statement and justify it based on a given situation (analytical reasoning); and (d) ability to effectively use counter examples, quantifiers, etc. to prove the truth or falsity of a statement. Proof is an essential component of doing, communicating, and recording mathematics. Learning the proving process empowers students with deductive reasoning capabilities, which are highly important for a successful career later in their lives.

In high school geometry, proofs can be presented in paragraph, flow-chart, or two-column formats. Different proving strategies include direct methods using analytic, synthetic, or a combination of these strategies, indirect method, and disproving by counter-examples. Dynamic geometry technology can be exploited to engage students in interesting and meaningful exploratory activities, challenge them to make and test conjectures, and obtain a better insight into the process of formal reasoning and proof.

Research studies determining the inter-relationship between students' ability to construct different types of proofs and their logical thinking ability are very few. The existing research studies reveal that students' ability to construct proofs is related to their level of logical reasoning ability. Several researchers have identified the cognitive levels students pass through before they are able to reason formally and deductively, in order to be successful in the abstract proof construction process.

Teachers' conceptions of the nature and functions of proof and their attitude toward the importance of teaching proofs and the use of technology in teaching proofs influences the development of students' logical thinking and positive attitude towards proofs, which are important to achieve success in constructing proofs.

CHAPTER 3: METHODS AND PROCEDURES

Chapter Overview

This chapter provides a description of the research methodology and procedures used to investigate high school geometry students' proof-construction and logical thinking abilities, and the impact of dynamic geometry software in enhancing these abilities. The first section restates the research questions that guided the research. The next four sections describe population selection, preliminary survey, instrumentation, and data collection procedures. The following section provides a brief description of specific data analysis procedures. The final section summarizes the organization and rationale used in reporting and discussing the results in the remaining chapters.

Research Questions

This study was primarily aimed to investigate (a) the role of yearlong geometry coursework in developing logical reasoning ability and proof-construction ability in students, (b) the linkage, if any, between students' logical reasoning and proof construction abilities, and (c) the impact of dynamic geometry software on student performance. Because the abilities under investigation were quite complex, the researcher ventured to determine a measure of the abilities by means of students' performance on two pre-tests and two post-tests. The research questions that guided this study were as follows:

1. Is there a significant difference in the performance of high school geometry students between the proof tests administered at the beginning and at the end of a yearlong geometry coursework based on:
   (a) the period of instruction?
   (b) use of dynamic geometry software?
   (c) type of geometry course?
   (d) an interaction between the use of dynamic geometry software and the type of geometry course?
2. Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework based on:
   (a) the period of instruction?
   (b) use of dynamic geometry software?
   (c) type of geometry course?
   (d) an interaction between the use of dynamic geometry software and the type of geometry course?
3. Is there a correlation between the students' performances on the tests of logical thinking and proof construction, administered at the beginning and at the end of a yearlong geometry coursework?
4. How well do high school geometry students learn to construct geometry proofs by the end of the high school geometry course?

Population and Sample Selection

The target population for this study was high school geometry students of the public high schools system governed by the Local Education Agency (LEA) in a large urban area in the southeastern region of the United States. The LEA used in this study is one of the nation's largest school systems, serving a diverse population of students of different national, religious, and ethnic origins. A total of 15 public high schools are affiliated to the LEA. Based on the random sample of 30 teachers from the preliminary survey, 15 from Group A (referred to as non-GSP group) and 15 from Group B (referred to as GSP group), the mathematics department chairs of ten schools were contacted by the researcher with a request to allow the geometry classes of these teachers to participate in this study. Approvals to conduct the research in the LEA were obtained from the Institutional Review Board (Appendix A) as well as the LEA's Board of Education (Appendix B) before contacting the mathematics department chairs of these schools. The chairs of four public high schools responded positively. For the purpose of anonymity, these schools are represented as Sch1, Sch2, Sch3 and Sch4 in this report. Thus, the sample of the study comprised of all geometry students of these four schools. Geometry classes of the main campuses of all the four schools as well as the ninth grade center of Sch3 participated in the study. This ninth grade center is referred to as Sch3F in the report. Details about the four schools regarding the student population were obtained from the Florida Department of Education website.

School 1 (Sch1) served a student population of about 3,500 and about 200 students were enrolled in the mastery geometry course during the academic year 2004-2005, when the study was conducted. The school had adopted the mastery geometry course designed by an independent company, covering all the Sunshine State Standards for geometry, where students had to master 80% of each geometry competency or better. The geometry teachers of Sch1 had designed their own course structure for all their geometry students with some modifications to the mastery course: Honors students studied 80% mastery coursework, with the remaining 20% added from honors coursework, whereas regular geometry students covered 90% of the mastery coursework, and the remaining 10% was left to the teachers' discretion, usually home work assignments. Of the six geometry classes in Sch1, five were regular and two were honors classes. As there was not much difference in the course structure of the regular and honors students in the mastery course, all students of Sch1 were treated as mastery course students in this study. Of the three teachers, the teacher who taught two regular and two honors classes confirmed using Geometer's Sketchpad (GSP) for classroom demonstrations. Thus, all the students studying under that teacher were categorized as the GSP group, and the remaining students were included in the non-GSP group.

School 2 (Sch2) had a student population of about 3,350, out of which approximately 650 geometry students taught by eight teachers during the academic year 2004-2005 participated in this study. This school offered regular and honors geometry courses designed by the LEA's Board of Education, based on the Sunshine State Standards. Four of these teachers used GSP for demonstration and/or small group work. All students enrolled in five honors and four regular classes taught by these four teachers were considered as GSP group, and other students were included in the non-GSP group. School 3 (Sch3) served a student population of 3,850 during the academic year 2004-2005, of which about 1,050 students were enrolled in the ninth grade center of this school (Sch3F). This school also offered regular and honors geometry courses designed by LEA's Board of Education. Approximately 180 students from the main center and 220 students from the ninth grade center of Sch3 participated in this study. All the five teachers from Sch3 confirmed that they did not use any dynamic geometry software in their instructional processes. Accordingly, all students enrolled in the nine honors and three regular courses of Sch3 were included in the non-GSP group.

School 4 (Sch4) served student populations of approximately 560 during the 2004-2005 academic year. Only one geometry teacher who was also the mathematics department chair of the school of Sch4 participated in this study. Sch4 also offered honors and regular geometry courses designed by the LEA Board of Education. As this teacher confirmed the use of GSP for instructional purposes, approximately 100 students enrolled in two regular and one honors courses taught by this teacher were put in the GSP group. The ethnic diversity of the participating schools is illustrated in Table 2, which shows the percentages of students of White, Black, Hispanic and other ethnic origins enrolled in each of the five schools during 2004-00 (OCPS, 2004).

The researcher made frequent visits to the participating schools during the academic year, and held informal conversations with most of the teachers. It could be concluded from these interactions that (a) all those teachers who used GSP had followed the same curricular structure as the curriculum followed by the non-GSP teachers in the same school, (b) some of the GSP teachers stated using GSP to help students understand the theorems and definitions during classroom demonstrations. These teachers also stated that they took their students to the computer laboratory once in a while to let them explore a particular topic using GSP, aided by worksheets containing instructional sequences for the specific topic, and (c) some other teachers stated that they brought a cart-full of laptop computers to the class and organized similar activities.

Table 2: Ethnic Distribution in Participating Schools (Percentages)

School code	Total Enrollment in 2004 Academic Year	Whites	Blacks	Hispanics	Others
Sch1	3,500	40	10	42	8
Sch2	3,350	50	25	16	9
Sch3	2,500	61	17	15	7
Sch3F	1,000	59	17	18	6
Sch4	560	19	62	13	6

Table 3 displays a distribution of students who were administered pre-tests I and II in both groups from these five schools. The strategy used for teacher sample selection was criterion-based purposeful sampling (Creswell, 1998). Although random selection was applied initially to a restricted extent, due to the difficulty encountered by the researcher in enlisting the cooperation of all the selected teachers, the final selection of teachers resulted in convenience sampling.

Table 3: School-wise Distribution of Students in Courses and Groups

School	Teacher	No. of students	Class	Group
Sch1	T1	35	Mastery	A
	T2	102	Mastery	B
	T3	54	Mastery	A
Sch2	T4	69	Regular	A
	T5	116	Honors	B
	T6	68	Regular	B
	T7	58	Honors	B
	T8	71	Regular	B
	T9	68	Regular	A
	T10	99	Regular	A
	T11	93	66 Regular; 27 Honors	A
Sch3	T12	64	Honors	A
	T13	59	Honors	A
	T14	60	Regular	A
Sch3F	T15	20	Regular	A
	T16	196	Honors	A
Sch4	T17	95	61 Regular 34 Honors	B

Preliminary Survey

A survey was conducted in the spring semester of 2004, to investigate the perceptions of high school geometry teachers about proof in geometry, and to determine if experience and educational attainment of teachers influence their perceptions. A questionnaire containing 22 items was administered to high school geometry teachers from thirteen public high schools in the LEA. A total of 150 forms were distributed, of which 95 were filled out and returned. A copy of the survey questionnaire is provided in Appendix C.

The questionnaire included eight items on a Likert scale of agreement, to investigate the perceptions of teachers regarding the functions of proof and to understand their comfort level and their beliefs about their knowledge of different types of proofs. One item investigated teachers' ideas about the appropriate grade level in which students should be introduced to proofs in mathematics coursework. Four more items explored teachers' use of, and their beliefs about, the usefulness of dynamic geometry software in teaching geometry and proofs. Four items were included to determine their educational attainment and experience in teaching geometry. One open-ended item explored participants' perceived reasons for not using dynamic geometry software in their instructional processes, if they have stated they did not use any. The last four items were open-ended questions, asking the respondents to state the most difficult and most enjoyable aspects of teaching proofs, reasons for their choice to use or not to use dynamic geometry technology in their instructional processes, and about any other comments they would like to make.

Teachers' responses to all items other than the open-ended items were analyzed using descriptive statistics and cross tabulation to investigate the influence of teaching experience and educational attainment on their responses. In addition, teachers' responses to the multiple-choice items were analyzed using correlation tests. The open-ended questions were analyzed qualitatively by tabulation of responses, coding, and categorizing. Table 4 illustrates the correlation coefficients for the multiple-choice items with educational attainment and experience.

Analysis of the cross tabulations and the Spearman correlation tests suggested the following results:

1. The statements about the functions of proof as means of mathematical communication, explanation, and creation of knowledge had statistically significant positive correlations, though low, with experience. Experience also had a significant positive correlation with the comfort level of teachers for teaching proofs.
2. All statements had a low correlation with participants' educational attainment. Education had a significant positive correlation with only one statement, related to the comfort level of teachers for teaching proofs.
3. Very few teachers in the sample were familiar with dynamic geometry software for teaching geometry. Those familiar with such software invariably used GSP. Those who used GSP felt it was useful to teach geometry, but were not quite sure of its usefulness to teach proofs.
4. Educational attainment seemed to have a slightly negative correlation with teachers' use of and attitude towards dynamic geometry technology. Although many experienced teachers in the sample preferred not to use dynamic geometry technology in their classrooms for various reasons, those who used found it more

useful to teach proofs than the less experienced teachers who used dynamic geometry technology.

Table 4: Correlation Coefficients of the Statements Based on Experience and Education

Statements	Coefficients	
	Experience	Education
Formal proof develops logical reasoning	0.20	0.10
I am conversant with different styles of proofs	0.19	0.20
Proof plays an important role in mathematical understanding	0.11	0.20
Proof establishes truth of a statement	0.05	0.08
Proof is a means of mathematical communication	0.36**	0.20
One aspect of proof is explanation	0.25**	0.05
Proof creates knowledge and systematizes it	0.31**	0.17
I enjoy teaching proof	0.24**	0.24*
At what grade level students should be introduced to proofs?	0.10	0.01
How do you rate the usefulness of dynamic geometry software in teaching geometry?	0.04	-0.07
How do you rate the usefulness of dynamic geometry software in teaching proofs?	0.06	-0.07

* $p < .05$

** $p < .01$

5. Teachers with a masters or doctoral degree felt more confident about their understanding of different styles of proofs than undergraduates. This also confirmed

the findings of the research done by Knuth (2002), which showed that teachers exposed to more mathematics courses might tend to have better conception of proof.

6. The perceived understanding of different styles of proofs was more in experienced teachers than novices in geometry teaching.

Instrumentation

The tools of investigation for this study were broadly categorized into two types: a set of pre- and post- test instruments to assess students' logical thinking ability, and a set of pre- and post- instruments to test their proof construction ability. Two sets of the Test of Logical Thinking (TOLT) developed by Tobin & Capie (1980a; 1980b) were used as pre- and post-tests part I, with permission, to assess students' logical thinking ability. TOLT instruments consisted of ten items, with two items assigned to measure each of the five modes of formal reasoning: (a) controlling variables; (b) proportional reasoning; (c) probabilistic reasoning; (d) correlational reasoning; and (e) combinatorial reasoning. Investigations of TOLT validity and reliability on a sample of about 600 students established a high reliability (.85) and internal consistency (.56 -.82) (Capie & Tobin). Clinical interviews were also conducted for validating the items, and the correlation between performance on the interviews and TOLT was .76 (Capie & Tobin). Investigations of predictive validity of TOLT on different samples during 1979 through 1980 conducted by different researches produced a correlation in the range of .46 to .73 (Capie & Tobin, 1981).

As students could not be expected to actually construct proofs at the beginning of the geometry course, pre-test part II was designed to measure the pre-requisite abilities for success in proof construction: (a) spatial reasoning; (b) conditional reasoning; (c) ordinal reasoning; (d) ability to use and analyze quantifiers; (e) ability to use counter-examples; (f) ability to analyze arguments to identify valid inference patterns and understand the inadequacy of examples to prove the truth of a statement; and (g) declarative knowledge (knowledge and understanding of definitions, postulates, etc.). This test is also referred to as *proof construction test* in this report, in order to mark its difference from the proof writing test administered at the end of the year. The initial draft version of the instrument contained 8 tasks and a small survey questionnaire with 10 statements. After review and feedback from peers and faculty members, a revised version of the instrument consisted of seven tasks and the survey questionnaire at the end of the test instrument. The revised instrument was then sent to approximately seven mathematics educators at other universities whose research works were among the guiding forces behind the current study.

Z. Usiskin (personal communication, July, 2004) responded with a detailed task-by-task analysis. He opined that the test was a bit too long and too complex for entry-level geometry students. He also suggested focusing on one aspect of proof and reducing some items. E. G. Knuth (personal communication, July, 2004) also suggested removal of some specific tasks, explaining that they will be too complex for entry-level students, and suggested two other tasks, for inclusion. He also suggested inclusion of a task that could be directly related to use of GSP so that the performance difference between GSP and non-GSP groups would be more obvious. S. L. Senk (personal communication, August, 2004) confirmed the overall content validity of the test instrument. She pointed out some inconsistencies in numbering the tasks, sub questions, and statements. Senk also suggested inclusion of some items involving geometry content and vocabulary. After revising the instrument based on the feedback from the educators, a third version of the instrument was pilot tested in a local private high school.

This version included three main tasks, with two sub-tasks each, and the survey questionnaire. All six tasks required the students to have acquired a basic knowledge of classifications and properties of triangles and quadrilaterals (declarative knowledge). The two subtasks of task 1 aimed at testing students' spatial reasoning, ordinal reasoning, and ability to use counter-examples. Task 2 consisted of two sub-tasks that aimed at investigating students' ability to analyze conditional statements and to use quantifiers. The last and third task also had two sub-tasks, which investigated students' ability to analyze arguments. Analyzing arguments required students to recognize the validity and generalizability of statements and understand the inadequacy of proving the truth of a statement based on a few examples. This version was pilot tested at a local private high school (PHS). The mathematics department chair of this school consented to administer this test to two geometry classes, one following the regular course and the other following the honors course. A total of 69 geometry students took the pilot test. The completed test papers were graded by the researcher and two graduate students from the computer science department, based on the scoring scheme developed by the researcher. The development of the scoring scheme is discussed in detail in the following section. The item-wise scores allotted by the two raters were analyzed to establish inter-rater reliability and the average of the scores awarded by the two raters for each item were analyzed to investigate the reliability of the instrument. Item-wise analysis of the pilot test scores showed inter-rater reliability between 0 .97 and 0.99 (Cronbach alpha), and overall reliability of the instrument established a coefficient of .76.

Moreover, factor analysis extracted a single factor, thus confirming that all the tasks were measuring one common underlying dimension.

While administering the pilot test to the geometry classes, teachers were requested to record any questions/clarifications raised by students, and the adequacy of the stipulated 30 minutes for completing the test. Their feedback revealed many students' difficulties in understanding the survey questionnaire. Based on the teachers' observations and suggestions, the survey questionnaire was removed from the pre-test instrument. The time allotted was found to be sufficient to complete the test, with the exclusion of the survey items. Thus, the final version of the pre-test part II consisted of the three main tasks discussed earlier, with two sub-tasks each, which students were expected to complete in 30 minutes. Most of the items were selected from the Geometry Proof Test II, year 8 and year 9 proof surveys, developed by Martin and McCrone for their longitudinal proof project (1999-2003), and some were adapted from Chazan (1989) and Knuth (1999), and items suggested by Knuth (personal communication, July, 2004) while reviewing. The final version, which was administered to the participating students, is provided in Appendix D.

Post-test part II was developed by the researcher in the same manner. The first draft of the post-test included two sections. The first section was the *proof construction section* with three tasks similar to the three tasks of the pre-test. The second section was the proof writing section. The four tasks in this section ventured to investigate students' ability to write proofs. These tasks aimed at determining students' geometric content knowledge, ability to use logical implications, and ability to form deductive chains of reasoning to prove a statement. Table 5 illustrates the comparative structure of the pre- and post-tests part II.

Table 5: Comparative Structure of Pre- and Post- tests Part II

Task	Pretest II	Posttest II
	Investigation	Investigation
1.1 & 1.2	Spatial visualization, Ordinal analysis; Ability to use counter-examples.	Spatial visualization, Ordinal analysis; Ability to use counter-examples.
2.1 & 2.2	Ability to analyze conditional statements; Ability to use and analyze quantifiers	Ability to analyze conditional statements; Ability to use and analyze quantifiers
3.1 & 3.2	Ability to analyze arguments; Ability to understand the importance of deductive reasoning and the inadequacy of few examples to prove a statement.	Ability to analyze arguments; Ability to understand the importance of deductive reasoning and the inadequacy of few examples to prove a statement.
4		Ability to write a given statement in conditional (if-then) form; represent the given information in a diagram; identify and write explicitly the antecedent (Given information) and consequent (To prove).
5		Fill in the blanks proof
6		Synthetic proof with a hint.
7		Synthetic proof without hint

Thus, while pre-test II was a *proof construction test*, post-test II included two sections: the first section was a *proof construction test* and the second section was a *proof writing test*. After a review of peers and faculty members and their feedback, the first draft was revised and the second version was forwarded to mathematics educators at other universities. E. G. Knuth (personal communication, March 3, 2005) suggested some revisions to make the instrument less technical and more understandable to students. He also suggested inclusion of some items to test students' declarative knowledge, i.e., asking a few questions early on about the facts that they would need to know later in the test. T.S. Martin (personal communication, March 2, 2005) provided some improvements in the items so as to match the difficulty levels of the pre- and post- proof construction tasks, thereby helping the validity and consistency of the instrument.

Feedback and advice for improvement received from T. S. Martin (personal communication, March 2, 2005) include (a) varying the order of presenting the arguments in task 3, (b) maintaining the difficulty level of the tasks consistent with the corresponding tasks in the pre-test II; (c) revisiting the arguments to make them clearer and more parallel to those in the pre-test, and (d) making sure to record the various factors that may influence the results as limitations. The test instrument was revised based on the feedback of these educators and then pilot tested with the same students who took the pilot pre-test at PHS. Of the 69 students who took the pre-test, 66 wrote the post-test. The completed test papers were graded by two independent raters, one of them being the researcher herself.

The scoring scheme prepared by the researcher for both pre-test and post-test part II instruments was based on the analytical scoring scheme as suggested by McBride and Carifio (1995). Details of the scoring scheme are provided in the following section. Analysis of the pilot test scores established item-wise inter-rater reliability ranging between .89 and .98. The average of the scores awarded for each student by the two raters were analyzed, and a reliability coefficient of .85 was obtained. The final version of the post-test was prepared after making revisions as suggested by the teachers who administered the pilot test. Factor analysis extracted two factors, one for each section. This confirmed that the first three tasks tested a common underlying dimension, and tasks 4-7 tested a common dimension. The final version of the post-test II instrument was developed after making further revisions based on the feedback provided by the teachers who had offered to pilot test the instrument.

The final version of the post-test consisted of seven tasks. Tasks 1-3 were similar to the tasks 1-3 of the pre-test II. Task 4 presented two statements and asked the students to (a) write each of the statements in conditional form, (b) draw a figure representing the given information, appropriately labeling the information, (c) state the hypothesis, and (d) state the conclusion to be proved.

Task 5 was a *fill-in-the-blank* proof item. Tasks 6 and 7 presented statements accompanied by diagrams, and required students to craft proofs. While task 6 provided a hint, task 7 did not give any hint to help students with the proof. Most of the items in the instrument were selected from the Geometry Proof Test II and year 9 proof survey developed by Martin & McCrone, for their longitudinal proof project (1999-2003), and some were adapted from Chazan (1989) and Knuth (1999), and items suggested by Knuth (personal communication, March 3, 2005) while reviewing. The final version is included in the report as Appendix E.

Development of Scoring Scheme

For scoring the pre- and post- test part I (TOLT) responses, the scoring scheme provided by Tobin and Capie (1981) was used. Each of the test instruments contained ten items. The first eight items in each test were multiple-choice items, with five choices for answer and five choices for reason. A response with the correct choice of the answer as well as the reason was considered a correct response and awarded one point. For the last two items in each test, a student was awarded a point for an item only if all the combinations recorded by the student were correct and no repetitions were observed. Thus, students were awarded points ranging from 0 to 10 in each test, based on the number of items for which they had recorded correct responses.

The scoring scheme for the *proof construction tests* – pre-test part II and the first section of post-test II – was developed using a combination of the analytic scoring scale developed by McBride and Carifio (1995), Ireland (1973), and McCrone and Martin (2004). The results of the study conducted by McBride and Carifio (1995), to compare dichotomous, holistic, and analytic scoring methods confirmed that the analytic scoring scale developed by them had better reliability, validity, and internal consistency (McBride & Carifio). Moreover, an analytic method of scoring generated more normal distributions than the other two methods. Their findings also confirmed that (a) the analytic scoring criteria made possible a better interpretation of students' performance; (b) the scoring criteria themselves could be used as both independent and dependent variables; and (c) the normal distribution facilitates analyses such as ANOVA (McBride & Carifio). After considering the scoring criteria provided by McBride & Carifio, and guidelines provided by Martin & McCrone (2004) and Ireland (1973), the following scoring scheme was developed:

Code 0: Incorrect answer and invalid evidence/reasoning.

Code 1: Correct answer and invalid evidence/reasoning or incorrect answer and slightly valid reasoning

Code 2: Correct answer and vague/ambiguous reasoning.

Code 3: Correct answer and slightly valid reasoning.

Code 4: Correct answer and valid reasoning at the specific level (Local deduction).

Code 5: Correct answer and valid reasoning at a general level (Global deduction)

Students' responses to each item were rated using points 0-5. Average of the scores awarded by the two raters for each student for each item was taken as the student's score for the particular item. Thus, the total score of a student for each *proof construction test* ranged from 0 to 30. Task 4 of the post-test II presented two statements and required students to write the conditional form, draw appropriate labeled diagram, and identify the hypothesis and claim parts of the statement. A correct conditional statement and an appropriate, labeled, diagram earned two points each. The correct identifications of hypothesis and claim parts were awarded one point each. Thus the total points a student could earn in task 4 ranged from 0 to 12. Tasks 5 presented a proof, with five blanks, which students were required to fill in with appropriate responses. A correct response to a blank earned the student one point. Thus a maximum of 5 points were awarded to task 5. Tasks 6 and 7 were proof items. While task 6 provided a hint, task 7 did not provide any hint. Each of these two items was awarded a maximum of five points, as per the following coding criteria, based on the performance levels determined by Senk (1982) and McBride and Carifio (1995):

Code 0: No work or meaningless work done.

Code 1: Understanding of what is expected was evident, but has no clue about how to proceed.

Code 2: Understanding of what is expected is evident, but confusion on how to proceed.

Code 3: Understanding of correct procedure, but major errors invalidated the proof.

Code 4: Complete proof, but with minor errors.

Code 5: Complete and valid proof.

The sum of points awarded for tasks 4-7 in the *proof writing* section of the post-test II ranged from 0 to 27. Thus students were awarded points ranging from 0 to 57 for the post-test II.

Data Collection Procedure

In response to the request from the researcher seeking participation of the geometry students in the study, the mathematics department chairs of Sch1 and Sch2 invited the researcher to their monthly staff meetings. The researcher accepted the invitation and attended the staff meetings. During the meetings, the researcher presented a brief narrative about the project and procedure, and sought the help and cooperation of the geometry teachers for the success of the project. Three geometry teachers of Sch1 (T1, T2, and T3) and eight geometry teachers of Sch2 (T4-T11) offered to administer the pre- and post- tests to their classes.

The chair of Sch3 provided the email addresses of the three geometry teachers of the main campus and three geometry teachers of the ninth grade center. Subsequently, the researcher contacted the six teachers by email with a request to participate in the study. All three teachers from the main campus (T12-T14) and two teachers from the ninth grade center (T15 and T16) agreed to administer the pre- and post- tests to their geometry classes. The mathematics department chair of Sch4 was the only geometry teacher (T17) in Sch4. The chair responded positively to the researcher's email request, and offered to administer the tests to the geometry students.

Pre-tests I and II were distributed to the teachers of all participating schools by the third week of September and first week of October 2004, respectively. The pre-tests could not be administered earlier as planned by the researcher due to the unexpected school closures and subsequent need to "catch-up" faced by the teachers as a result of a series of hurricanes that battered Florida from late August through early September 2004. Post-tests parts I and II were administered the last week of April and first week of May 2005. The completed test papers were collected toward the end of October 2004 and May 2005 respectively. After assigning codes to each participating student, all papers were blind coded by two raters independently as per the scoring scheme. While the researcher was one rater for all tests, the second raters were selected from graduate students including two doctoral students. The researcher personally met each rater and explained the grading scheme and provided a copy of the detailed grading scheme. The researcher also contacted the participating teachers after the administration of tests to check the smooth conduct of the tests. Informal conversations during these visits were focused on teachers' instructional procedures, in particular with those teachers who were using GSP.

Data Analysis

Due to the relatively small number of students in the GSP group as compared to the non-GSP group, the researcher decided to grade all the pre-test papers. The grades assigned by both raters for both part I and part II tests were entered into SPSS for analysis. The average of the scores awarded by each rater for a task was taken to represent an individual student's score for each item. Statistical analysis to investigate the research questions were done as follows:

In order to investigate research question 1, the scores of students in the *proof construction tests* were used to create two dependent variables PROOF FCONSTRUCTION 1 and PROOF CONSTRUCTION 2. Types of courses formed one independent variable with three levels, regular (R), mastery (M), and honors (H). The second independent variable was the defined by the use of GSP, with two levels, non-GSP and GSP groups. Statistical analysis was completed using descriptive statistics and repeated measures ANOVA tests. Investigation of research question 2 was conducted in a similar manner after creating two dependent variables LOGICAL THINKING 1 and LOGICAL THINKING 2 using the scores of participating students in the pre- and post- tests part I. The correlation between students' performance on the tests of logical thinking and proof construction was analyzed using all the four variables created for the investigation of research questions 1 and 2. In addition, a fifth dependent variable was created (PROOF POSTTEST) using the total scores of students in both *proof construction* and *proof writing* sections of the post-test II, and a correlation analysis was done by replacing the variable PROOF CONSTRUCTION 2 with this new variable. Lastly, students' performance on the proof test at the end of the year was investigated by means of descriptive statistics using three variables. A new variable (PROOF WRITING) besides PROOF CONSTRUCTION 2 and PROOF POSTTEST was created for this purpose, representing students' total scores in the *proof writing* section of the post-test II. Figure 1 illustrates the variables created for the analyses of the data.

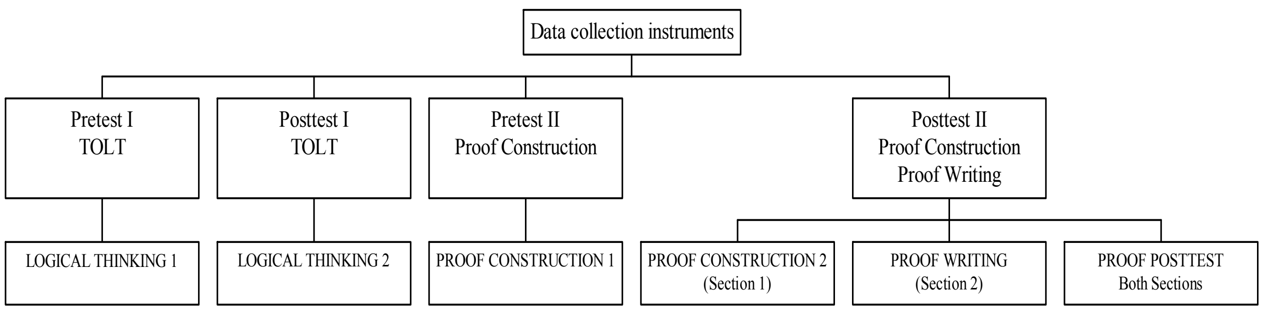


Figure 1: Details of Dependent Variables for the Data Analyses

CHAPTER 4: DATA ANALYSIS

Chapter Overview

The purpose of the current study was to investigate (a) the role of a yearlong geometry course in developing logical reasoning and proof-construction abilities in students, (b) the linkage, if any, between students' logical reasoning and proof writing abilities, and (c) the impact of dynamic geometry software in developing these abilities. In addition, this study also ventured to determine if the type of geometry course had any impact on students' logical thinking and proof writing abilities. As the abilities under consideration were complex, students' performance on the tests of logical thinking and proof at the beginning and at the end of the year long geometry course were used to determine a measure of these abilities.

The sample for this study consisted of 1,325 high school geometry students enrolled for the 2004-2005 academic year at 4 public high schools and 1 ninth grade center attached to one of the schools. All the schools were affiliated with the Local Education Agency (LEA). The study was primarily investigational, with pre- post- design. Data were collected by means of two pre-tests and two post-tests. Part I of the pre- and post- tests were logical thinking, and part II of the pre- and post- tests were proof tests. This chapter presents the results of the analyses of the data collected from the pre- and post- tests to investigate the research questions and is followed by a summary of findings based on the analyses.

Analyses of the Test Scores

As described in Chapter 3, six dependent variables and two independent variables were created in order to analyze the data in light of the research questions. The research questions guiding the analyses are reiterated below for ready reference.

1. Is there a significant difference in the performance of high school geometry students between the proof tests administered at the beginning and at the end of a yearlong geometry coursework based on
   (a) the period of instruction?
   (b) use of dynamic geometry software?
   (c) type of geometry course?
   (d) an interaction between the use of dynamic geometry software and the type of geometry course?
2. Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework based on:
   (a) the period of instruction?
   (b) use of dynamic geometry software?
   (c) type of geometry course?
   (d) an interaction between the use of dynamic geometry software and the type of geometry course?

3. Is there a correlation between the students' performances on the tests of logical thinking and proof construction, administered at the beginning and at the end of a yearlong geometry coursework?
4. How well do high school geometry students learn to construct geometry proofs by the end of the high school geometry course?

Of the 1,325 students who participated in the study, 1,159 students completed the pre-test I (logical thinking pre-test), 1,124 students wrote pre-test II (proof construction pre-test), 1,031 responded to post-test 1(logical thinking post-test), and 971 took the proof construction section of post-test II. From the data collected, four variables were created initially: LOGICAL THINKING 1 represented the students' cumulative scores in the pre-test I; LOGICAL THINKING 2 represented students' scores in post-test I, PROOF CONSTRUCTION 1 was made of students' scores in pre-test II, and PROOF CONSTRUCTION 2 consisted of students' scores in the proof construction section of post-test II. The number of cases in each of the variables were 1,159, 1,031, 1,124, and 971 respectively. However, for the purpose of analyses of research questions 1-3, only the scores of students who wrote both pre- and post- tests in each part were considered. In addition to these four variables, two more variables were created to aid analyses of research questions 3 and 4. PROOF POSTTEST included students' cumulative scores in the two sections of post-test II, and PROOF WRITING represented their scores in the proof writing section of the post-test II. Table 6 illustrates the descriptive statistics of the six dependent variables.

Table 6: Descriptive Statistics for Pre- and Post- tests

Variable	N	Maximum points	Mean	SD
PROOF CONSTRUCTION 1	841	24	9.34	4.28
PROOF CONSTRUCTION 2	841	29	10.02	5.59
LOGICAL THINKING 1	931	10	4.90	3.09
LOGICAL THINKING 2	931	10	5.03	3.13
PROOF POSTTEST	949	49	18.38	10.92
PROOF WRITING	683	27	11.41	5.96

Analysis of Table 6 indicated that the mean scores of students in the pre- and post- *proof construction* tests showed an increase. Similarly, the mean scores of the students who took both the pre- post- tests of part I (LOGICAL THINKING 1 and LOGICAL THINKING 2) also showed a marginal increase. The standard deviations for these variables suggested that the variance in the pre- and post- *proof construction* tests increased substantially, while the variance in the pre- and post- tests of logical thinking increased marginally.

Investigation of Students' Performance on Tests

Comparative performance of students on tests of proof construction and logical thinking were investigated by means of repeated measures analysis of variance using General Linear Model. Repeated measures design was used because the same subjects were tested under different conditions. Moreover, the General Linear Model of the repeated measures test affords the most valid results when the levels in the factors are unequal, like the factors involved in the current research (Becker, 1999). This procedure reduces the error variance caused by between-group differences (Field, n.d.). Repeated measures ANOVA holds only if the variances of difference between the two levels of treatments are equal. This equality of variances of the differences between the treatment levels is known as Sphericity, also sometimes called circularity. Mauchly's test for sphericity examines the hypothesis that the variances of the differences between the conditions are equal. If the tests showed a probability of more than .05, homogeneity of error variance could be assumed and if the probability were less than .05, homogeneity of error variance could not be assumed; and accordingly, a revision/correction of F- ratios could be considered (Field).

Research questions 1 and 2 were investigated using the scores of students who had taken all the four tests, with the intention of controlling for the mortality of subjects to some extent. It was interesting to observe that only 689 students had written all four of the tests, although 841 students had written both pre- and post- tests of *proof construction* and 931 students had responded to both the pre- and post- tests of logical thinking. Some of the many reasons for this high rate of mortality could be (a) student mobility during the academic year, (b) absence during the days in which the tests were administered, and (c) students' disinterest to respond to the tests due to various reasons. Teachers were instructed to inform the students before taking the test that their responses to the tests were requested for research purposes and the points earned by them on these tests would not be considered for their grades. This could have been a non-motivating factor, leading to the high mortality rate.

Repeated measures tests were conducted to investigate the overall effect of the period of instruction, the type of course, the use of GSP, and the interaction of these two factors on the performance of students on the tests. For investigating students' performance on proof construction tests, the dependent variables PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2 were entered as the dependent factor PRFCONST with two levels and LOGICAL THINKING 1 and LOGICAL THINKING 2 were entered as the two levels of the dependent factor LOGTHINK for investigating the performance of students on tests of logical thinking. The variables COURSE with three levels (Mastery, Regular, and Honors) and GROUP with two levels (non-GSP and GSP) were entered as independent factors.

Of the 689 students who wrote all four tests, 51 were enrolled in mastery, 290 in regular, and 348 in honors courses. Mastery students were made up of 14 non-GSP and 37 GSP groups. Regular students were comprised of 192 non-GSP and 98 GSP groups. Among honors students, 191 were in the non-GSP, and 157 were in the GSP groups. Overall, 397 non-GSP and 292 GSP students took all of the four tests. Table 7 shows the within-subjects effects for both LOGTHINK and PRFCONST along with the multivariate test Wilks' Lambda value.

The repeated measures analysis for the dependent factor PRFCONST showed a significant difference in the variance of scores between the two levels PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2 based on the period of instruction ($F = 5.04$, $p < .05$), but only about 1% of the variance could be accounted for by the period of instruction. The Wilks' Lambda value showed that about 99% of the variance in scores was unaccounted for by the period of geometry instruction.

Table 7: Wilks' Lambda and Within-Subjects Effects for Each Factor

Factors	Wilks' Lambda	Type III SS	F	Partial Eta Squared
PRFCONST	0.99	35.99[a]	5.04	0.01
PRFCONST x COURSE	0.97	162.66[b]	11.39**	0.03
PRFCONST x GROUP	1.00	1.70[a]	0.24	0.00
PRFCONST x COURSE x GROUP	0.98	74.20[b]	5.20**	0.02
LOGTHINK	0.61	5113.79[a]	431.12**	0.39
LOGTHINK x COURSE	0.98	188.52[b]	7.95**	0.02
LOGTHINK x GROUP	1.00	16.79[a]	1.42	0.00
LOGTHINK x COURSE x GROUP	1.00	4.32[b]	0.18	0.00

Note. Error degrees of freedom for PRFWRITE = 835.

Error degrees of freedom for LOGTHINK = 925.

[a] degree of freedom = 1

[b] degree of freedom = 2

* $p < .05$

** $p < .01$

Data showing within-subjects effects showed the following results:

(a) There was a statistically significant difference in the variance in scores of students between the pre- and post- proof construction tests based on the type of course ($F = 11.39$, $df = 2$, $p < .01$). Approximately 3% of this variance could be accounted for by the type of the course. Wilks' Lambda value showed that about 97% of the variance is unaccounted for by the type of course, confirming this observation.

(b) There was no statistically significant difference in the variance of scores of students between the pre- and post- - proof construction tests based on the use of GSP ($F = 0.24$, $df = 1$, $p > .05$). This observation was also confirmed by the Wilks' lambda value (1.000) and the partial eta squared value (0.000).

(c) There was a statistically significant difference in the variance of scores of students between the pre- and post- - proof construction tests based on an interaction between the type of course and use of GSP with the period of instruction ($F = 5.20$, $df = 2$, $p < .01$). However, only 2% of the variance could be accounted for by this interaction, as was confirmed by the Wilks' lambda value (.98).

Within-subjects effects for LOGTHINK showed a significant difference in the variance of the scores based on the period of instruction ($F = 431.12$, $p < .01$) and type of course ($F = 7.95$, $p < .01$), as against highly insignificant ($p > .01$) F values for the use of GSP and the interaction between type of course and use of GSP. Moreover, the percentage of variance accounted for by the period of instruction (main effect) was 39% whereas the type of course accounted for only 2% of the variance in scores between the pre- and post- tests of logical thinking. Further investigations using profile plots showing estimated marginal means were desired to get a deeper insight into the behavior of the variables in these two analyses.

Estimated marginal means, also called unweighted means, are more reliable measures when the factors have unequal groups which is absolutely applicable to the current analysis, as the number of students in each type of course and each group are different (Field, n.d.).

Figures 2 and 3 illustrate the estimated marginal means of the three sub groups of COURSE within the two levels of GROUPS: non- GSP (Group A) and GSP (Group B) respectively.

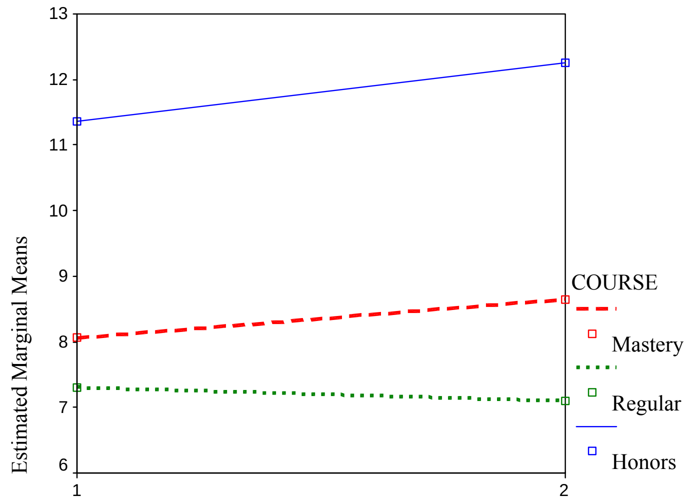


Figure 1: (M = 17; R = 236; H = 237)

Figure 2: Estimated Marginal Means for PRFCONST at GROUP: Non-GSP

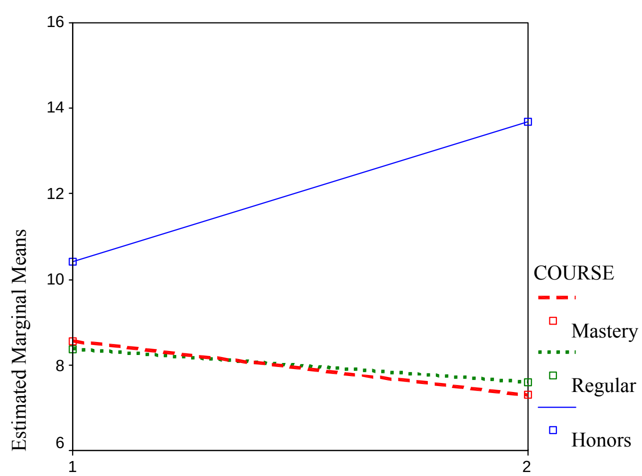


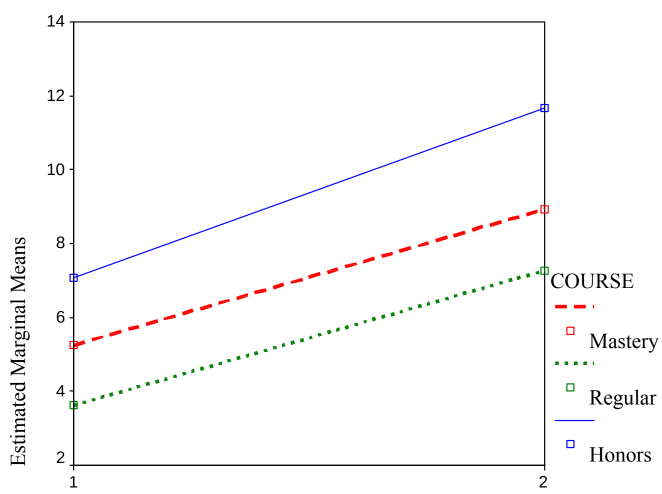
Figure 2: (M = 54; R = 125; H = 172)

Figure 3: Estimated Marginal Means for PRFCONST at GROUP: GSP

The profile plot in Figure 2 illustrating changes in the estimated marginal means for the three types of courses between PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2 for the non-GSP group showed: (a) Honors students have significantly outperformed their counterparts in mastery and regular courses; (b) Mastery students scored a higher mean than regular students in both tests; and (c) While both honors and mastery students had an increase in mean in the post-test, regular students showed a slight decline in the mean score. The profile plot in Figure 3 illustrating the changes in the estimated marginal means for the three types of course between PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2 for the GSP group showed: (a) Honors students have significantly outperformed their counterparts in mastery and regular courses; (b) While mastery students had a marginally higher mean in the pre-test, their mean dropped slightly below that of regular students' in the post-test; and (c) Both regular and mastery students showed a decline in their mean from pre-test to post-test.

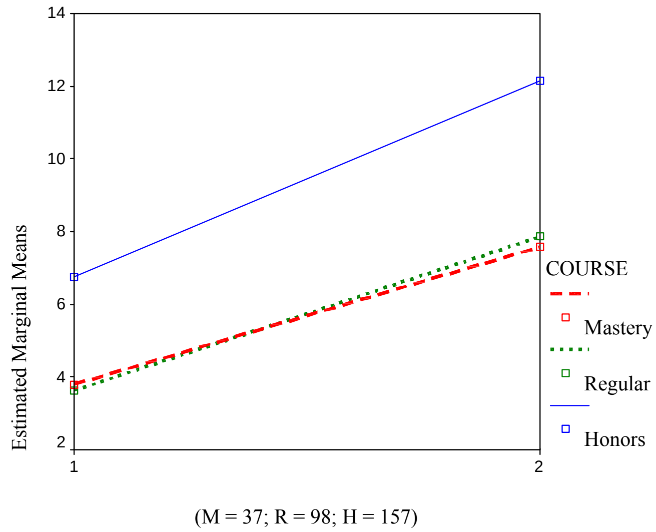
A closer observation of both Figures 2 and 3 indicated that GSP students in the honors course had benefited more by the period of instruction than the GSP students in regular and mastery courses. At the same time, regular and mastery students in the non-GSP group had shown better response to the instruction than their GSP counterparts. This observation suggested either that GSP was used more purposefully with the honors students, or honors students were more motivated to exploit the GSP environment to improve their knowledge. As there was a significant interaction effect between the variables COURSE and GROUP with the variable COURSE having three levels, a post hoc comparison of means of the three levels of COURSE was done to obtain a deeper insight into the performance of the students enrolled in different types of courses. Scheffe's test was selected for this purpose because the sizes of the levels were unequal. Sheffe's test is the most widely acceptable post hoc comparison of statistical means and of complex combinations of means for unequal group sizes (Shavelson, 1996). Sheffe's post hoc comparison of means indicated a significant difference in the mean scores between (a) honors and mastery courses, and (b) honors and regular courses. Both differences were significant at the level of .01. There was no significant difference in the means between regular and mastery courses ($p > .05$). The post hoc tests thus confirmed the observations derived from the two profile plots presented below.

The estimated marginal means for the tests of logical thinking did not show much difference in the performance of the students based on the types of courses and groups, which confirmed the results of the analysis for the variable LOGTHINK. However, the profile plots showing the estimated marginal means for the GSP group showed some interesting results. Figures 4 and 5 illustrate the estimated marginal means of students enrolled in the three different types of courses based on the two groups for the two tests of logical thinking.



(M = 14; R = 192; H = 191)

Figure 4: Estimated Marginal Means for LOGTHINK at GROUP: Non-GSP



:

Figure 5: Estimated Marginal Means for LOGTHINK at GROUP: GSP

A scrutiny of the profile plots in Figures 4 and 5 revealed that (a) the differences in the estimated marginal means for all the three courses in the non-GSP group were almost equal, with the regular course students earning the lowest mean both in pre-test and post-test (Figure 4), and (b) while mastery students had a slightly higher mean than regular students in the pre-test, regular students scored a slightly higher mean than mastery students in the post-test (Figure 5.

Relation between Students' Performance on Logical Thinking and Proof Tests

In order to investigate the relation between the four variables (PROOF CONSTRUCTION 1, PROOF CONSTRUCTION 2, LOGICAL THINKING 1, and LOGICAL THINKING 2), multiple regression seemed a good option, as more than two variables were involved. Multiple regression is a useful test to learn more about the relationships among several independent variables (Shavelson, 1996). The term multiple regression was first used by Pearson, in 1908 (StatSoft, n.d.). The main purpose of using multiple regression is prediction. In order to obtain a good insight into the relationship among the four variables under consideration, three multiple regression analyses were done. The first analysis explored the predictability of a student's score in LOGICAL THINKING 2 if the that student's scores in the other three variables LOGICAL THINKING 1, PROOF CONSTRUCTION 1, and PROOF CONSTRUCTION 2 were known. The second analysis explored the possibility of predicting a student's score in PROOF CONSTRUCTION 2 when the scores of PROOF CONSTRUCTION 1, LOGICAL THINKING 1, and LOGICAL THINKING 2 were known. The third analysis ventured to predict the a student's total scores in the PROOF POST-TEST if that student's total scores in PROOF CONSTRUCTION 1, LOGICAL THINKING 1, and LOGICAL THINKING 2 were known.

For the first analysis, LOGICAL THINKING 2 was taken as the dependent (outcome) variable, and the other three variables were entered as predictor variables, with LOGICAL THINKING 1 in the first block and PROOF CONSTRUCTION 1 & PROOF CONSTRUCTION 2 in the second block. For the second analysis, PROOF CONSTRUCTION 2 was taken as the outcome variable, PROOF CONSTRUCTION 1 as the predictor variable in the first block, and LOGICAL THINKING 1 and LOGICAL THINKING 2 were included as predictor variables in the second block. Table 9 presents the pair-wise correlations among the four variables for these two analyses.

Table 8: Pair-wise Pearson Correlations for Analyses 1 and 2

Variables	LOGICAL THINKING 1	LOGICAL THINKING 2	PROOF CONSTRUCTION 1	PROOF CONSTRUCTION 2
Analysis # 1				
LOGICAL THINKING 1	1.00	0.73**	0.40**	0.50**
LOGICAL THINKING 2	0.73**	1.00	0.42**	0.52**
PROOF CONSTRUCTION 1	0.40**	0.42**	1.00	0.48**
PROOF CONSTRUCTION 2	0.50**	0.52**	0.48**	1.00
Analysis # 2				
LOGICAL THINKING 1	1.00	0.73**	0.40**	0.50**
LOGICAL THINKING 2	0.73**	1.00	0.42**	0.52**
PROOF CONSTRUCTION 1	0.40**	0.42**	1.00	0.50**
PROOF CONSTRUCTION 2	0.50**	0.52**	0.50**	1.00

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

As evident from Table 8, the Pearson Correlation tests showed that (a) LOGICAL THINKING 1 correlated the most with LOGICAL THINKING 2 ($r = .73$, $p < .01$), then with PROOF CONSTRUCTION 2 ($r = .50$, $p < .01$) and PROOF CONSTRUCTION 1 (r .40, $p < .01$) in both analyses; (b) Both analyses indicated a significant positive correlation between LOGICAL THINKING 1 and PROOF CONSTRUCTION 1 ($r = .42$, $0 < .01$), and LOGICAL THINKING 2 and PROOF CONSTRUCTION 2 ($r = .52$, $p < .01$); (c) There was a significant positive correlation between PROOF CONSTRUCTION 1 and PROOF

CONSTRUCTION 2. A close look at Table 8 shows slight difference in the correlation coefficient between PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2 between the two analyses. This difference may be attributed to the fact that PROOF CONSTRUCTION 2 was entered as a predictor along with PROOF CONSTRUCTION 1 in the first analysis, whereas it was the outcome variable in the second analysis.

For the third and final analysis, the total of all tasks of post-test II (Total 57 points) was taken as dependent variable (PROOF POST-TEST), with PROOF CONSTRUCTION 1 in the first block, and LOGICAL THINKING 1 and LOGICAL THINKING 2 in the second block of predictor variables. Of the total 1,325 participants, SPSS identified 694 valid cases for all four variables. Table 9 illustrates the pair-wise correlations between the variables the third analysis.

Table 9: Pair-wise Pearson Correlations for Analysis 3

Variable	PROOF POST-TEST	PROOF CONSTRUCTION 1	LOGICAL THINKING 1	LOGICAL THINKING 2
PROOF POST-TEST	1.00	0.53	0.56**	0.54**
PROOF CONSTRUCTION 1	0.53**	1.00	0.43**	0.46**
LOGICAL THINKING 1	0.56**	0.43**	1.00	0.73**
LOGICAL THINKING 2	0.54**	0.46**	0.73**	1.00

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

A close look at Table 9 showed that (a) LOGICAL THINKING 1 correlated the most with LOGICAL THINKING 2 ($r = .73, p < .01$), then with PROOF POST-TEST ($r = .56, p < .01$) and PROOF CONSTRUCTION 1 ($r\ \ .43, p < .01$); (b) There were significant positive correlations between LOGICAL THINKING 1 and PROOF CONSTRUCTION 1 ($r = .46$, $0 < .01$), and LOGICAL THINKING 2 and PROOF POST-TEST ($r = .54$, $p < .01$); and (c) There was a significant positive correlation between PROOF CONSTRUCTION 1 and PROOF POST-TEST ($r = .53$, $p < .01$). Table 10 presents the model summary for each analysis, showing the multiple correlations between the value of the outcome variable and its predicted values based on the factors (R), and the proportion of variance in the outcome variable accounted for by the other variables in the independent factors (R^2) for each of the three analyses.

Table 10: Model Summary for the Outcome Variables

Outcome Variables	Predictor Variables	R	R Square	R Square Change
LOGICAL	LOGICAL THINKING 1	0.73[a1]	0.54	0.54
THINKING 2	PROOF CONSTRUCTION 1	0.76[b1]	0.57	0.04
	PROOF CONSTRUCTION 2			
PROOF	PROOF CONSTRUCTION 1	0.50[a2]	0.25	0.25
CONSTRUCTION 2	LOGICAL THINKING 1	0.62[b2]	0.38	0.13
	LOGICAL THINKING 2			
PROOF POSTTEST	PROOF CONSTRUCTION 1	0.53[a3]	0.28	0.28
	LOGICAL THINKING 1	0.65[b3]	0.43	0.15
	LOGICAL THINKING 2			

The multiple correlation coefficients (R Values) for the first analysis, as indicated by Table 10, showed that coefficient of correlation between the value of the dependent variable LOGICAL THINKING 2 and its predicted value based on LOGICAL THINKING 1 was .73, and the coefficient based on the combined effects of all three predictor variables was .76. Moreover, Table 10 also showed that, of the total of 58% of the variance in LOGICAL THINKING 2 that could be accounted for by the three predictor variables, approximately 54% was accounted for by LOGICAL THINKING 1, and the remaining 4% by the combined effects of PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2.

For the second analysis, the coefficients of multiple correlation between the value of the outcome variable PROOF CONSTRUCTION 2 and its predicted value based on PROOF CONSTRUCTION 1 was .50 (Table 10), and the coefficient based on the combined effects of all three predictor variables was .62. The multiple correlation coefficients for the third analysis showed that the coefficient of correlation between the value of the outcome variable PROOF POST-TEST and its predicted value based on PROOF CONSTRUCTION 1 was .53, and the coefficient based on the combined effects of all three predictor variables was .65. Moreover, a total of 43% variance in the outcome variable could be accounted for by the three predictor variables, out of which 28% could be accounted for by PROOF CONSTRUCTION 1, and the remaining 15% by the combined effects of LOGICAL THINKING 1 and LOGICAL THINKING 2.

Table 11 presents the coefficients and constants that could be applied to predict a student's score in the outcome variable in each of the three analyses if that student's scores in the three predictor variables were known. The t-values in the last column confirm that the differences in the mean scores of these variables are all statistically significant ($p < .01$ for all the t-values).

Table 11: Constants and Coefficients to Predict the Outcome Variables

Analysis #	Predictor Variables	Unstandardized Coefficients		Standardized Coefficients	t
		B	Std. Error	Beta	
Analysis # 1	(Constant)	1.52	0.16		9.58**
	LOGICAL THINKING 1	0.74	0.03	0.73	28.23**
	(Constant)	0.63	0.20		3.11**
	LOGICAL THINKING 1	0.62	0.03	0.61	20.90**
	PROOF CONSTRUCTION 1	0.06	0.02	0.09	2.96**
	PROOF CONSTRUCTION 2	0.09	0.02	0.17	5.41**
Analysis # 2	(Constant)	4.09	0.44		9.34**
	PROOF CONSTRUCTION 1	0.65	0.04	0.50	15.27**
	(Constant)	2.03	0.43		4.70**
	PROOF CONSTRUCTION 1	0.42	0.04	0.33	9.75**
	LOGICAL THINKING 1	0.35	0.08	0.19	4.24**
	LOGICAL THINKING 2	0.44	0.08	0.24	5.41**
Analysis # 3	(Constant)	6.10	0.90	0.90	6.77**
	PROOF CONSTRUCTION 1	1.19	0.07	0.07	16.24**
	(Constant)	2.55	0.84	0.84	3.03**
	PROOF CONSTRUCTION 1	0.72	0.07	0.07	9.71**
	LOGICAL THINKING 1	1.01	0.15	0.15	6.71**
	LOGICAL THINKING 2	0.66	0.15	0.15	4.35**

Analysis # 1: Outcome variable LOGICAL THINKING 2 ** $p < .01$

Analysis # 2: Outcome variable PROOF CONSTRUCTION 2

Analysis # 3: Outcome variable PROOF POST-TEST

As the variables were of unequal size, applying the standardized coefficients and constants from Table 11, the following regression equation could be formed to predict the outcome variable LOGICAL THINKING 2 from the analysis # 1:

Predicted LOGICAL THINKING 2 score = 1.52 + 0.73 (LOGICAL THINKING 1 score)

Predicted LOGICAL THINKING 2 score = 0.63 + 0.61 (LOGICAL THINKING 1 score) + 0.09 (PROOF CONSTRUCTION 1 score) + 0.17 (PROOF CONSTRUCTION 2 score).

The first equation would mean that for an increase of one unit in the LOGICAL THINKING 1 score, one could predict an increase of 0.73 units in the LOGICAL THINKING 2 score. From the second equation, it could be inferred that a one-unit increase in LOGICAL THINKING 1 would result in an increase of .61 units in the predicted LOGICAL THINKING 2 score if the other two variables were kept constant. The coefficients indicated that PROOF CONSTRUCTION 1 would contribute the least to the predicted score when other variables were kept constant. This observation confirmed the pattern observed in the correlation table. The following regression equations were formed for predicting the PROOF CONSTRUCTION 2 scores from the analysis # 2:

Predicted PROOF CONSTRUCTION 2 score = 4.09 + 0.50 (PROOF CONSTRUCTION 1 score)

Predicted PROOF CONSTRUCTION 2 score = 2.03 + 0.33 (PROOF CONSTRUCTION 1 score) + 0.19 (LOGICAL THINKING 1 score) + 0.24 (LOGICAL THINKING 2 score).

Applying a similar inference pattern to analysis # 2, it could be noted from the above equations that LOGICAL THINKING 1 contributed the least to the prediction of PROOF CONSTRUCTION 2 score when other variables were kept constant. This confirmed the observation from the correlation matrix in Table 9. Finally, the regression equations that could be obtained from the analysis # 3 were:

Predicted PROOF POST-TEST score = 0.90 + 0.07 (PROOF CONSTRUCTION 1 score)

Predicted PROOF POST-TEST score = 0.84 + 0.07 (PROOF CONSTRUCTION 1 score) + 0.15(LOGICAL THINKING 1 score) + 0.15 (LOGICAL THINKING 2 score).

In this case, it was noted that PROOF CONSTRUCTION 1 contributed the least for the prediction of PROOF POST-TEST score when the other two variables were kept constant. This was also in line with the correlation matrix in Table 9.

Students' Performance on the Proof Test at the End of the Year

The final analysis of the current study was to investigate how well high school geometry students had performed in the proof test administered at the end of the year. The proof test consisted of two sections. The first section constituted three tasks comparable to the *proof construction test* administered at the beginning of the academic year. Tasks 4-7 were *proof writing tests*. These four tasks focused on students' proof writing ability. So, it seemed most sensible to analyze these questions individually and collectively to answer the fourth and final research question. Accordingly, tasks 4-7 were analyzed both individually and collectively, using frequencies, measures of central tendency and dispersion. The total points for task 4 ranged from 0 to 12. The total points for tasks 5, 6, and 7 each ranged from 0 to 5. The first analysis investigated descriptive statistics of these four tasks based on types of courses and groups. Tables 12 and 13 provide the relevant statistics for these tasks.

Table 12: Descriptive Statistics for Tasks 4-7 Based on Groups

Task #	N		Mean		Standard Deviation	
	Non-GSP	GSP	Non-GSP	GSP	Non-GSP	GSP
Task 4	444	321	5.55	6.61	3.06	3.34
Task 5	424	310	2.23	2.56	1.25	1.64
Task 6	314	214	1.14	2.07	1.08	1.72
Task 7	291	199	1.41	2.27	1.24	1.54

Table 12 indicated that out of the 949 students who wrote the post-test II, 765 students responded to the task 4 (81%), 734 responded to task 5 (77%), 528 responded to task 6 (56%), and 490 responded to task 7 (52%). Of the 765 students who responded to task 4, 444 were in non-GSP group and 321 were in GSP group. 424 non-GSP students and 310 GSP students responded to task 5. Respondents to tasks 6 and 7 included 314 and 291 non-GSP students, and 214 and 199 GSP students respectively. GSP students seemed to have performed better than non-GSP students in each of the four tasks. Nevertheless, the scores of GSP students were more dispersed than their non-GSP counterparts in each task. It is interesting to note that only about 72% of the participants of the study completed the post-test II. Of those who wrote the post-test II, the percentage of students who responded to the tasks decreased according to the difficulty level of the task. Table 13 provides the mean scores and standard deviations for each course in each task.

Table 13: Descriptive Statistics for Tasks 4-7 Based on Courses

Task #	Mean (N)			S.D.		
	M	R	H	M	R	H
Task 4	3.64 (58)	4.64 (289)	7.26 (418)	2.92	2.88	2.92
Task 5	1.63 (56)	1.58 (270)	2.99 (408)	1.37	1.12	1.32
Task 6	0.89 (28)	0.78 (166)	1.93 (334)	1.10	0.97	1.51
Task 7	1.22 (18)	0.96 (146)	2.14 (326)	1.06	1.02	1.45

Note. M: Mastery course

R: Regular course

H: Honor course

It was interesting to observe from Table 13 that while honors students had always scored a higher mean in all the tasks, regular students had scored a higher mean than mastery students in task 4, and mastery students had scored a higher mean than regular students in all other tasks. Regular students had a lower dispersion than the other two types in each of the four tasks. Nevertheless, the overall mean scores of respondents for each task fell below 50% of the maximum points for each task. (50% for task 4, 47% for task 5, 30% for task 6 and 35% for task 7). The trend observed in the Tables 13 and 14 were further explored by means of box plots in Figures 6-13, illustrating the mean and inter-quartile range for each task, based on the type of course and the use of GSP, which are presented in the following sub sections.

Analysis of Students' Responses to Task 4.

In order to further investigate the relative performance of students based on the type of course and the use of GSP, pictorial representation of mean scores and inter-quartile range (IQR) were considered suitable measures of central tendency and dispersion respectively. Figures 6 and 7 illustrate the comparison of mean scores and the inter-quartile ranges of the two groups and the three courses respectively.

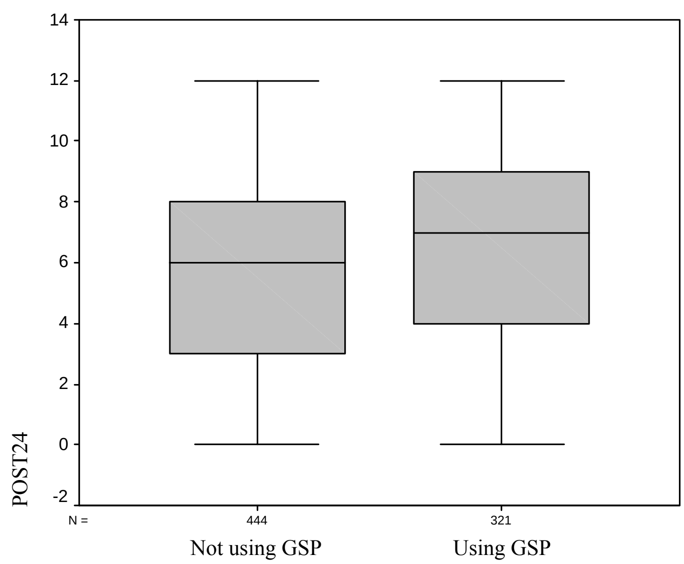


Figure 6: Comparison of Scores between Groups: Task 4

Figure 6 confirmed the data provided in Table 13. This indicated that GSP students obtained a higher mean and a higher IQR than non-GSP students. Figure 7 illustrated clearly that honors students significantly outperformed the regular and mastery students. Mastery students had the lowest mean and IQR for task 4

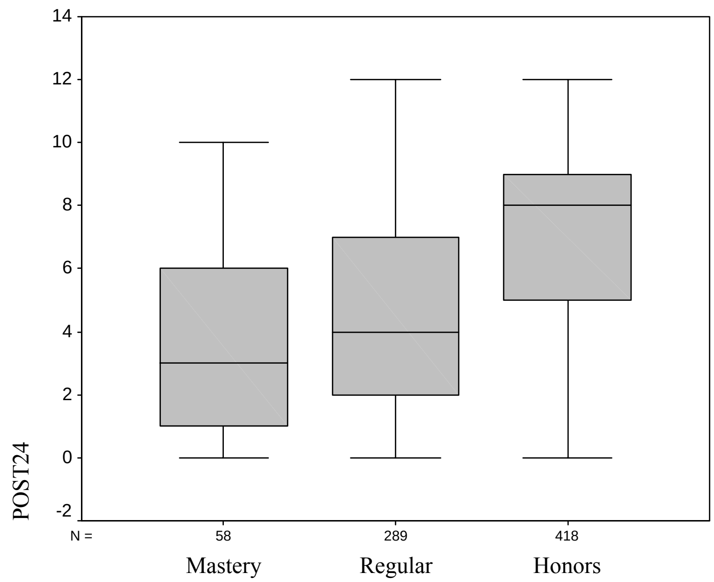


Figure 7: Comparison of Scores between Courses: Task 4

Analyses of Scores for Tasks 5-7

Of all the tasks in the post-test II, the least number of responses was observed for task 7. This task was expected to be the most difficult item. Box plots illustrating students' mean scores and inter-quartile range were used to explore the trend further for individual tasks, based on the type of course and the use of GSP. Figures 8 and 9 provide pictorial illustrations of the comparison of scores based on groups and types of courses respectively for task 5.

It is evident from Figure 8 that GSP students performed better than non-GSP students in task 5, with higher mean and IQR, which confirmed the observations from Table 12. Figure 9 confirms the data provided in Table 13, showing that honors students outperformed mastery and regular students. Moreover, while both regular and mastery students earned approximately same mean scores, the regular students spanned a lower IQR than their mastery course counterparts. It is also interesting to observe that the maximum points scored by regular students on task 5 was 3, while some mastery and honors students earned 5 points.

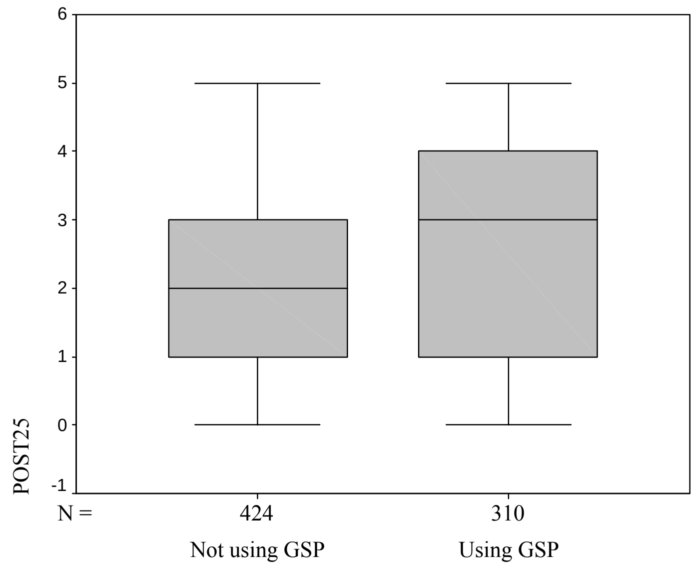


Figure 8: Comparison of Scores between Groups: Task 5

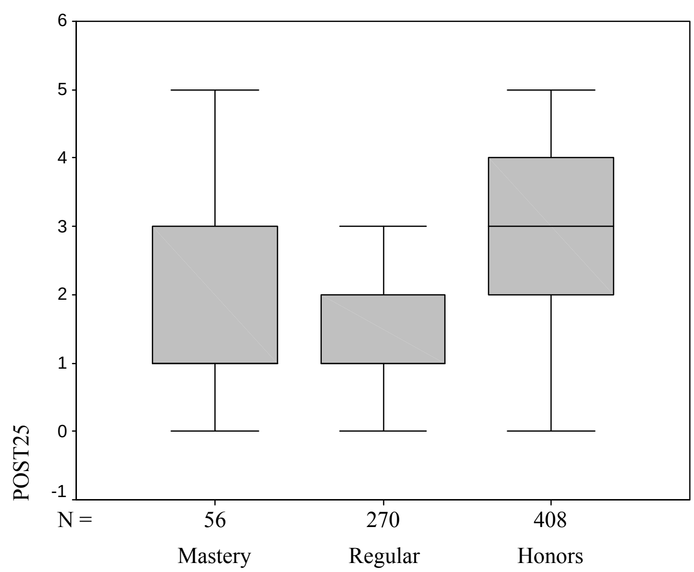


Figure 9: Comparison of Scores between Courses: Task 5

For a better understanding of the course-wise and group-wise performance of students in task 6, box plots in Figures 10 and 11, illustrating pictorially the comparison of scores based on groups and courses respectively, were used.

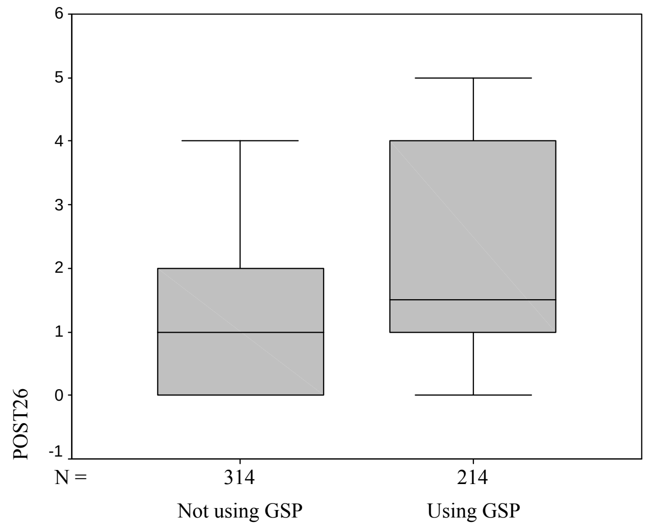


Figure 10: Comparison of Scores between Groups: Task 6

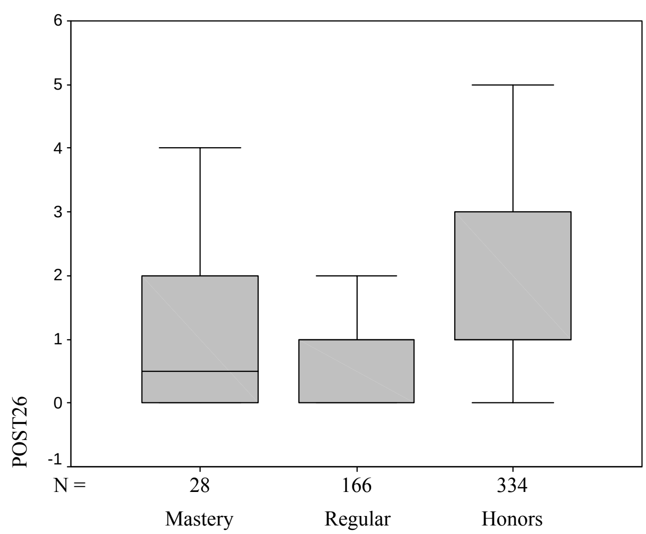


Figure 11: Comparison of Scores between Courses: Task 6

It was clear from Figure 10 that GSP students had a higher IQR than non-GSP students and a higher mean. This confirmed the observations made from Table 12 for task 6. Also, none of the non-GSP students earned the full credit of 5 points for task 6. The box plot in Figure 11 suggested that honors students scored the highest mean and IQR, whereas regular students had the lowest in both. Also, the maximum points earned by a regular course student was 2, while mastery students reached a maximum of 4 and some honors course students could get the full credit of 5 points. Figures 12 and 13 provide illustrations of comparisons of scores based on groups and courses respectively for task 7..

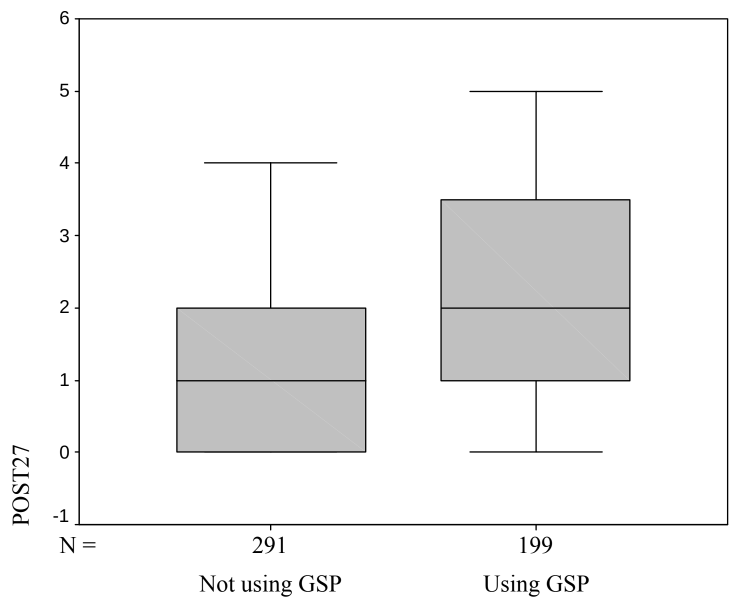


Figure 12: Comparison of Scores between Groups: Task 7

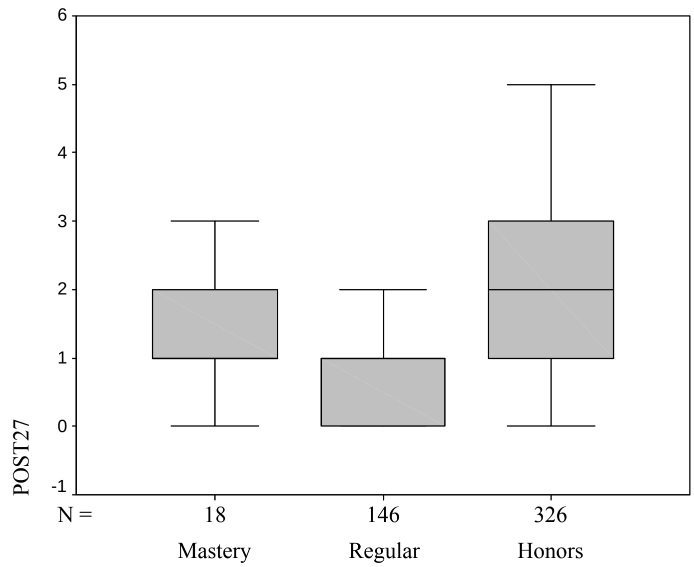


Figure 13: Comparison of Scores between Courses: Task 7

The box plots in the Figures 12 and 13 revealed that the performance of students in task 7 followed the similar patterns as their performance on task 6. GSP students showed better performance than non-GSP students, and honors students outperformed the regular and mastery students, while regular students earned the lowest mean and inter-quartile range. Because each of these tasks ventured to assess some of the skills needed for success in proof writing, students' performances on proof writing were also analyzed using the total scores obtained by adding the scores earned by each student on tasks 4, 5, 6, and 7 (PROOF WRITING), to better understand their performance on proof writing. The maximum points for PROOF WRITING thus became 27 (12 + 5 + 5 + 5 = 27). Because one dependent variable and two independent variables (COURSE and GROUP) were considered, a two-way Analysis of Variance (ANOVA) was used.

ANOVA was invented in the 1920s to partition the variance of the dependent variable into uncorrelated parts (Yvette, 1996). This test is effective and efficient for comparing the means of different levels of the dependent variable and to decide whether the differences between them represent a chance occurrence or a systematic effect (Shavelson, 1996). In the current study, two-way ANOVA was used to investigate end-of-year performance of students once with the dependent variable, PROOF WRITING, which represented students' score in the *proof writing* section, and a second time by entering PROOF POST-TEST, which included both *proof construction* and *proof writing* sections, as the dependent variable, in order to consider the overall performance of students on the post-test II.

Of the 949 students who wrote the post-test II, 683 students had responded to at least two of the four tasks in the *proof writing* section. Responses of these 683 students were included in this analysis. Of these, 394 were in the non-GSP group, and 289 were in GSP group. Moreover, 37 students were enrolled in mastery, 249 in regular, and 397 in honors courses. Table 14 presents the mean score and standard deviation of students of each group enrolled in the three types of courses under consideration.

Table 14: Descriptive Statistics for PROOF WRITING and PROOF POST-TEST

Variable	Course	Mean		Standard Deviation		Total # of students	
		Non-GSP	GSP	Non-GSP	GSP	Non-GSP	GSP
PROOF WRITING	Mastery	7.92	8.63	4.59	4.37	13	24
	Regular	7.52	8.10	4.34	5.08	147	102
	Honors	12.20	16.55	4.48	5.77	234	163
PROOF POST-TEST	Mastery	8.39	12.30	8.03	6.66	49	60
	Regular	12.35	14.36	7.42	8.30	251	145
	Honors	23.61	29.12	8.84	8.66	266	178

As is evident from Table 14, honors students had scored significantly better than mastery or regular students in both groups. Regular students had the lowest mean score in both groups. GSP students had a higher mean score than their non-GSP friends in each of the three courses. Table 15 presents the between-subjects effects of the two factors COURSE and GROUP, and an interaction effect of these two for both variables. Analyses of both variables indicate similar trends for the between-subjects effects as well as the interaction effects.

Table 15: Between-Subjects Effects for PROOF WRITING and PROOF POST-TEST

Variable	Factors	Type III Sums of Squares	F	Partial Eta Squared
PROOF WRITING	COURSE	6780.30[a]	142.56**	0.30
	GROUP	217.35[b]	9.14**	0.01
	COURSE x GROUP	566.04[a]	11.90**	0.03
PROOF POST-TEST	COURSE	42901.22[a]	319.18**	0.40
	GROUP	2276.81[b]	33.88**	0.04
	COURSE x GROUP	606.42[a]	4.51**	0.01

Error degrees of freedom = PROOF WRITING = 677; PROOF POST-TEST = 943.

[a] degree of freedom = 2

[b] degree of freedom = 1

* Significance $p < .05$

** Significance $p < .01$

As the levels in the independent variables had unequal number of cases, it was expected that observing a profile plot of estimated marginal means would contribute to explaining the performance of students by determining any discrepancies. Figures 14 and 15 present the profile plots illustrating the estimated marginal means of each type of courses in each group for the two variables.

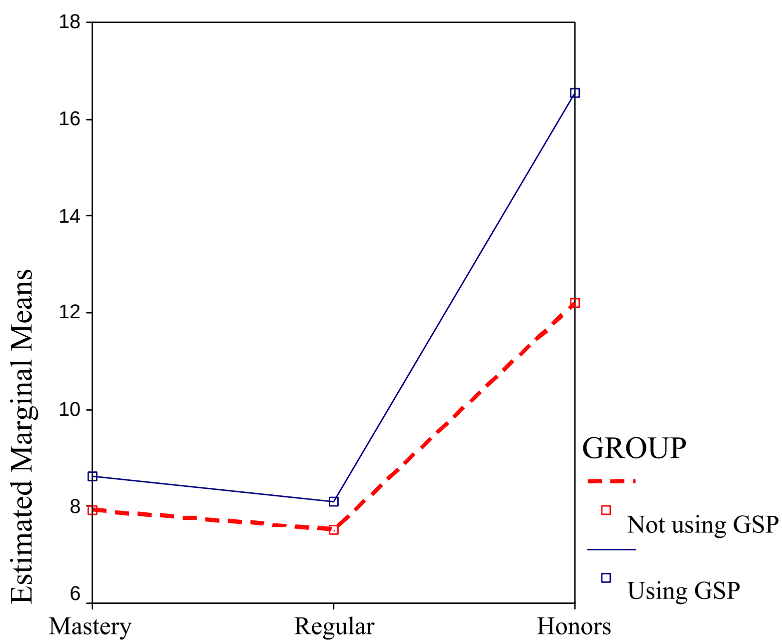


Figure 14: Estimated Marginal Means based on Courses: PROOF WRITING

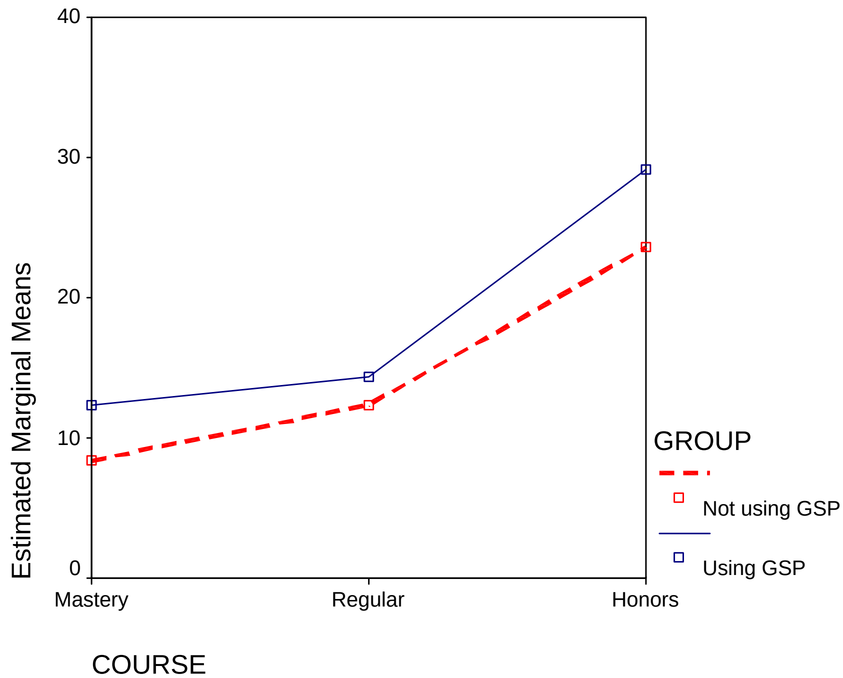


Figure 15: Estimated Marginal Means based on Courses: PROOF POST-TEST

These profiles confirmed the observations recorded above. Overall, students taught using GSP seemed to have performed better than those taught without the use of GSP. Honors students appeared to have benefited more than regular and mastery students from the use of GSP, while regular students seemed to have benefited the least. Moreover, mastery students in both GSP and non-GSP groups had performed better than their regular course counterparts in proof writing section, while they lagged behind when both proof construction and proof writing sections were considered. As there was a significant interaction effect between the variables COURSE and GROUP with the variable COURSE having three levels, a post hoc comparison of means of the three levels of COURSE was done to obtain a deeper insight into the performance of the students enrolled in different types of courses. Sheffe's post hoc comparison of means for both dependent variables PROOR WRITE and PROOF

POST-TEST indicated a significant difference in the means between (a) honors and mastery courses, and (b) honors and regular courses for both variables. All theses differences were significant at the level of .01. There was no significant difference in the means between regular and mastery courses in PROOF WRITE ($p > .05$) while the multiple comparisons showed a significant difference in means between regular and mastery in the variable PROOF POST-TEST at the significant level .05. The post hoc tests thus confirmed the observations derived from the two profile plots presented in Figure 14 and Figure 15.

Summary

Although initially 1,325 students responded to this study by taking the pre-tests, by the end of the academic year, this number had decreased substantially. Particularly, the number of students who completed the post-test II was the lowest (841). Due to this difference in the numbers of respondents for each test, more than a single analysis was considered necessary to investigate the research questions. The various analyses presented in this chapter led to the following observations:

Repeated measures analysis for students' performance on proof construction tests indicated a significant difference in the variance between the tests administered in the beginning and at the end of the academic year based on the type of course and an interaction between the types of course and use of GSP. There was no significant difference based on the use of GSP alone. The profile plots confirmed that GSP group students enrolled in honors course showed marked improvement in their performance by the end of the academic year beyond that of their regular and mastery counterparts. Mastery students in the non-GSP group showed a slight increase in mean score in the proof construction section of the post-test, while their counterparts in the GSP group recorded a slight decrease in mean. The post-hoc test confirmed the observations of the profile plots that honors students outperformed regular and mastery students in both GSP and non-GSP groups.

Correlation tests showed that pre- and post- tests of logical thinking had the maximum correlation and the pre-tests of logical thinking and proof construction were the least correlated. However, the post-tests of logical thinking and proof construction had a higher correlation than the pre-tests. Multiple regression analyses confirmed this by revealing highest predictability of the score for post-test of logical thinking based on the pre-test of logical thinking. Performance of students on the proof writing section of the post-test revealed some interesting patterns. Honors students were ahead of regular and mastery students in both the GSP and non-GSP groups. Nevertheless, GSP honors students appeared to have performed better than their non-GSP counterparts. Regular students had the lowest mean scores in the proof-writing tasks, individually as well as collectively. Finally, the overall end-of-year performance of GSP group appeared to be better than that of non-GSP group, and honors students topped the comparison of performance based on type of course, followed by mastery students, with regular students attaining the last place.

CHAPTER 5: FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Chapter Overview

Logical reasoning and proving have been long-standing goals of school mathematics coursework around the world. Particularly, success in proof writing has been an important goal of high school geometry courses. Research findings suggest that very few high school students achieve success in this endeavor (Senk, 1983; Stover, 1989). The study reported here was an effort to replicate and extend the existing research in this field by investigating the link between students' performance on tests of logical thinking and proof constructing, and the impact of using dynamic geometry technology in their performance. In this chapter, the findings resulting from the analyses reported in Chapter 4 are summarized in light of the research questions and the conceptual framework in which the study was grounded. The findings of the statistical analyses are supplemented with a detailed discussion about students' responses in the following section with typical examples from students' responses wherever applicable. Reflections about the merits and flaws of the design and methodology are also discussed. The final section of this chapter includes suggestions for future research on proof in secondary school mathematics.

Summary of Findings and Conclusions

Of the 1,325 students who participated in the study, 1,159 students completed the pre-test I, 1,124 students wrote the pre-test II, 1,031 wrote post-test 1, and 971 wrote the *proof construction* section of the post-test II. While 931 students wrote both pre- and post- tests of logical thinking, only 841 students answered the pre- and post- proof construction tests. Only 689 students responded to at least two of the four tasks in the *proof writing* section of the post-test II. In this section, the findings of various analyses completed to investigate the research questions and conclusions that could be derived are presented. The performance of students on the tests of logical thinking (LOGICAL THINKING 1 and LOGICAL THINKING 2) and the proof construction tests (PROOF CONSTRUCTION 1 and PROOF CONSTRUCTION 2) administered at the beginning and at the end of the academic year were analyzed using repeated measures with two independent variables--the type of course and use of GSP. As the different levels in the dependent and independent factors were unequal, closer scrutiny of the estimated marginal means was done to provide more insight into the issues investigated. Thus, profile plots of estimated marginal means of the pre- and post- levels of the dependent variables for the three types of course in each of the two groups were explored. The relationships among the performance of students on the two pre-tests and two post-tests were investigated using Pearson correlations and multiple regressions. Lastly, students' proof writing performance was investigated using the scores earned by them in post-test II.

Students' Performance on Proof Construction Tests

The four sub questions of research question 1 investigated the difference, if any, in the performance of students on proof construction tests given at the start and at the end of the high school geometry course. The results of the statistical analyses could be summarized with reference to the research questions as follows:

Research Question 1(a): Overall Performance

Is there a significant difference in the performance of high school geometry students between the *proof construction* tests administered at the beginning and at the end of a yearlong geometry coursework?

Repeated measures analysis results showed a significant difference in the variance of scores between the pre- and post- *proof construction* tests based on the main effect of the period of instruction, but just .7% of the variance could be accounted for by the effect of the period of instruction. The yearlong geometry course seemed to have had little effect on the performance of students on *proof construction* tests.

Research Question 1(b): Performance Based on Use of GSP

Is there a significant difference in the performance of high school geometry students between the *proof construction* tests administered at the beginning and at the end of a yearlong geometry coursework, based on the use of dynamic geometry software?

There was no significant difference in the variance of scores between the pre- and post- *proof construction* tests based on the use of GSP. Considering the fact that both dependent and independent factors in the analyses were unequal might have influenced the test results, analyses of estimated marginal means for the two groups was completed using profile plots. Although the profile plots indicated influence of use of GSP in the honors and mastery courses, there was no overall influence for all the types of courses.

Research Question 1(c): Performance Based on Type of Course

Is there a significant difference in the performance of high school geometry students between the *proof construction* tests administered at the beginning and at the end of a yearlong geometry coursework, based the type of geometry course?

Repeated measures ANOVA test indicated a significant difference in the variances between the scores of pre- and post- *proof construction* tests, based on the type of course. Nevertheless, only a 4% influence of the type of course on the performance of students was indicated. Estimated marginal means were observed using profile plots, in order to determine if there was a difference in the results due to the unequal number in each factor. The plots indicated a significant difference in performance among the three types of courses. The honors students had a higher performance level than the regular and mastery students in both groups and also showed a significant improvement in performance between pre- and post-tests. Regular students did not show any difference in performance between pre- and post- *proof construction* tests in both groups. Mastery students appeared to have been influenced more by the use of GSP than the course structure.

Research Question 1(d): Performance Based on Interaction Effect

Is there a significant difference in the performance of high school geometry students between the *proof construction* tests administered at the beginning and at the end of a yearlong geometry coursework, based on the interaction between the type of geometry coursework and the use of dynamic geometry software?

There was a significant difference in the variance of scores between the pre- and post-*proof construction* tests based on an interaction between the types of course and use of GSP in both analyses; only 2% of this difference was accounted for by the interaction effect. As the factors and the levels of the factors were unequal, estimated marginal means (unweighted means) and post hoc tests were considered to gain more insight into the investigation. Analysis of the profile plots was particularly important with regard to the impact of GSP use due to the fact that the repeated measures tests did not show a statistically significant influence of the use of GSP on the variance. Use of GSP seemed to have adversely affected the performance of mastery students and had little impact on regular students, while its use appears to have benefited the honors students.

The patterns discussed above suggest that the way the software is used may be an important influence on the impact of GSP in developing the reasoning skills in students. Effective use of GSP depends to a great extent on the attitudes of teachers and their efficiency and comfort level in using the software in instructional processes. Studies conducted by Zhonghong (2002), Pandiscio (2002), Roberts & Stephens (1999), and DeVilliers (1999) also support the idea that efficient use of GSP can help students develop the reasoning skills needed for success in proof construction. Analysis of descriptive statistics as well as the profile plots revealed a slight improvement in the performance of non-GSP group students in both honors and regular courses but a decline in the performance of mastery course non-GSP students. In the GSP group, honors students showed significant improvement while regular and mastery students recorded a slight decline in the unweighted means. Students enrolled in regular geometry coursework did not seem to have acquired the pre-requisite reasoning skills needed for success, not only in proof construction, but also in any argumentation involving demonstrative (deductive) reasoning. Closer scrutiny of the profile plots revealed that although the means of both groups declined to the same extent, the initial mean of the regular students in the GSP group was higher than that of non-GSP students. This observation suggests that one of the important factors influencing their performance could be the course structure. It is supported by the observation that both honors and mastery students performed better than the regular students in the non-GSP group.

Students' Performance on Logical Thinking Tests

The analyses of the tests of logical thinking ventured to examine the four sub questions of research question 2. The results of the statistical analyses could be summarized with reference to the research question as follows:

Research Question 2(a): Overall Performance

Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework?

The repeated measures ANOVA test for the tests of logical thinking indicated a significant difference in the variance of scores between the pre- and post- tests of logical thinking based on the period of instruction, with a substantial 39% of the difference in variance attributed to the effect of the period of instruction. Estimated marginal means for unequal factors also showed different results for the two tests. Observation of the profile plots showed a substantial increase in the estimated marginal means for all the three courses in both groups, from pre-test of logical thinking to post-test of logical thinking. It could be concluded that the yearlong geometry course did have some effect on the logical thinking abilities of students.

A close scrutiny of the percentage of correct responses for the five reasoning modes tested by the instrument exposed interesting insights. Each of the five reasoning modes was tested by two items in the instrument. Percentage of average number of correct responses in each pair is displayed in Table 16.

Table 16: Percentage of Correct Responses for Each Reasoning Modes

Reasoning Modes	Pre-test I	Post-test I
Proportional reasoning	37	52
Controlling Variables	55	31
Probabilistic reasoning	49	54
Correlational reasoning	42	55
Combinatorial reasoning	40	56

Table 16 shows that the overall performance of students on four of the five reasoning modes recorded an increase, which resulted in the overall percentage increase. Analyzing the responses of students to the two items related to controlling variables would shed more light into the common response patterns of these students and could possibly support the common response patterns identified by Tobin and Capie (1981). A deeper analysis of the tests of logical thinking was beyond the scope of this research. In any case, based on the fact that the percentage of total correct responses has shown an increase from pre-test to post-test (pre-test: 45%; post-test = 50%) and the results of the other statistical tests discussed above, the researcher tends to conclude that students' overall performance on the test of logical thinking has improved due to a yearlong geometry coursework.

Research Question 2(b): Performance Based on Use of GSP.

Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework, based on the use of dynamic geometry software?

There was no significant difference in the variance in scores between the pre- and post- tests logical thinking. Investigation of the estimated marginal means through profile plots confirmed this result. Use of GSP did not appear to have had a major overall impact in the difference in the performance of students between the pre- and post- tests.

Research Question 2(c): Performance Based on Type of Course.

Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework, based on the type of geometry course?

The repeated measures analysis indicated a significant difference, though only 2% of the variance was attributed to the influence of the type of course. Review of the estimated marginal means in the profile plots and the post hoc tests further confirmed the results suggested by the statistical tests. A general observation after reviewing the profile plots, which was also confirmed by the post hoc test results, was that honors students showed a higher logical thinking ability than did their regular and mastery counterparts in both GSP and non-GSP groups.

Research Question 2(d): Performance Based on Interaction Effect.

Is there a significant difference in the performance of high school geometry students between the tests of logical thinking administered at the beginning and at the end of a yearlong geometry coursework based on the interaction between the use of dynamic geometry software and the type of geometry coursework?

The statistical tests indicated that there was no significant difference in the variance of scores between the pre- and post- tests of logical thinking. Further examination of the profile plots of estimated marginal means did lead to some interesting observations. While the regular and mastery students in both GSP and non-GSP groups were found to have approximately the same level of logical thinking in the first test, the second test placed the performance of mastery students in the non-GSP group slightly higher than those of mastery GSP group, the regular GSP group, and the regular non-GSP groups.

Relationship between Students' Logical Thinking and Proof Performances

Research Question 3: Is there a relation between the measured changes in the performances on the tests of logical thinking and the proof tests administered at the beginning and at the end of the yearlong geometry coursework?

For the investigation of research question 3, the correlation between students' scores in the tests of logical thinking and *proof construction* tests were analyzed using multiple regression analyses. Pair-wise Pearson correlation showed significant positive correlation between each pair of these variables with the highest correlation between tests of logical thinking followed by the tests of proof construction. Each pair was at least 40% correlated. This would imply that if a geometry course were proof oriented, the course would help students develop higher levels of logical reasoning skills. Furthermore, the value of TOLT instruments for predicting students' performance on proof construction tests was established by the regression equations with positive coefficients ranging from .07 to .74.

High School Students' Performance on Proof Test

Research Question 4: How well do high school geometry students learn to write geometry proofs by the end of the high school geometry course?

Initially, the tasks 4-7 of the *proof writing* section of the post-test II was analyzed individually as well as collectively to explore research question 4. Further investigations were completed with the sum of students' scores in all tasks in post-test II. While individual explorations were conducted basically using descriptive statistics, students' overall performances on the proof writing tasks and in the entire post-test II were analyzed using a General Linear Model with a two-way ANOVA. The analysis of tests involving individual and collective scores and some conclusions that could be derived are discussed in the following sub sections.

Performance on Proof Writing Tasks

Overall, more than 50% of the respondents to each task from 4 to 7 scored below 40% of the total point worth of that task. Moreover, the number of respondents to each task decreased, with the increase in the level of difficulty (task 4: 765; task 5: 734; task 6: 528; task 7: 490). It appears that many students did not attempt to complete the more difficult tasks. This trend was also supported by the observations of McCrone and Martin (2002) that high school students feel uncomfortable with the open nature of logical deductions. Also, while scoring students' responses, the researcher felt that many students had proof phobia, i.e., they were put off by the very idea of having to write a proof. A detailed discussion of this issue is presented in the next section.

The analyses of variance completed for the total score of students in tasks 4-7 indicated statistically significant differences in the variance based on the type of course, the use of GSP, and an interaction between the two. Investigation of the estimated marginal means using profile plots indicated that in all the three courses, GSP students had performed better than their non-GSP counterparts in all the types of courses, and the difference in the performance of GSP and non-GSP groups in each course was highest in honors course.

Nevertheless, the mean scores of both groups fell below 50% of the total point worth of each task. Post hoc test on the independent factor COURSE indicated that honors students outperformed regular and mastery students in both GSP and non-GSP groups.

Performance on the Proof Construction Tasks

Most of the students who took the post-test II had responded to the first two tasks (task 1: 930 and task 2: 915). Task 3 required students to analyze arguments, establish the generalizability of each, and identify the most valid argument. This task had a lower response rate than the first two tasks (780). The overall performance of students on the *proof construction* section did not show significant improvement by the end of the academic year.

Overall performance on pre- and post- proof tests

In order to get a deeper insight into students' performances on the individual tasks in both pre- and post- proof tests, five levels of performance were developed based on the criteria suggested by McBride and Carifio (1995) and Senk (1982). A brief description of the performance levels suggested by these educators and the tables pertaining to the task-wise analyses of percentages of students placed in each performance level are presented in Appendix F. A close scrutiny of the tables indicated the following trends:

1. In task 1, students who were placed in performance level 0 and 1 in pre-test had shown improvement in post-test. This was evident from the fact that the percentage of students placed at performance levels 3 and 4 had increased in the post-test. Nevertheless, the percentage of students reaching level 4 declined substantially.

2. In task 2, some students at level 1 had progressed and some had regressed, as indicated by the increase in the percentage at levels 0, 2, and 3 and a decline in the percentage at level 1 in the post-test.
3. In task 3, students at levels 0 and 1 had improved, as indicated by the increase in percentage in levels 2, 3, and 4.
4. In task 4, 30% of the respondents had reached level 3, but only 15% could perform at level 4. Many students appeared to have confusion in forming conditional forms of the given statements.
5. More than half of the students who had attempted each of the tasks 5-7 were placed in level 0 or 1, i.e., below 40% of the total points for each of the tasks.
6. Percentage of students performing at level 4 showed a steady decline from 10% for task 5, to 6.6% for task 6 and 5.5% for task 7. Very few students had reached performance level 4 for any of the three proof construction tasks.
7. Many respondents found task 6 more difficult than task 7. This trend was interesting as task 6 was supposed to be relatively easier than task 7, due to the inclusion of a hint.
8. The analysis of performance levels of students in the pre- and post- proof construction tests confirmed the results of the other statistical tests, which showed no significant difference in the variance of scores between the pre- and post- proof construction tests.

In sum, the results of the analysis point to the fact that the high school geometry students who participated in this study had performed at a fairly low level on both *proof writing* and *proof construction* tests. This would mean that though more than half of the participating students could understand what was expected of them in responding to a given task, they were confused as to how to proceed, or made some major errors that invalidated their responses. Common errors and misconceptions are discussed in the following section.

Discussion and Recommendations

The results of the analysis, after probing into the four sub questions of research question 1, showed a very slight impact of the type of course and even less impact of the interaction of the use of GSP and the different types of courses, on the performance of students on *proof construction* tasks. The yearlong geometry instruction did not seem to have influenced the development of students' proof construction skills. The *proof construction* tasks were framed to investigate the performance of students with respect to (a) spatial reasoning, (b) conditional reasoning, (c) ordinal reasoning, (d) ability to use and analyze quantifiers, (e) ability to use counter-examples, (f) ability to analyze arguments to identify valid inference patterns and understand the inadequacy of examples to prove the truth of a statement, and (g) declarative knowledge (knowledge and understanding of definitions, postulates, among others). Some persistent confusions and errors observed by the graders are discussed below.

The most common error pattern observed in the responses was overt dependence on physical measurements to validate an argument. For example, task 1.1 presented two identical shapes overlapping a portion of each other. Students were required to determine if the non-overlapping portions had the same area. Typical erroneous responses to this item referred to lack of specific measurements or symmetry of the overlapped portion.

Nevertheless, some responses were strikingly justifiable. Examples of the best and the worst responses are presented below:

> *"No. Because the diagram may not be accurate because it could have more area on one side of the triangle and less on the other."*
>
> *"Yes. Because the middle shape is a diamond which makes the total of each side to be equal."*
>
> *"Yes. Because the overlapping areas are equal. So since the 2 tiles are identical then the rest is also equal so that non overlapping area does [do] have the same area."*

Most students view a mathematical proof as a method to examine and verify a particular case and tend to judge the validity of a proof by its appearance (Chazan, 1989; Martin & Harel, 1989). This trend was observed in students' responses to tasks requiring analysis of arguments. Chazan referred to this type of misconception as measurement of examples. In task 3.1, a set of three arguments to justify the statement "*Base angles of isosceles triangles are congruent*" were presented and the students were required to choose the best argument, in their opinion, which would prove the statement for <u>all</u> isosceles triangles. The first argument presented a single case by paper folding. The second was a proof that used an auxiliary construction of altitude to the base and was based on congruency of triangles. The third argument tested five different isosceles triangles to generalize the statement. Many students chose the third argument and justified their choice stating that the third argument tested with many examples, not just one. Examples of some typical responses are presented below:

"Argument B – because Shea did not measure the lengths of the sides or the angles but gave them variable[s] which means that for all isos. triangles, the base angles are congruent."

"Argument A because argument B was full of detail, but this only works on right triangles."

"I think all look like they are true because there [they're] all triangles."

"I believe that the first argument is correct due to the fact that Rick stated and concluded his argument real well."

"Argument C because Ling tested all isosceles triangles & not just one like Argument A & B did."

The ability to use counter examples to prove a statement appears to be present in most high school students. For instance, task 1.2 presented a diagram illustrating a cyclic quadrilateral in which the diagonals passed through the center of the circle and a statement announcing: "Whatever quadrilateral I draw with corners on a circle, the diagonals will always cross at the center of the circle." The students were asked to state whether they agreed with the statement and justify their choice. While many students agreed with the statement and justified their choice by drawing a confirmation from the diagram presented, many students responded negatively. Some of the justifications for this task include:

"Yes. He's right because every quadrilateral in the circle has their midpoint in the middle."

"Yes. Because the quadrilateral is touching the circle it will have the same center and diagonals touch the center of the quadrilateral."

"No. Kites are quadrilaterals, but their diagonals won't always cross at the center point" (diagram)

"No. Quadrilaterals w/90 deg. angles will have diagonals that go though the center but some such as a kite will not."

There was a fair amount of improvement in the ability of students in analyzing conditional statements as was evident from the most common responses to the two items testing this skill. Identifying, analyzing, and using quantifiers were also observed to have developed in more than 60% of the participants. Nevertheless, some good responses were invalidated by errors in class analysis, use of specific terms and phrases, and reluctance to articulate the justifications properly. For instance, task 2.1 presented the following item:

Statement: All squares are rectangles.

Based on the above statement, check (✔) the <u>one</u> statement that you think is true:

If a four-sided figure is not a square, then it is not a rectangle.________

If a four-sided figure is a rectangle, then it is a square.________

If a four-sided figure is a square, then it is a rectangle._______

Students were asked to explain the reason for their choice. Many students cited the given statement and simply said that their choice was the conditional form of that statement. This was a good response provided their choice was the correct one. Some students argued that as a rectangle has four right angles and a square also has four right angles, a rectangle should be a square, and justified their choice of statement (*ii*). While scoring, the graders encountered responses, which indicated inherent confusion about the class and classification of quadrilaterals in general. Ireland (1973) cites Robert Kane's (1960) study, wherein he found that both high school and college students tended to accept invalid reasoning as valid if it leads to a conclusion with which they agree and to reject valid reasoning as invalid if it leads to a conclusion with which they disagree. This tendency suggests that to some degree, high school and college students have not reached the formal operational level described by Piaget (Ireland, 1973).

The findings and observations discussed above prompt one to wonder about the actual emphasis placed by teachers on proving activities in high school geometry classes. The teachers' role in setting up an activity, asking appropriate questions and forcing children to think is essential to the successful teaching of proofs (Andrew, 1995). Teachers are also primarily responsible for developing positive attitudes toward proof in students. Apart from teachers' content knowledge and expertise in teaching methods, their attitudes towards the subject and the enthusiasm displayed by them in the class are important factors that help develop the much-needed positive attitudes in students towards proving activities. This fact was quite evident in many papers graded by the researcher, on which students wrote comments such as "I hate proofs," "Proof sucks," "my teacher did not teach this," and "This is true because I said so."

During the frequent visits to the participating schools, the researcher had opportunities to hold informal conversations with many participating teachers including all teachers using GSP. These conversations suggested that: (a) Of the six teachers who used GSP in their classes, one teacher taught mastery course (T2), two teachers (T5 and T7) taught honors course, two teachers (T6 and T8) taught regular courses, and one teacher (T18) taught both honors and regular courses; (b) The teachers T2, T6, T8, and T18 confirmed that they used GSP only for classroom demonstrations, quite infrequently, while teachers T5 and T7 said they organized individual and small group work with GSP for their students, besides using it in demonstrations in their geometry classes. These observations suggest that the way in which the software is used in the classroom is a major factor influencing the impact of the dynamic geometry technology on enhancing the performance investigated in this study. A replication of this research with a sample where groups could be controlled such that the GSP group students get appropriate exposure and experience with GSP activities would lead to a better understanding of the use of GSP on student performance on proof tests.

The results of analysis for investigating the four sub questions of research question 2 showed that about 35% of the improvement in students' performance on logical thinking tests could be attributed to the yearlong geometry coursework, and a further 2% could be due to the type of course. The findings for research question 3 indicated that high school geometry students' performance on general logical thinking tests and tests of proof construction skills were positively correlated, though to a small extent. This is a comforting observation in view of the researcher's theory and the perceptions of many educators that high school geometry is an important vehicle for developing logical thinking and problem solving skills in students (Fawcett, 1938; Ireland, 1973; Stover, 1989; Knuth, 1999; McCrone et. al., 2002,). It is interesting to note that only honors students recorded a significant improvement in their performance on the post-test of logical thinking. The analyses presented here indicate that proper use of GSP could develop students' logical thinking. It is also observed that incorporation of proving activities in the regular and mastery geometry courses would help students' general logical thinking abilities.

Students' engagement with proving activities that challenge and encourage them in making conjectures, testing and validating them, and justifying their conclusions to themselves and their classes, would help develop their general logical reasoning skills in addition to the reasoning skills needed to succeed in proof writing. Their ability to apply logical implications is very essential to success in proof construction (Hoyles & Kuchemann, 2002). As Prince (1998) rightly contends, students need the rigor of logical reasoning in order to face the innumerable complexities and challenges of the 21st century digital world.

The results of the investigation of the last research question, which was also the most important and significant to the current research, presented a somewhat dismal picture of the performance of students on the proof test administered towards the end of the year. First, of the 949 students who wrote the post-test, many students did not attempt the tasks related to proof writing (tasks 4-7). Second, of those who attempted tasks 4 and 5, which required only the basic concept of proof writing from the students, many could not or did not, attempt the proof-writing items in tasks 6 and 7. This observation is aligned with the findings of many studies that have established deductive reasoning skill among students who have studied high school geometry as being very low (Burger & Shaughnessy, 1986; Usiskin, 1982). Moreover, research results on students' conceptions of proof show that most high school and college students have not been appropriately exposed to the process of proving and justifying the mathematical processes (Dreyfus, 1999). Careful examination of the responses to the four tasks provided insight into some common response patterns and misconceptions of students.

The analyses presented here indicate that the percentage of students at performance levels 0 and 4 was much smaller than the percentage at levels 1-3. The performance levels in the proof-construction test at the beginning of the year and the corresponding part in the post-test do not reveal much difference in the levels of students. These observations suggest that probably most of the participating students were still at the van Hiele level 2 and hence could not develop formal reasoning skills needed for success in proof construction. Research studies investigating the cognitive processes needed for higher level geometrical thinking suggest that the formal reasoning skill needed for understanding, appreciating, and writing proofs require students to attain van Hiele level 3 (Battista & Clements, 1996; DeVilliers, 1999; Senk, 1989). This may be one of the reasons for students' observed difficulty with proof writing.

The response patterns observed by the graders supported the comments of Ireland (1973) that students tend to derive inferences from if-then statements in a consistent and erroneous way. He also notes that implications suggested by if-then statements should not be taken for granted with students of grade 10 or below. For instance, task 4.1 presented the statement "Diagonals of a parallelogram bisect each other" and required students to (a) rewrite the statement in if-then form, (b) draw and label a figure showing the given information, and (c) using the figure, write what is given and what is to be proved. Many students had written the if- then statement as "*If a parallelogram has diagonals, then they bisect each other*" or "*If the diagonals bisect each other, then it is a parallelogram*." Students' ability to draw appropriate diagrams was also found to be erroneous. For example, many students had drawn a rectangle in place of a parallelogram, or a segment joining the midpoints of the opposite sides of the quadrilateral as diagonals.

Task 6 required the students to know and use properties of isosceles triangles and properties of alternate interior angles when a transversal cuts parallel lines theorems whereas task 7 required them to know and use properties of parallelogram and triangle similarity theorems in addition to the alternate interior angles property. Apart from this difference, there was not much difference in the two tasks with regard to the abilities assessed by each. Also, use of the isosceles triangles theorem was provided as a hint in task 6; however, task 6 appeared to be more difficult to students than task 7, as suggested by the data in Table 22. Students who responded to task 7 were comfortable with identifying and stating vertical angles and similarity of AAA similarity property of triangles. Nevertheless, quite frequently responses like "*Triangle KLP similar to triangle NQP b/c they are similar in shape and that's it"* were observed.

These students were not able to recognize the two parallel lines from the parallelogram when one of the opposite sides were extended and, apply alternate interior angles theorem. Another common error pattern observed was a lack of ability to recognize and apply logical implications. This was frequently observed in the responses to task 5 and 6. It was very common to see the blank space next to the statement *"*$\Delta ABD \cong \Delta ACE$*"* filled in with the response "*CPCTC*" instead of "*ASA test of congruency*." As suggested by Ireland (1973), explicit instruction in valid and invalid inference patterns may help students overcome this barrier to success in proof writing.

Apart from the error patterns mentioned above, other error patterns that suggest serious lack of pre-requisite skills/knowledge for success in proof writing were: (a) declarative knowledge: Many students who attempted task 6 were not able to identify the parallel lines and the alternate interior angles in spite of the fact that it was mentioned in the statement of the task. Many students did not seem to understand the correct meaning of the tests of congruency and similarity, and wrote ASA as reason for similarity of two triangles in task 7; (b) tendency to inappropriately use short forms: the acronym CPCTC was used quite frequently at wrong places. (c) Linguistic skill: many of the terms and phrases frequently used in geometry were misspelled or misused. Commonly misspelled words include "vrticle" for "vertical," "than" for "then," "lenght" for "length," "logickule" for "logical." One of the important pre-requisites for success in proof writing is the ability to write the commonly used terms and phrases correctly (Stover, 1989). It is the researcher's view that students should be encouraged to write full sentences wherever applicable in any mathematical activity involving writing work. For that matter, teachers themselves should restrict use of abbreviations and shortened sentences in their instructional processes that involve writing.

Students exposed to GSP explorations performed generally better than non-GSP students in each of the tasks 4-7. This trend would suggest that students exposed to GSP may have developed other pre-requisite skills for success in proof-writing in addition to the reasoning skills investigated earlier. For example, task 4 tested the ability of students to recognize and state the hypothesis and claim parts of a geometric statement and to represent the given information visually by an appropriate sketch. One of the major skills tested in tasks 5-7 is the ability of students to recognize and apply various postulates and connect them with logical implications. GSP students seemed to have developed these skills better than their non-GSP counterparts.

Honors students seemed to have outperformed students enrolled in regular or mastery courses, in both GSP and non-GSP groups, in all the four tasks. This may be due to the fact that honors geometry courses are more proof-oriented than the other two courses. Also, honors students in the GSP group had significantly higher mean scores than non-GSP honors students in all tasks. In the regular course, GSP students had performed better than non-GSP students in tasks 4, 6, and 7. The mastery GSP students had performed better than non-GSP mastery students only in task 6. In view of the fact that tasks 5-7 basically tested students' declarative knowledge and ability to recognize and apply logical implications, these observations did not reveal any underlying trend or pattern. This would suggest that, apart from the type of course and the use of GSP, some other factors may have major influence on students' performance on the proof writing test.

Limitations

Planning for the research started about two months before the commencement of the academic year with an expectation that teachers could be contacted well in advance so that pre-tests could be administered about two weeks after the start of the academic year. But, due to the unforeseen weather conditions that prevailed in the region for more than a month, normal school work was delayed. Consequently, many of the selected teachers were reluctant to offer their class for the study. As a result, the sample for the study was one of convenience. Furthermore, the number of students in the different types of courses and different groups could not be controlled by the researcher and were unequal. Another important factor that might have had a major influence on the results of the current study were the variations in the types of participating schools. Apart from the ethnic diversity discussed in chapter 3, the school differences included the level of their performance on the state mandated standardized exam and the Scholastic Aptitude Tests.

The Florida Comprehensive Assessment Tests (FCAT) is administered by the Florida Department of Education each year to all students in Grades 3-10. Its purpose is to assess students' achievement in the higher cognitive skills recommended by the Sunshine State Standards and to provide a sound basis for comparison of Florida's student achievement with that of other states (Florida Department of Education, 2005). The standardized tests use criterion-referenced as well as norm-referenced tests. All students from grades 3-10 are assessed in reading and mathematics, while all students in grades 4, 8, 10 take writing tests, and students in grades 5, 8, and 10 take a science test. The scores that assess the Sunshine State Standards are reported in terms of five achievement levels, 1 being the lowest and 5 being the highest. The Department of Education website provides school wide data of FCAT mathematics mean scores and percentage of students in each level of achievement. As per the records shown in this website, about 63% of students enrolled in Sch1 and Sch2, 73% enrolled in Sch3, and only 57% enrolled in Sch4 during the academic year were placed at levels 3 and above. A difference in the performance of students on SAT exams (2003-2003 year) also could be observed from the Table 17.

Moreover, The FCAT 2005 results placed Sch1 and Sch2 in 'B Grade' level, while Sch3 was placed in 'A grade' level and Sch4 was placed in 'C grade' level. As is evident from Table 17, the difference in the test-taking ability levels of participating students might not have been uniform. This could have interfered with their performance on the proof and logical tests. Future research controlling the extraneous variables discussed above would probably yield a more generalizable result.

Table 17: Performance Levels of Participant Schools on FCAT and SAT (Percentages)

School code	FCAT 2004 Students levels			SAT 2002-2003 year	
	Level 1	Level 2	Level 3 and above	% taken	Mean score
Sch1	19	19	62	47	484
Sch2	16	21	63	62	493
Sch3	13	14	73	68	544
Sch4	13	31	56	42	443

Students' performance on two tests may not be sufficient to measure students' ability in proof writing and logical thinking. Future research with more structured tests supplemented with focus group interviews would provide further insight into the investigations of these abilities. The use of GSP for teaching proof related topics in geometry was entirely at the discretion of the participating teachers. The factors that might have interfered with the investigation of the impact of use of GSP might include (a) frequency of classroom demonstrations using GSP might vary with different teachers; (b) organization of the instructional process--some teachers used GSP for demonstration while some others were able to set up individual and / pair work with GSP for their students; and (c) the questioning method adopted by teachers might vary, which might influence the development of deductive reasoning skills in students. Future research with carefully structured activities for participating teachers followed by well planned socio-mathematical norms for argumentation in class would help achieve a more reliable investigation. Finally, future research in this field could include a nationwide sampling, with appropriate randomization so that the findings of the current research could be extended and generalized.

APPENDIX A: APPROVAL LETTER FROM IRB

Revised 6/04

THE UNIVERSITY OF CENTRAL FLORIDA
INSTITUTIONAL REVIEW BOARD (IRB)

IRB Committee Approval Form

PRINCIPAL INVESTIGATOR(S): Lalitha Subramanian **IRB #: 04-1983**

PROJECT TITLE: High Schools Geometry Students' Understanding of and Ability to Write Geometry Proofs: Effectiveness of Dynamic Geometry Software

Full Board
[ ] Contingent Approval
Dated: ____________

[ ] Final Approval
Dated: ____________

[ ] Expiration
____
Date: ____________

Committee Members:

Dr. Theodore Angelopoulos: ____________
Ms. Sandra Browdy: ____________
Dr. Jacqui Byers: ____________
Dr. Ratna Chakrabarti: ____________
Dr. Karen Dennis: ____________
Dr. Barbara Fritzsche: ____________
Dr. Robert Kennedy: ____________
Dr. Gene Lee: ____________
Ms. Gail McKinney: ____________
Dr. Debra Reinhart: ____________
Dr. Valerie Sims: ____________

Chair
[x] Expedited Approval
Dated: 28 July 2004

[ ] Exempt
Dated: ____________

[x] Expiration
Date: 27 July 2005

Chair, IRB

Signed: ____________
Dr. Sophia Dziegielewski

NOTES FROM IRB CHAIR (IF APPLICABLE): ____________

APPENDIX B: APPROVAL LETTER FROM LEA

Submit this form and a copy of your proposal to:
Program Services
P.O. Box 271
Orlando, FL 32802-0271

Orange County Public Schools

RESEARCH REQUEST FORM

Your research proposal should include: Project Title; Purpose and Research Problem; Instruments; Procedures and Proposed Data Analysis

Requester's Name: **Lalitha Subramanian** Date: **July 15, 2004**

Address Home: **2550, N. Alafaya Trail, Apt. # 6203, Orlando, FL 32826** Phone: **(407) 282 5516**

Business::**University of Central Florida, Orlando Campus, FL21817** Phone: **(407) 823n 1378**

Project Director or Advisor: **Dr. Michael Hynes** Phone: **(407) 823 2005**

Address: **College of Education, University of Central Florida, 4000, Central Florida Blvd., Orlando, FL 32817-1250**

Degree Sought (check one) ☐ Associate **X** Doctorate ☐ Bachelor's ☐ None ☐ Master's ☐ Specialist

Project Title: **High School Geometry Students' Understanding of and Ability to Do Proofs: Impact of Dynamic Geometry Software.**

ESTIMATED INVOLVEMENT

PERSONNEL/CENTERS	NUMBER **(Approx.)**	AMOUNT OF TIME (DAYS, HOURS, ETC.)	SPECIFY/DESCRIBE GRADES, SCHOOLS, SPECIAL NEEDS, ETC.
Students	**600**	**5 days (approx.)**	**High school geometry classes (grades 9/10)**
Teachers	**20**	**10 days (approx.)**	**High school geometry teachers**
Administrators			
Schools/Centers	**10**	**10 days (approx.)**	**Public high schools**
Others (specify)			

Specify possible benefits to students/school system: **A major benefit would be providing a deeper insight into the effectiveness of dynamic geometry software in enhancing high school geometry students' understanding of and ability to do geometric proofs. For more detail, please refer to the detailed methodology provided earlier.**

ASSURANCE

Using the proposed procedures and instrument, I hereby agree to conduct research in accordance with the policies of the Orange County Public Schools. Deviations from the approved procedures shall be cleared through the Senior Director of Program Services. Reports and materials shall be supplied as specified.

Requester's Signature: (Sd.) Lalitha Subramanian Sent via e-mail

Approval Granted: ☒ Yes ☐ No Date: 8-10-04

Signature of the Senior Director for Accountability, Research, and Assessment Lee Baker

9PL202 FORM ID #GB0103/23-1/1FY REV 11/00

APPENDIC C: PRELIMINARY SURVEY QUESTIONNAIRE

Survey Questionnaire

For questions 1 – 8, please indicate your agreement or disagreement with each of the following statements, with 1 being strong disagreement and 5 being strong agreement. Circle your responses.

Start Here

How far do you agree with………	Strongly Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Strongly Agree
1. Teaching formal proof in geometry develops logical reasoning ability in students.	1	2	3	4	5
2. I am conversant with different styles of proof Examples: flow-chart, paragraph, two-column, etc.).	1	2	3	4	5
3. Proof plays an important role in mathematical understanding.	1	2	3	4	5
4. The primary role of proof is to establish the truth of a statement.	1	2	3	4	5
5. Proof is a means of mathematical communication.	1	2	3	4	5
6. One aspect of proof is explanation.	1	2	3	4	5
7. Proof plays an essential role in creating knowledge and systematizing it.	1	2	3	4	5
8. I enjoy teaching proofs in my geometry class.	1	2	3	4	5

Please continue on the next page

Continue Here

For questions 9 to 17, indicate your responses by placing an **X** in the appropriate box. Please specify your response wherever needed.

9. In your opinion, at what grade level should students start learning to write proofs in mathematics?
   - ☐ K - 5
   - ☐ 6 - 8
   - ☐ 9 – 10
   - ☐ 11 – 12

10. How long have you been teaching mathematics at the high school level?
    - ☐ 1 – 3 years
    - ☐ 4 – 6 years
    - ☐ 7 – 9 years
    - ☐ More than 9 years

11. Please specify the high school mathematics courses you have taught. Place an '**X**' in the box beside **all** the courses you have taught.
    - ☐ Algebra I
    - ☐ Algebra II
    - ☐ Geometry
    - ☐ Pre-calculus
    - ☐ Trigonometry
    - ☐ Others (Please specify): ____

12. Please specify the high school mathematics courses you are currently teaching. Place an '**X**' in the box beside **all** the courses you are teaching.
    - ☐ Algebra I
    - ☐ Algebra II
    - ☐ Geometry
    - ☐ Pre-calculus
    - ☐ Trigonometry
    - ☐ Others (Please specify): ____

13. What is the highest level of mathematics you have studied?
    - ☐ Undergraduate
    - ☐ Masters
    - ☐ Doctorate
    - ☐ Others (Please specify): ____

Please continue on the next page

Continue Here

14. Are you using dynamic geometry software (Capri, Sketchpad, Logo, etc.) for teaching geometry?
    - ☐ Yes, CONTINUE WITH QUESTION 17
    - ☐ No, SKIP TO QUESTION 20

15. Which dynamic geometry software are you using in your geometry class?
    - ☐ Geometer's Sketchpad
    - ☐ Capri
    - ☐ Logo
    - ☐ Other (Specify): ____

16. How do you rate the usefulness of the software for teaching geometry?
    - ☐ Very useful
    - ☐ Useful to some extent
    - ☐ Rarely useful
    - ☐ Not at all useful

17. How do you rate the usefulness of the software for teaching proofs in geometry?
    - ☐ Very useful
    - ☐ Useful to some extent
    - ☐ Rarely useful
    - ☐ Not at all useful

For questions 18 - 22, please write your responses in brief.

18. If you are not using any dynamic geometry software in your class, please explain briefly the reason, if any, for your decision.

____

____

____

____

Please continue on the next page

Continue Here

19. In your experience, what are the **most challenging** aspects about teaching proof in geometry?

____

____

____

____

20. In your experience, what are the **easiest** aspects about teaching proof in geometry?

____

____

____

____

21. Which textbook are you using for your geometry class this year? How do you like this book?

____

____

____

____

22. For any other comments, please use the following lines.

____

____

____

____

Thank you for responding to this survey. I really appreciate your valuable input.

APPENDIX D: PROOF CONSTRUCTION PRE-TEST

Entry-level proof Construction Survey

Name: ____ Class: __________________

School: ____ Gender: M or F: __________

Geometry Teacher's name: ____

<table>
<tr>
<td rowspan="2">
You have 30 minutes to answer these questions.

In this packet, there are five tasks. Some tasks ask your opinion about statements or diagrams. Some tasks present arguments about a statement and ask you questions about the arguments. We are interested in your thinking as well as your answers. So, please show all your rough working on the same page as your answer.

In most questions, you will be asked for explanations. Make these as clear as you can, but don't make them longer than necessary.

Use a pen. You may cross thing out, but do not rub out any of your work and do not use correction fluid.

You might find some of the questions quite difficult. Don't worry. If you get stuck on a question, leave it and proceed with the next. You can return to it later.

This test is given to you as part of a research project. The results of this test will not be used in your school grades. This research is about changes in mathematics education that are made possible by the use of computers. Please take this test seriously and do your best.
</td>
<td># A

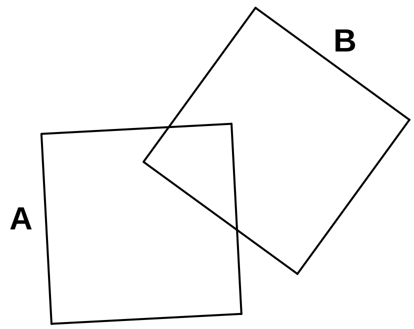
For office use only</td>
</tr>
<tr>
<td>School: ☐

Class: ☐

Student: ☐</td>
</tr>
</table>

Task # 1: In this task, two questions based on diagrams are given. Answer the questions by writing ‘Yes’ or ‘No’ in the blanks provided after the questions. Please explain your answers with appropriate justifications.

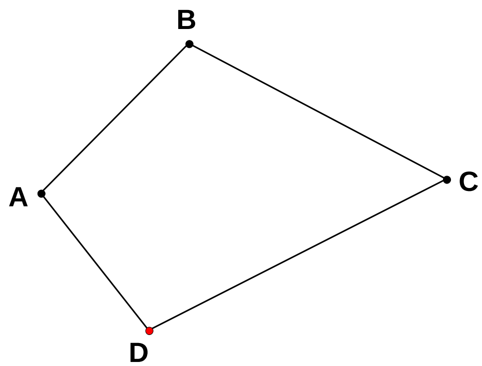
1. The diagram shows two identical square tiles A and B. The tiles overlap.

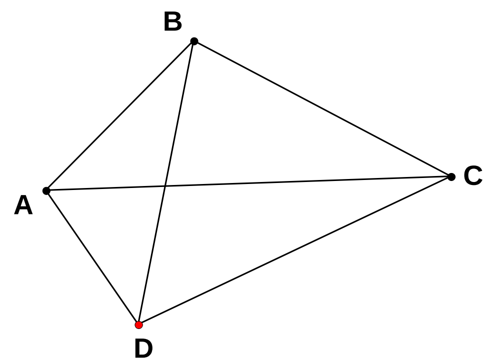


a) Do the non-overlapping regions have the same area? ____________________

Explain your answer:

2. Tim sketches a quadrilateral. He draws diagonals to the quadrilateral.





Tim notices that one of the diagonals has cut the area of the quadrilateral in half. He says: "Whatever quadrilateral I draw, one of the diagonals will cut the area of the quadrilateral in half.

Is Tim right? ___________________________

Explain your answer:

Task # 2:
Directions: In this task, two sets of statements are given. Each set consists of mathematical facts, followed by sub questions. Please read the statements carefully and answer the sub questions.

1. Statement: **All rectangles are parallelograms.**

Based on the above statement, check (✔) the <u>one</u> statement that you think is true:

i) *If a four-sided figure is not a rectangle, then it is not a parallelogram.*

ii) If a four-sided figure is a parallelogram, then it is a rectangle.

iii) If a four-sided figure is a rectangle, then it is a parallelogram.

Explain your choice:

Reason:

2. Statements: **a) Diagonals in rectangles always have the same length.**
b) Some rhombuses are rectangles.

Based on the statements above, check (✔) the <u>one</u> statement that you think is true:

i) Diagonals in rhombuses never have the same length. _____

ii) Diagonals in rhombuses sometimes have the same length. _____

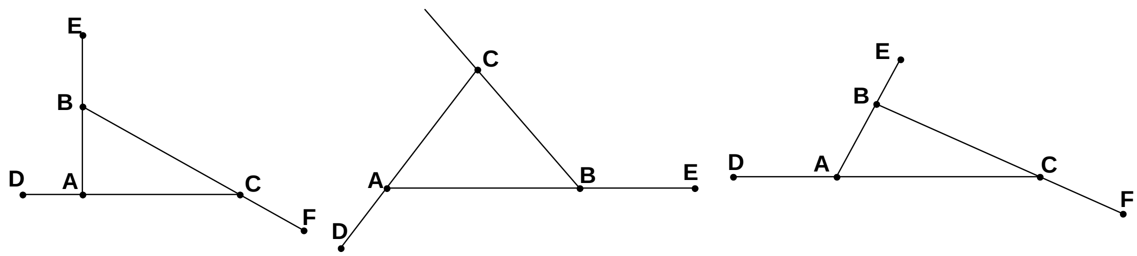
iii) Diagonals in rhombuses always have the same length. _____

Explain your choice.

Task # 3:
Directions: In this task, you are given two statements, which may or may not be true, followed by arguments that attempt to justify each statement. Please read the statements and arguments and respond to the numbered sub questions. Be as specific as possible in your responses.

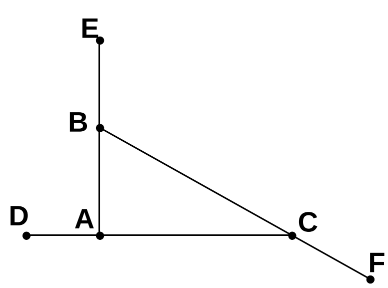
1. Statement: **In a triangle, the sum of the exterior angles is 360°.**

Argument 1:



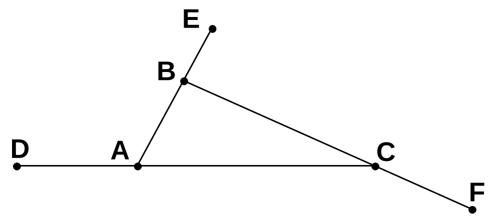
I drew three different triangles. I labeled each triangle ABC. Each triangle also has three exterior angles labeled ∠BAD, ∠ ACF, and ∠ CBE. I measured the exterior angles in each of the three triangles and in each case the sum of the exterior angles was 360°. Since I checked all three kinds of triangles, namely, right, acute, and obtuse, I can be sure that Statement 1 is always true.

Argument 2:



If you start at vertex A facing north (UP), then turn clockwise and walk along the sides of the triangle until you end up at vertex A again, you will end up facing the way you began. Because each of the three turns you make on your trip is the same number of degrees as each of the three exterior angles on the triangle, the number of degrees your body turns is the same as the exterior angle sum. As a result, the full 360 degrees turn (in which you end up facing the same direction you started) you completed shows that the sum of the exterior angles also must be 360°.

Argument 3:



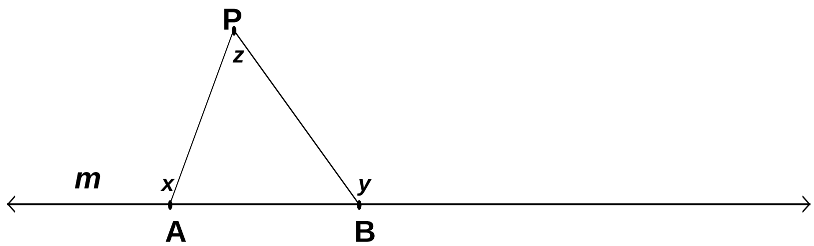
In a triangle, each exterior angle forms a linear pair with its associated interior angle. You can see that the sum of the angle measures in each linear pair is 180°, we can subtract 180 from 540 (sum of all interior and exterior angles) to be left with the sum of the triangle's exterior angles only. We know that 540 – 180 = 360, so the statement is true for all triangles.

Which of the above arguments show that the statement is true for all triangles? Please justify your answer.

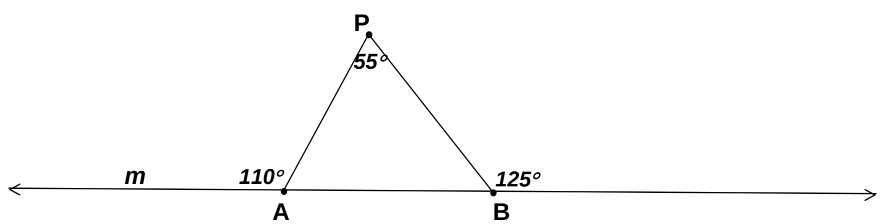
Your answer:

2. In the diagram, A and B are two fixed points on a straight-line *m*. Point P can move, but stays connected to A and B (the straight lines PA and PB can stretch or shrink).

Statement: **$x°$ + $y°$ is equal to 180° + $z°$**

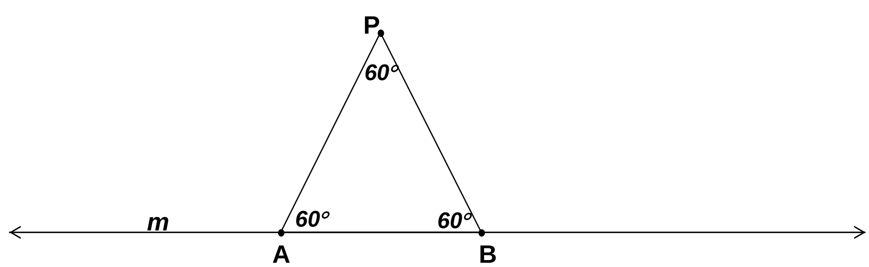


<u>Argument 1:</u>



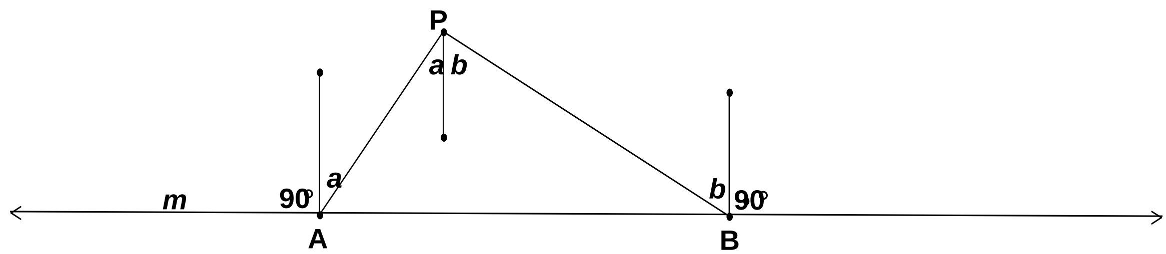
I measured the angles in the diagram and found that angle x is 110°, angle y is 125°, and angle z is 55°. 110° + 125° = 235°, and 180° + 55° = 235°
So, the statement is true.

<u>Argument 2:</u>



I can move P so That the triangle is equilateral, and its angles are 60°. So, x is 120°, and y is 120°. 120° + 120° is the same as 180° + 60°. So, the statement is true.

Argument 3:



I drew three parallel line segments. The two angles marked as '*a'* are the same and the two marked as '*b*' are the same. Angle x is 90° +a, and angle y is 90° + b. So, x plus y is 180° + a + b, which is 180° + z.
So, the statement is true.

Which of the above arguments show that the statement is true for all positions of the point P. Please justify your answer.

Your answer:

Thank you for completing this test. Have a good day.

APPENDIX E: PROOF POST-TEST

Geometry Proof Questionnaire

Name: ____Period #: _________________

Class: Regular____________ Honors: ____________Other (specify): ______________

School: ____ Gender: M or F: __________

GeometryTeacher'name____

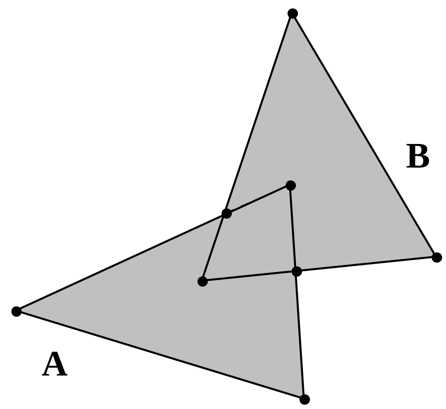
You have 50 minutes to answer these questions. In this packet, there are eight tasks. Some tasks ask your opinion about statements or diagrams. Some tasks present arguments about a statement and ask you questions about the arguments. We are interested in your thinking as well as your answers. So, please show all your rough working on the same page as your answer. In most questions, you will be asked for explanations. Make these as clear as you can, but don't make them longer than necessary. Use a pen. You may cross things out, but do not erase any of your work and do not use correction fluid. If you get stuck on a question, leave it and proceed with the next. You can return to it later. This test is given to you as part of a research project. The results of this test will not be used in your school grades. This research is about changes in the teaching and learning of geometry that are made possible by the use of computers. Please take this test seriously and do your best.	**A** For office use only School: ☐ Class: ☐ Course: ☐ Student: ☐

Section 1: Proof Construction

Task # 1: In this task, two questions based on diagrams are given. Answer the questions by writing ‘Yes’ or ‘No’ in the blanks provided after the questions. Please explain your answers with appropriate justifications.

1.1: The two triangles A and B have the same area. The triangles overlap.

Do the two non-overlapping regions have the same area? _______



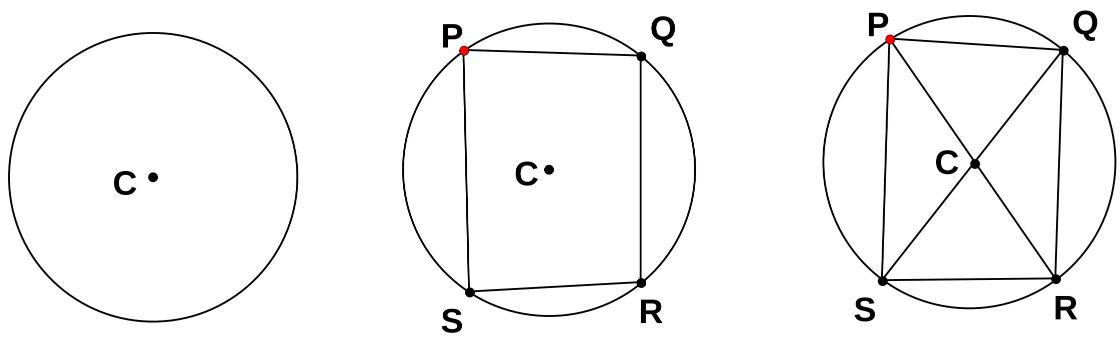
Explain your answer:

1.2:

Darren sketches a circle. He calls the center C.

He then draws a quadrilateral PQRS, whose corners lie on the circle

He then draws the diagonals of the quadrilateral.



Darren says: “Whatever quadrilateral I draw with corners on a circle, the diagonals will always cross at the center of the circle.”

Is Darren right? ___________________________

Explain your answer:

Task # 2: Directions: In this task, two sets of statements are given. Each set consists of mathematical facts, followed by three statements. Please read the statements carefully and answer the questions.

2.1: Statement: **All squares are rectangles.**

Based on the above statement, check (✔) the <u>one</u> statement that you think is true:

(i) *If a four-sided figure is not a square, then it is not a rectangle.*________

(ii) *If a four-sided figure is a rectangle, then it is a square.*________

(iii) *If a four-sided figure is a square, then it is a rectangle.*_______

Explain the reason for your choice:

2.2: Here are two statements:

a. **Diagonals of a rectangle always have the same length**.
b. **Some parallelograms are rectangles.**

Based on the above statements, check (✔) the <u>one</u> statement that you think is correct:

i) *Diagonals of parallelograms never have the same length.*_______

ii) *Diagonals of parallelograms sometimes have the same length.*______

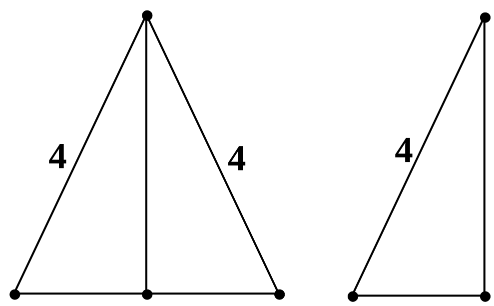
iii) *Diagonals of parallelograms always have the same length.*______

Explain the reason for choosing your answer:

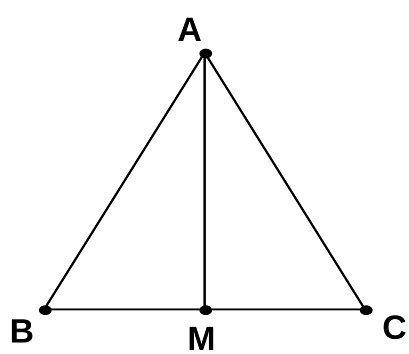
Task # 3: In this task, you are given two statements, which may or may not be true, followed by arguments that attempt to justify each statement. Please read the statements and arguments carefully and respond to the sub questions. Be as specific as possible in your responses.

3.1: **Statement:** **The base angles of an isosceles triangle are congruent**

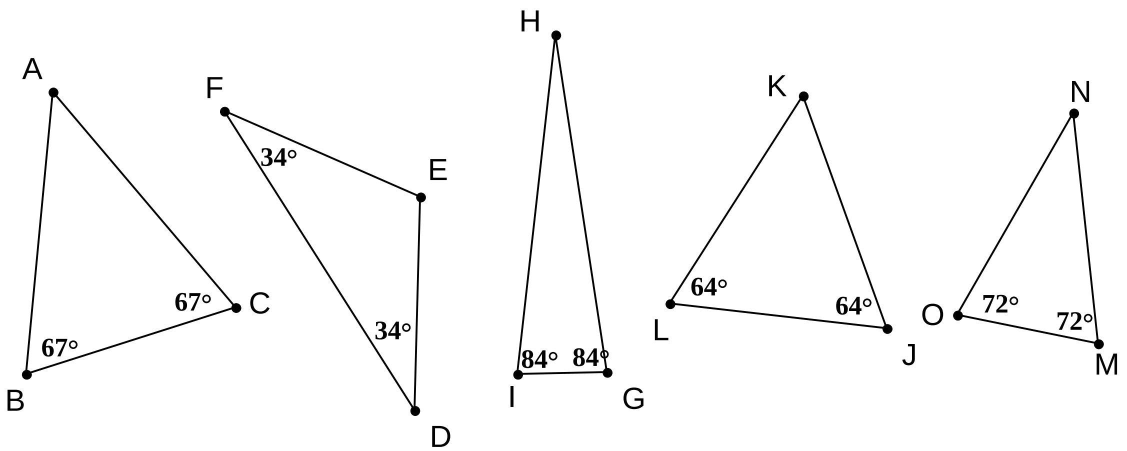
Argument A: Rick drew an isosceles triangle on a sheet of paper. He then cut out the triangle, and folded it over the median that connects the midpoint of the base to the vertex angle. He observed that the triangle is divided into two congruent parts. He concluded that base angles of all isosceles triangles are congruent.



Argument B: In order to prove that the base angles of any isosceles triangle are congruent, Shea drew an isosceles triangle Δ ABC, with AB = AC. She also sketched an extra line segment AM, such that AM is perpendicular to BC. Shea said that the two triangles ABM and ACM are right triangles, with congruent hypotenuses and a common side AM. Then she proved that Δ AMB is congruent to Δ AMC, and concluded that ∠B is congruent to ∠C.

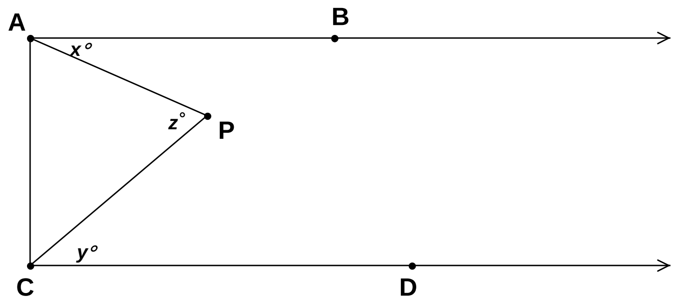


Argument C: Ling drew five different isosceles triangles, and measured the base angles of each. He found that in each case, the base angles had the same measure. So, he concluded that the statement is true for all isosceles triangles.



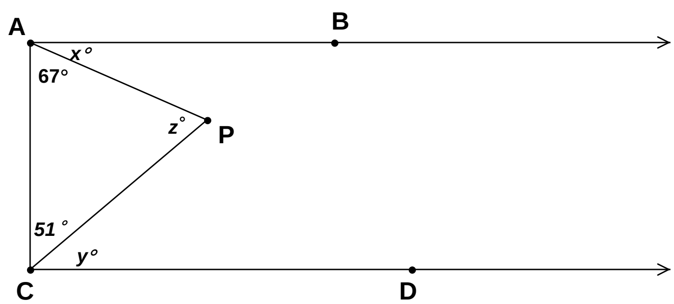
In your opinion, which of the above arguments is true for all isosceles triangles. Please justify your answer.

3.2: In the diagram, ray AB is parallel to ray CD, and AC is at right angles to both rays. Points A, B, C, and D are fixed. Point P can move anywhere between AB and CD, but stays connected to A and C. (The straight line segments PA and PC can stretch or shrink).

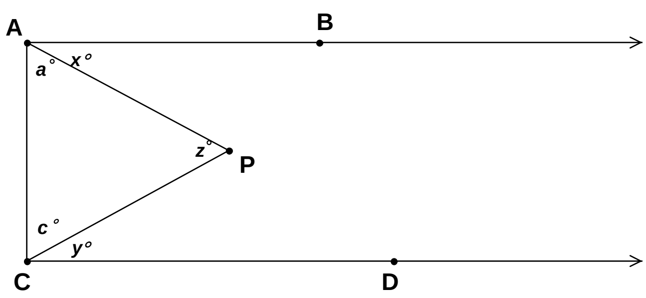


Statement: $x° + y° = z°$

<u>Argument A:</u> Rick said: I could have a triangle APC with the angles as shown in the diagram. Then, $z = 180 – 51 – 67 = 62$. I also have $x = 90 – 67 = 23$, and $y = 90 – 51 = 39$. But, $39 + 23 = 62$. Thus, I get $x + y = z$. So, I conclude that the statement is always true.



<u>Argument B:</u> Shea said: The angle sum of a triangle is 180°. So, $z + a + c = 180°$. Angles BAC and DCA are 90°. So, I can write $90 – x$ for a, and $90 – y$ for c. Then, I can write the first equation as $z + (90 – x) + (90 – y) = 180$. This gives me $z - x - y + 180 = 180$. So, $z – x – y = 0$. Thus, we get $z = x + y$ is always true.



Argument C: Ling said: I measured the angles in the original diagram. I then moved P to another place and measured the angles again. I made the following table:
Both times I found that $x + y = z$. So, the statement is always true.

x	y	z
23	40	63
17	32	49

In your opinion, which of the above arguments is the best argument. Please justify your answer.

Section 2: Proof Writing

Task # 4.1:

Read the following statements carefully. Suppose you wished to prove each of these statements. (i) Rewrite each statement in conditional form, i.e., if- then form.

(ii) Draw and label a figure showing the given information.

(iii) Using your figure, write what is given and what is to be proved.

a) Statement: **Diagonals of a parallelogram bisect each other.**

(i) Conditional form:

____

____

(ii) Figure:

(iii) Given:___________________________

To Prove:_________________________

b) Statement: **A segment joining the midpoints of two sides of a triangle is parallel to the third side.**

(i) Conditional form:

____

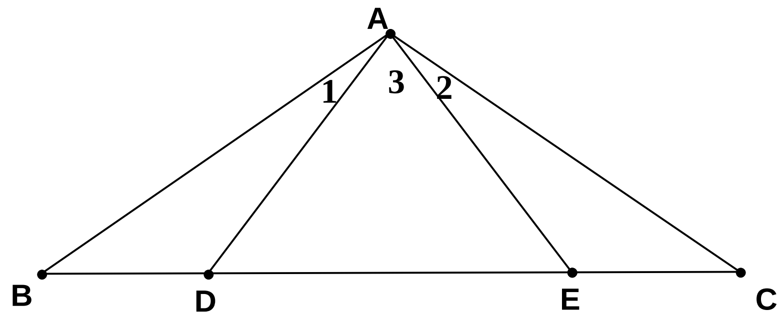
____

(ii) Figure:

(iii) Given: __________________________

To Prove: _______________________

Task # 5:
Fill in the blanks to complete the proof:

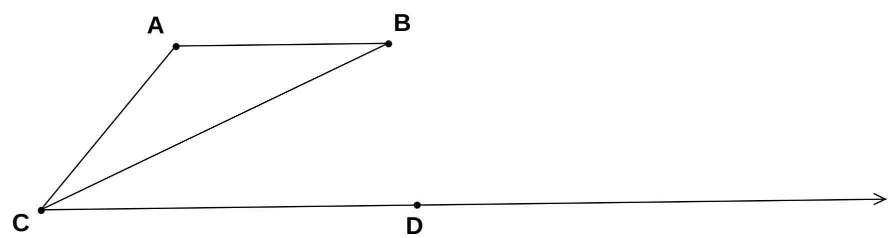


Given: Δ ABC is an isosceles triangle with AB ≅ AC.
$\angle 1 \cong \angle 2$

Prove: Δ ADE is isosceles.

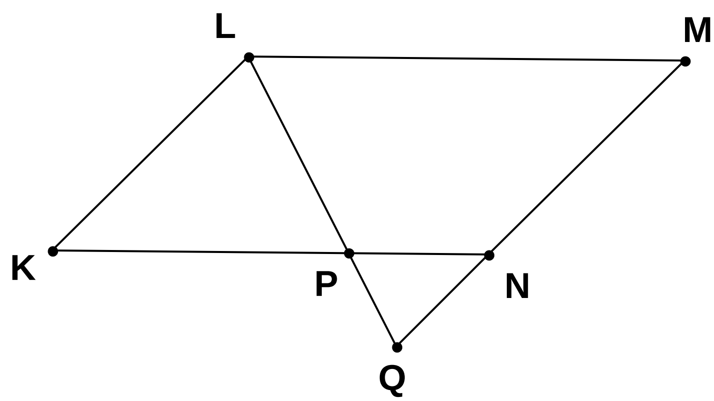
Statements	Reasons
1. AB ≅ AC	______________________________
2. ______________________________	Base angles of an isosceles triangle are congruent (equal in measure)
3. $\angle 1 \cong \angle 2$	Given
4. Δ ABD ≅ Δ ACE	______________________________
5. ______________________________	Corresponding parts of congruent triangles are congruent.
6. Δ ADE is isosceles	______________________________

Task # 6:
In the following figure, AB is parallel to CD. AB = AC. Using properties of isosceles triangles, prove that CB bisects $\angle$ACD. State appropriate reasons for each statement you write for the proof.



Write your proof here:

Task # 7: Write a Statement and Reason proof:
Given: KLMN is a parallelogram. N is on line MQ, and LQ and KN intersect at P
Prove: Δ KLP ~ Δ NQP.



Write your proof here:

Thank you for completing this test. Have a good day.

APPENDIX F: PERFORMANCE LEVELS

In order to aid the analysis of students' performances on the individual tasks, five levels of performance were developed based on the criteria suggested by McBride and Carifio (1995) and Senk (1982). A brief description of the performance levels suggested by these educators and the tables pertaining to the task-wise analyses of percentages of students placed in each performance level are presented in Table E 1.

Based on the categorization suggested by McBride and Carifio (1995) and Senk (1982), the following five levels were developed for investigating students' performance on proof writing tasks:

Level 0: No work or meaningless work done.

Level 1: Understanding of what is expected was evident, but confusion on how to proceed.

Level 2: Understanding of correct procedure, but major errors invalidated the proof.

Level 3: Complete proof, but flawed by minor errors.

Level 4: Complete and valid proof.

Categorization of students in the various performance levels were determined using the coding scheme presented in Chapter 3 as follows:

Level 0: A score of 0 in each task.

Level 1: Scores between 1 and 3 for task 4; either 1 or 2 for tasks 5-7.

Level 2: Scores between 4 and 6 for task 4; 3 for tasks 5-7.

Level 3: Scores between 7 and 9 for task 4; 4 for tasks 5-7.

Level 4: Scores between 10 and 12 for task 4; 5 for tasks 5-7.

Table E 1: Description of the Categorization of Proof Construction Levels

Level	McBride & Carifio (1995, p.7)	Senk (1982, p.7)
0	Unschooled: those who have no idea what to do.	Noncommencement: No work or only meaningless work was done.
1	Novices: Those who understand what they are expected to attempt but are confused.	Approach: Some meaningful work was done, but an early impasse was reached.
2	Intermediate: Those who show an understanding of the process but appear to be missing knowledge of a key concept that would enable the necessary deductions.	Substance: Sufficient detail indicated that the student proceeded toward a rational solution, but major errors invalidate the proof.
3	Competents: Those who appear able to solve most problems or write most proofs.	Results: Minor errors flaw an otherwise valid proof
4.	Experts: Those who can write eloquent proofs or problem solutions.	Completion: a complete valid proof was produced.

Students' performances on individual tasks were analyzed in light of the above levels. Table E 2 presents the percentages of students placed at each performance level for the preand post- proof construction tests. Table E 3 illustrates the number and percentage of students at each performance level in each of the tasks 4-7.

Table E 2: Performance Level: Pre- and Post- Proof Construction Tasks

Level	Task 1		Task 2		Task 3	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
0	39.9	38.4	19.5	35.9	51.3	31.2
1	27.7	19.4	34.7	23.7	40.4	31.0
2	18.1	32.5	35.6	36.8	7.3	20.3
3	15.8	17.0	9.5	21.5	7.3	21.6
4	4.8	1.3	1.7	0.2	0.3	6.0

Table E 3: Performance Levels: Proof Writing Tasks

Level	Task 4		Task 5		Task 6		Task 7	
	N	Percent	N	Percent	N	Percent	N	Percent
0	35	4.6	68	9.3	133	25.2	98	20
1	164	21.4	338	46	292	55.3	260	53.1
2	215	28.1	167	22.8	30	5.7	60	12.2
3	234	30.6	87	11.8	38	7.2	45	9.2
4	117	15.3	74	10.1	35	6.6	27	5.5

LIST OF REFERENCES

Allen, R., Channac, S., & Trillling, L. (2001). Parallels between constructing dynamic figures and constructing computer programs. *The Journal of Computers in Mathematics and Science Teaching, 20*(2), 179-97.

Andrew, P. (1995). Proof in secondary mathematics: The necessary place of the concrete. *Mathematics in School, 24,* 40-42.

Battista M. (2002, February). Learning geometry in a dynamic computer environment. *Teaching Children Mathematics, 8*(6), 333 - 339.

Battista, M. T., & Clements, D. H. (1996, May). Geometry and proof. *Mathematics Teacher, 89*(5), 386-388.

Becker, L. A. (1999, November). *GLM: Unequal n designs.* Retrieved June 4, 2005, from http://web.uccs.edu/lbecker/SPSS/glm_uneqn.htm

BouJaoude, S. (2004, January). Relationships between selective cognitive variables and students' ability to solve chemistry problems. *International Journal of Science Education, 26* (1), 63-84.

Burger, W. F., & Shaughnessy, J. M. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*(1), 31-48.

Callahan, P., & Stanley, D. (n. d.). Content statement: Secondary geometry. Retrieved November 25, 2004, from http://tepd.ucop.edu/ois/ucop_admin/getfile.php?file=CS_Sec_Geom.pdf&infoID=75&fileType=application%2Fpdf

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin Company.

Capie, W., & Newton, R. (1981, May 30). *Developmental patterns among formal reasoning skills.* Paper presented at the Eleventh Annual Symposium of the Jean Piaget Society, Philadelphia.

Chazan, D. (1989). *Ways of knowing: High school students' conception of mathematical proof.* (Doctoral dissertation, Harvard University, 1989).

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24* (7), 359 – 87.

Clements, D. H., & Meredith, J. S. (1993). Research on Logo: Effects and efficacy. *Journal of Computers in Child Education,4*, 263 – 290.

Coxford A. (1991). Geometry from multiple perspectives. In Hircsh, C.R., (Ed.), *Curriculum and Evaluation Addenda series*. Reston, VA: National Council of Teachers of Mathematics.

Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage Publications Inc.

DeVilliers, M. (1990, November). The role and function of proof in mathematics. *Pythogoras, 24*, 17-24.

DeVilliers, M. (1986) The role of acclimatization in mathematics and mathematics teaching. *Research Unit for Mathematics Education University of Stellanbosch (RUMEUS)*, South Africa.

DeVilliers, M. ((1999). The role and function of proof with Sketchpad. *Excerpt from Introduction to De Villiers M. (1999) Rethinking Proof with Sketchpad*. Emeryville, CA: Key Curriculum Press.

DeVilliers, M. (2002). Developing understanding for different roles of proof in dynamic geometry. *Paper presented at ProfMat,* Visue, Portugal, October 2-4, 2002.

Dennis, A. (2000, November). A survey of mathematics undergraduates' interaction with proof: some implications for mathematics education. *International Journal of Mathematical Education in Science & Technology, 31*(6), p.869-890.

Dennis, A. (2001). Pupils' proof potential, *International Journal of Mathematical Education in Science & Technology, 0020-739X*, 32 (1).

Dennis, A. (2003) Engendering proof attitudes: Can the genesis of mathematical knowledge teach us anything? *International Journal of Mathematical Education in Science and Technology, 34*(4), 479-488.

Dennis, C. C. (1977, January). The relative effectiveness of three geometric proof construction strategies. *Journal for Research in Mathematics education,* 62-67.

Dessart D. J. (1981). *Curriculum: Mathematics education research: Implications for the 80's.* Reston, VA: National Council of Teachers of Mathematics.

DeVault, M. V., & Weaver, J. F. (1970). A history of mathematics education in the United States and Canada. *National Council of Teachers of Mathematics, 32nd year book*. Reston, VA: National Council of Teachers of Mathematics.

Dreyfus, T. (1999). Why Johnny can't prove (with apologies to Morris Cline). *Educational Studies in Mathematics, 38*, 85-109.

Dye, B. (2001). The impact of dynamic geometry software on learning. *Teaching mathematics and its applications, 20*(4), 157-169.

Edutopia (n.d.). *National Standards for Technology in Teacher Preparation*. Retrieved December 20, 2004 from http://www.edutopia.org/php/resources.php?id=item_216859

Fawcett, H. (1938). Nature of Proof. *National Council of Teachers of Mathematics 13th yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Feller, B. (2004, November 16). In second term, Bush seeks to build on his education base. The Detroit News: Politics and Government. Retrieved December 7, 2004 from http://www.detnews.com/2004/politics/0411/16/-6185.htm

Field, A. (n.d.). A bluffer's guide to … sphericity. *Retrieved June 3, 2005, from* http://www.sussex.ac.uk/Users/andy/research/articles/sphericity.pdf

Florida Department of Education Website. (n.d.). *Evaluation and reporting*. Retrieved February 2, 2005, from: http://www.firn.edu/doe/evaluation/home0018.htm

Galindo, E. (1998, January). Assessing justification and proof in geometry classes taught using dynamic software. *Mathematics Teacher, 91*(1), 76-82.

Hanna, G. (1996). The ongoing value of proof. *PME XX (1).* Valencia, Spain.

Hannafin R. D., Burruss, J. D., & Little C. (2001). Learning with dynamic geometry programs: Perspectives of teachers and learners. *The Journal of Educational Research, 94*(3), 133-144.

Heinze, A., & Kwak, J. Y. (2002). Informal prerequisites for informal proofs*. ZDM, 34*(1), 9-16.

Herbst, P. G. (2002). Establishing a custom of proving in American school geometry: evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics, 49*, 283-312.

Herbst, P. G. (2002, May). Engaging students in proving: A double bind on the teacher. *Journal of Research in Mathematics Education, 33*(3), 176-203.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24,* 389-399.

Hirsch, C. R. (Ed.). (1985, March). Deductive and analytical thinking: activities. *Mathematics Teacher, 78,* 188-194.

Hoffer, A. ((1981, January). Geometry is more than proof. *Mathematics Teacher,* 11-18.

Honavar, V. (2005, January 5). High Schools Must Demand More: Achieve urges states to prepare for college, work: Report. *Edweek, 24*(16), 3. Retrieved May 30, 2005 from http://www.edweek.org/ew/articles/2005/01/05/16achieve.h24.html?querystring=Honawar

Hoyles, C., & Healy, L. (2004). Justifying and proving in school mathematics. *Project funded by ESRC (1/11/95 – 31/10/98).* Retrieved November 20, 2004 from http://www.ioe.ac.uk/proof/proof.htm

Hoyles, C., & Kuchemann, D. (2002). Students' understandings of logical implication. *Educational Studies in Mathematics, 51*, 193-223.

Ireland, S. H. (1973). *The effects of a one-semester geometry course which emphasizes the nature of proof of student comprehension of deductive processes.* (Doctoral dissertation, University of Michigan, 1973).

Iyvette, E. (1996, January). Understanding that ANOVA effects are perfectly uncorrelated. *Paper presented at the annual meeting of Southwest Educational Research Association*: New Orleans, LA.

Izen, S. P. (1998, November). Proof in modern geometry. *Mathematics Teacher, 91*(8), 718-720.

Knuth, E. J. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal of Research in Mathematics Education,* 33(5), 379-405.

Knuth, E. J., & Elliott, R.L. (1998). Characterizing students' understanding of mathematical proof, *Mathematics Teacher, 91*(8), 714-731.

Kunth, E.J. (1999). Secondary school mathematics teachers' conceptions of proof. (Doctoral dissertation, University of Wisconsin, Madison, 1999).

Lindquist, M. M., & Clements, D. H. (2001, March). Geometry must be vital. *Teaching Children Mathematics, 7*(7), 400-415.

Liu, L., & Cummings, R. (1997). Logo and geometrical thinking: concrete-abstract thinking and abstract-concrete thinking. *Computer in Schools, 12* (1-2), 1 – 10.

Liu, L., & Cummings, R. (2001). A learning model that stimulates geometric thinking through use of Logo and Geometer's Sketchpad. *Computers in the Schools, 17*(1/2), 85-104.

Malloy, C. E., & Friel, S. N. (1999, October). Perimeter and area through the van Hiele model. *Mathematics Teaching in the Middle School, 5*(2).

Manoucherhri, A. (1999). Computers and school mathematics reform: Implications for mathematics teacher education. *The Journal of Computers in Mathematics and Science Teaching, 18*(1), 31-48.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal of Research in Mathematics Education, 20* (1), 41-51.

McBride, B., & Carifio, J. (1995, April). Empirical results of using an analytic versus holistic scoring method to score geometric proofs: Linking and assessing Greeno, Bloom, and van Hiele views of student abilities to do proof. *Paper presented at the Annual Meeting of the American Educational Research Association.* April 18-22, 1995. San Francisco, CA.

McCrone, S. S., & Martin, T. S. (n.d.) *Tools for investigating high school students' understanding of geometric proof*. Retrieved October 21, 2003, from http://www.west.asu.edu/cmw/pme/resrepweb/PME-rr-mccrone.htm

McCrone, S. S., Martin, T. S., Dindyal, J., & Wallace, M. L. (2002). An investigation of classroom factors that influence proof construction ability. *Proceedings of the annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Athens, GA. October 26-29, 2002. (Obtained from ERIC Document No: ED 471 776).

McCrone, S. S., Martin, T. S., Dindyal, J., & Wallace, M. L. (2004). A longitudinal Study of mathematical reasoning: Student development and school influences: 1999–2003. *Technical Report: Y8, Y9 Proof Survey.* Retrieved January 15, 2005 from http://www.ioe.ac.uk/proof/techreps.html

McGehee, J. J. (1998, March). Interactive technology and classic geometry problems. *Mathematics Teacher, 91*(3), 204-209.

Mingus, T. Y., & Grassl, R. M. (1999). Preservice teachers' beliefs about proofs. *School Science and Mathematics, 99*(8), 438-444.

Moore, B. (2001, Jan/Feb). Taking stock of teacher technology use. *Multimedia Schools, 8*(1). 26-31.

National Council of Teachers of Mathematics (1980). *Agenda for action.* Reston, VA: Author.

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1973). *Geometry in the mathematics curriculum. National Council of Teachers of Mathematics 36th yearbook*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics.* Reston, VA: Author.

Olson, L. (2005, March 2). States take steps to put more rigor into high schools: College, work readiness are focus of governors. *Edweek, 24*(25), 1, 22.

Orange County Public Schools (n.d.). *Orange County Public Schools Enrollment Summary.* Retrieved December 22, 2004 from http://www.ocps.k12.fl.us/pdf/enroll.pdf

Oxford American Dictionary and Thesaurus. 2nd Ed. (2001, July). NY: Berkley Books.

Pandiscio, E. A. (2002, May). Exploring the link between preservice teachers' conception of proof and the use of dynamic geometry software. *School Science and Mathematics, 102*(25), 216-222.

Papert, S. (1980). *Mind-Storms: Children, computers, and Powerful Ideas.(all about LOGO and how it worked* NY: Basic Books. (2nd Edition 1993).

Polya. (1957). *How to solve it: A new aspect of mathematical method.* Garden City, NY: Doubleday. (Original work published in 1945).

Prince, A. A. (1998). Prove it! *Mathematics Teacher, 91*(8), 726-773.

Raman, M. (2003). Key ideas: What are they and how they can help us understand how people view proof? *Educational Studies in Mathematics, 52*, 319-325.

Rasmussen, S. (1992). *Engaging proofs using the Geometer's sketchpad*. Berkeley, CA: Key curriculum press.

Roberts, D. L., & Stephens, L. J. (1999). The effect of the frequency of usage of computer software in high school geometry. *The Journal of Computers in Mathematics and Science Teaching, 18*(1), 23-30.

Santos-Trigo, M., & Espinosa-Perez, H. (2002). Searching and exploring properties of geometric configurations using dynamic software. *International Journal of Mathematical Education in Science and Technology, 33*(1), 37-50.

Schacter, J. (1994). The impact of education technology on student achievement: What the most current research has to say. A meta analysis. Milken Exchange on Education Technology.

Segal, J. (1999, February). Learning about mathematical proof: conviction and validity. *The Journal of Mathematical Behavior, 18*(2), 191-210.

Senk, S. L., (1982, March). *Achievement in writing geometry proofs*. (Obtained from ERIC Document No: ED 218091).

Senk, S. L. (1983). Proof-writing achievement and Van Hiele levels among secondary school geometry students**.** *Unpublished doctoral dissertation*, University of Chicago.

Senk, S. L., & Usiskin, Z. (1983, February). Geometry proof writing: a new view of sex differences in mathematics ability. *American Journal of Education, 91*, 187–201.

Senk, S. L. (1985, September). How well do students write geometry proofs? *Mathematics Teacher, 78,* 448-456.

Senk, S. L. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education, 20* (3), 309-321.

Senk, S. L., & Thompson, D. R. (2003). *School mathematics curricula: Recommendations and issues. Standards-based School Mathematics Curricula: What are they? What do students learn?* Mahwah, NJ: Lawrence Erlbaum Associates.

Sharpiro, J., & O'Brien, T. C. (1970) Logical thinking in children ages six through thirteen. *Child Development, 41*, 823-829.

Shaughnessy, M., & Burger, W. F. (1985), September). Spadework prior to deduction in geometry. *Mathematics Teacher, 78,* 419-428.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences (3rd Ed.)* Boston, MA: Allyn and Bacon.

Smith E. P., & Henderson, K. B. (1989) *Proof. The growth of mathematical ideas, grades K – 12. National Council of Teachers of Mathematics, 24th yearbook*. Washington, DC: National Council of Teachers of Mathematics.

Sowder, L., & Harel, G. (1998). Types of students' justification. *Mathematics Teacher, 91*(8), 670-674.

StatSoft Electronic Textbook Inc. *Multiple regression.* Retrieved June 2, 2005 from http://www.statsoft.com/textbook/stmulreg.html

Steen, L. A. (1990). On the shoulders of giants: New approaches to numeracy. *National Research Council.* Washington, D. C.: National Academy Press.

Stover, N. F. (1989). An exploration of students' reasoning ability and van Hiele levels as correlates of proof-writing achievement in Geometry. (Doctoral dissertation, University of Oregon, 1990).

Susanna S. E. (2003, December). The role of logic in teaching proof. *The Mathematical Association of America, 110*, 886-899.

Tobin, K. G., & Capie, W. (1980a, April). The development and validation of a group test of logical thinking. *Paper presented at the American Educational Research Association Meeting*, April 1980. Boston, MA.

Tobin, K. G., & Capie, W. (1980b, April). The test of logical thinking: Development and applications. *Paper presented at the Annual Meeting of the National Association for Research in Science Teaching*, April 1980. Boston, MA.

Tobin, K. G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement, 41,* 413 – 423.

Usiskin, Z. (1982). Van Hiele levels and achievement in secondary school geometry. *Final report of the Cognitive Development and Achievement in Secondary School Geometry Project,* Chicago: University of Chicago, Department of Education.

Waring, S. (2001). Proof is back. *Mathematics in School, 30* (1), 4-8.

White House Office of the Press Secretary. (2004, September). The President's New Education Proposals. Retrieved September 4, 2004, from http://www.whitehouse.gov/news/releases/2004/09/print/20040902-3.html.

Winicki-Landman, G. (2002, March). Students as initiators of proofs. *Mathematics in School, 31*(2), 2–7.

Zhonghong, J. (2002). Developing preservice teachers' mathematical reasoning and proof abilities in the Geometer's Sketch environment. *North American Chapter of the International Group for the Psychology of Mathematics Education*, 24th Annual meeting. GA, October 26-29, 2002.