NAME OF AUTHOR/*NOM DE L'AUTEUR* EDGAR ROLAND WILLIAMS

TITLE OF THESIS/*TITRE DE LA THÈSE* AN INVESTIGATION OF SENIOR HIGH SCHOOL STUDENTS' UNDERSTANDING OF MATHEMATICAL PROOF

UNIVERSITY/*UNIVERSITÉ* UNIVERSITY OF ALBERTA

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* Ph. D.

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1979

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* DR. A. T. OLSON

DATED/*DATÉ* Dec 11, 1978 SIGNED/*SIGNÉ* E. R. Williams

PERMANENT ADDRESS/*RÉSIDENCE FIXÉ* 4 TAYLOR PLACE
ST. JOHN'S, NEWFOUNDLAND
A1A1L5

THE UNIVERSITY OF ALBERTA

AN INVESTIGATION OF SENIOR HIGH SCHOOL STUDENTS UNDERSTANDING OF THE NATURE OF MATHEMATICAL PROOF

by

EDGAR ROLAND WILLIAMS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

SPRING, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled AN INVESTIGATION OF SENIOR HIGH SCHOOL STUDENTS UNDERSTANDING OF THE NATURE OF MATHEMATICAL PROOF submitted by Edgar Roland Williams in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

....
(Supervisor)

....

....

....

....
(External Examiner)

Date Dec. 8, 1978

ABSTRACT

The purpose of this study was to assess the extent to which grade eleven students, who are enrolled in a mathematics program designed to prepare them for the study of post-secondary mathematics, subjectively understand the nature and role of mathematical proof. Specifically, the aim was to identify, categorize and describe some of the subjective thinking processes used by these students in attempting to justify a variety of mathematical generalizations and conclusions presented to them.

Twelve items constructed by the investigator after extensive pilot testing were used to assess the extent to which high school students understand a number of selected concepts related to proof in mathematics. These items were administered to two hundred fifty-five grade eleven students in ten randomly selected classes from nine different senior high schools of the Edmonton Public and Separate School systems. The items were administered during regular school hours and each student responded to each item in writing.

The responses of the students to each of the items were summarized and response categories for each item were developed on the basis of the type of responses that students actually gave to each of the items. The responses of students to each item were then categorized by three independent judges to ensure that each response was assigned to an appropriate response category. As a result, a distribution of responses by category for each item was obtained which

was used to analyze statistically the response patterns of students and also to describe the kinds of thinking exhibited.

The major findings can be summarized as follows:

1. Only those students who were classified as high achievers by their teacher, less than thirty percent of the sample, exhibited any understanding of the meaning of proof in mathematics.

2. Approximately half of the students sampled do not see any need to prove a mathematical proposition which they consider to be intuitively obvious.

3. At least seventy percent of the students sampled do not distinguish between inductive and deductive reasoning and hence do not realize that induction is inadequate to support mathematical generalizations.

4. Approximately eighty percent of the students sampled do not always realize the significance of hypotheses and definitions in mathematical arguments. Approximately sixty percent of the students sampled are unwilling to reason, for the sake of argument, from any hypothesis which they consider is false.

5. Less than twenty percent of the students sampled understand the method of indirect proof.

6. Almost eighty percent of the students do understand the concept of counter-example.

7. There was no evidence from this study to suggest that high school students understand that a mathematical statement and its contrapositive are logically equivalent or that students realize that a statement and its converse are not logically equivalent.

8. There were no significant differences between the overall level of responses given by females and that given by males.

ACKNOWLEDGEMENTS

I would like to acknowledge my indebtedness and express my sincere gratitude to a number of individuals whose time, efforts and encouragement has contributed to this study. In particular, I would like to thank:

Dr. Al Olson for his expert guidance, invaluable assistance, and, perhaps, extreme patience.

Dr. Tom Kieran for his many helpful suggestions and assistance in analyzing some of the data.

Dr. Sol Sigurdson, for his encouragement.

The members of my Supervisory Committee for their trouble and time in reading and commenting upon this thesis.

The principals, teachers and students of the nine schools from the Edmonton Public and Separate School systems for their interest and cooperation.

Dr. Ted Hobbs for his encouragement and assistance.

Ms. Elaine Boone and Ms. Rosalind Genge for their typing of the manuscript.

Mr. W. J. Blundon, Dr. J. H. Burry, and Dr. B. Gardner of the Department of Mathematics, Memorial University for their cooperation.

Dr. M. O. Morgan, President of Memorial University for the opportunity and financial assistance to carry out this study.

My wife and children for their understanding, encouragement and continuous assistance.

TABLE OF CONTENTS

CHAPTER PAGE

I INTRODUCTION 1

Background to the Problem 1

Statement of the Problem 10

Definitions 11

Significance of the Study 13

Description of the Study 15

Hypotheses Tested 18

Limitations 19

Outline of the Report 21

II REVIEW OF RELATED LITERATURE 23

III RESEARCH PROCEDURES 34

The Pilot Study 34

Instrumentation 42

Situation One 44

Situation Two 46

Situation Three 47

Situation Four 49

Situation Five 50

Situation Six 51

Situation Seven 52

Situation Eight 54

Situation Nine 54

CHAPTER PAGE

Situation Ten 55

Situation Eleven 56

Situation Twelve 57

The Sample 57

Testing Procedures 58

The Categorization Scheme 63

The Scoring and Analysis of Responses 75

IV RESULTS OF THE STUDY

Distribution of Responses by Category 78

Situation One 78

Situation Two 82

Situation Three 86

Situation Four 91

Situation Five 94

Situation Six 99

Situation Seven 103

Situation Eight 107

Situation Nine 110

Situation Ten 112

Situation Eleven 114

Situation Twelve 116

Results of Testing the Hypotheses 118

Hypothesis One 121

Hypothesis Two 121

CHAPTER PAGE

Hypothesis Three 124

Hypothesis Four 126

Summary 127

V SUMMARY, CONCLUSION, DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS 129

Summary of the Study 129

The Instrument 129

Procedures 131

The Sample 132

Discussion of Some of the Findings 132

Question One 132

Question Two 136

Question Three 138

Question Four 139

Question Five 142

Question Six 143

Question Seven 144

Discussion of the Results of the Hypothesis Testing 145

General Observations 152

Education Implications 155

Suggestions for Further Research 158

Epilogue 161

BIBLIOGRAPHY 163

APPENDIX 169

VITA 181

LIST OF TABLES

Table	Description	Page
I	The Sample	59
II	Mathematics Achievement Levels by Group	60
III	Response Scoring Scheme	77
IV	Distribution of Responses by Category for Situation One	79
V	Distribution of Responses by Category for Situation Two	83
VI	Distribution of Responses by Category for Situation Three	87
VII	Distribution of Scores on Situations Two and Three	88
VIII	Distribution of Responses by Category for Situation Four	92
IX	Distribution of Responses by Category for Situation Five	95
X	Distribution of Responses by Category for Situation Six	100
XI	Distribution of Responses by Category for Situation Seven	104
XII	Distribution of Responses by Category for Situation Eight	108
XIII	Distribution of Responses by Category for Situation Nine	111
XIV	Distribution of Responses by Category for Situation Ten	113
XV	Distribution of Responses by Category for Situation Eleven	115
XVI	Distribution of Responses by Category for Situation Twelve	117

Table	Description	Page
XVII	Mean Achievement and Situation Scores	119
XVIII	Distribution of Scores for each Situation	120
XIX	Analysis of Variance on Mean Total Scores for Each Class	122
XX	Correlations between Total Score and Achievement for each Group	123
XXI	Intercorrelations between Achievement and Situation Scores	125
XXII	Comparison between the Scores of Males and Females	128

CHAPTER I

INTRODUCTION

I. BACKGROUND TO THE PROBLEM

An introduction to the deductive nature of modern mathematics has been widely accepted as a desirable and important goal of the high school mathematics curriculum, especially for students who plan to continue their study of mathematics at the post-secondary level. This goal has perhaps taken an added significance in this century in view of the rapid growth of new knowledge and new conceptualizations of old knowledge in all existing fields of mathematics. During the past half century, however, there has been a great deal of debate in the mathematics community concerning the most appropriate means of attaining this goal.

The central notion of deductive mathematics is that of proof. E. T. Bell (1936) stated:

> Unless the student who gets no farther than a first course in algebra or geometry acquires as part of his mentality for life a clear, cold perception of what "proof" means in any deductive argument or system of deductive reasoning, his time and effort will have been wasted (p.138).

The significance of proof as one of the pivotal ideas in mathematics is conveyed by Polya (1957):

> In mathematics, as in the Physical sciences, we may use observation and induction to discover general laws. But there is a difference. In the Physical sciences, there is no higher authority than observation and induction; but in mathematics there is such an authority: rigorous proof (p.117).

Kline (1973) supports these comments by stating:

> There is no question that deductive proof is the hallmark of mathematics. No result is accepted into the body of mathematics until it has been proved deductively on the basis of an explicit set of axioms (p.188).

If these comments are at all representative, then it is apparent that many mathematics educators during the past fifty years have pointed to the importance of providing high school students with some understanding of the nature and role of mathematical proof. This point was perhaps best summarized by Fawcett (1938) when he stated:

> There is no disagreement concerning the educational value of any experience which leads children to recognize the necessity for clarity of definition, to weigh evidence, to look for the assumptions on which conclusions depend, and to understand what proof really means (p.6).

In the traditional high school mathematics curriculum, Euclidean geometry, to which Fawcett was referring above, was the sole vehicle by which high school students were introduced to deductive mathematics. Allendoerfer (1957) states that one of the main objectives of teaching Euclidean geometry at the high school level is "to teach the deductive method as it is applied to mathematical reasoning and thus give the students a first taste of the nature of mathematical proof (p.65)." While this objective was widely accepted as a reason for teaching Euclidean geometry during the first half of this century, at the same time many mathematics educators began to express doubts and concern about the extent to which this objective was actually being achieved in the usual manner in which Euclidean geometry was taught.

Bell (1936) questioned "why Mathematics, of all Sciences the most progressive and the most prolific in its research activities, is the most backward pedagogically (p.139)." He went on to state:

> Who but a demented reactionary would teach physics to boys of fifteen out of Aristotle's "Physics"? Yet the equivalent of that unthinkable stupidity is precisely what we do in geometry (p.139).
>
> Beyond the few trivial applications, what is there in elementary geometry but a training in deductive reasoning? Nothing. And unless the training is modernized, the habits of "reasoning" which are drilled into the pupils are about as bad as they could be. The pity of it is that a decent job would be no harder to do than the awkward muddle consecrated by tradition and sanctioned by mential inertia. Endless generations of committees on the teaching of geometry have proposed timid patches here and there on a corpse that has lain in state for generations (p.140).

Fawcett (1938) reported that while "there is almost unanimous agreement that demonstrative geometry can be so taught that it will develop the power to reason logically (p.8)", there is also almost unanimous agreement that most teachers of geometry do not ordinarily teach it in such a way as to secure the desired learning outcomes. In elaboration of this point, Fawcett states:

> While verbal allegiance is paid to these large general objectives related to the nature of proof, actual classroom practice indicates that the major emphasis is placed on a body of theorems to be learned rather than on the method by which these theorems are established. The pupil feels that these theorems are important in themselves and in his earnest effort to "know" them he resorts to memorization (p.1).

Many other mathematics educators offered similar criticisms of the traditional Euclidean geometry curriculum. Rosskopf (1957) suggests that in the traditional manner of teaching mathematical proof, "the average student never really learns what a proof is, but rather learns how to write correct proofs through imitation of his instructor and his textbooks and by adjusting his efforts to their authority (p.21)." Kline (1973) makes a similar point when he states:

> The concept of proof is fundamental in mathematics, and so in geometry the students have the opportunity to learn one of the great features of the subject. But since the final deductive proof of a theorem is usually the end result of a lot of guessing and experimenting and often depends on an ingenious scheme which permits proving the theorem in proper logical sequence, the proof is not necessarily a natural one, that is, one which would suggest itself readily to the adolescent mind. Moreover, the deductive argument gives no insight into the difficulties that were overcome in the original creation of the proof. Hence the student cannot see the rationale and he does the same thing in geometry that he does in algebra. He memorizes the proof (p.7).

A related but perhaps more significant deficiency in the traditional teaching of Euclidean geometry centres around the kind of tests that were used to assess learning outcomes in geometry. Fawcett (1938) claimed that "the tests most commonly used emphasize the importance of factual information (p.1)." More importantly, however, he states:

> It is probably safe to assume that the values emphasized in the testing program of any school are those values which receive emphasis in the classroom, and a study of commonly used tests

> in geometry is sufficient to reveal that little, if any, attempt is made to measure the degree to which the purposes claimed for demonstrative geometry are realized (p.9).

Fawcett further states that "there is little, if anything, in these tests which by any stretch of the imagination could be interpreted as examining children on their understanding of 'the nature of proof' (p.9)."

The lack of and need for items that do test many of the objectives stated for Euclidean geometry was recognized by Fehr (1966). He stated:

> The major areas which have been neglected in mathematics test development are: the understanding of proof, consistency of a mathematical system and a whole spectrum of related topics. The need is to address oneself to the development of better questions at this level, on this vast spectrum that I would call the proof spectrum (p.61).

In a more general sense, Wilson (1971), in the context of Bloom's Taxonomy (1956), classifies objectives in the cognitive domain of mathematics in a hierarchy as follows: (a) computational; (b) comprehensive; (c) application and (d) analysis. In commenting on these categories, Wilson states:

> Mathematics teachers often state their goals of instruction to include all cognitive levels. They want their students to be able to solve problems creatively, etc. But then their instruction, their testing, and their grading tend to emphasize the lower behaviour levels, such as computation ard comprehension (p.650).

It is evident that the important objectives stated for teaching Euclidean geometry in the past, particularly those dealing with the notion of proof, are at the higher levels of Wilson's hierarchy. It is also evident that it is much more difficult to assess in a meaningful way learning outcomes at the analysis level than say at the computational level. If teachers experience difficulty in their attempts or do not know how to evaluate meaningfully the learning outcomes implied by a specific goal of the curriculum, then it is doubtful whether any meaningful teaching for that goal will or can take place. Hence, the degree of difficulty associated with assessing learning outcomes may be a prime determinant of whether or not certain objectives are included in the curriculum. If meaningful teaching can only take place when teachers can assess the results of their teaching and if meaningful learning can only take place when students can determine the extent to which they have learned, then it seems certain that many of the desired learning outcomes in the area of mathematical proof have not been attained in the past because the method of evaluation most often used did not assess learning outcomes relative to the stated goals but rather have assessed in many cases learning outcomes that many would consider irrelevant. Hence the lack of or tendency not to use test items appropriate to assess the extent to which students have attained the stated objectives of the geometry curriculum may be one of the main reasons for many of the criticisms directed at Euclidean geometry.

Euclidean geometry, however, came under fire for other reasons. In particular, while the content of Euclidean geometry remained basically unchanged for centuries, the maturity level of many of the students who enrolled in it changed dramatically. A century ago, Euclidean geometry was taught only at the college level, presumably to a select group of students. However, in the middle of the twentieth century, it was being taught to rapidly increasing numbers of high school students. The fact that a much larger proportion of the population were staying on to complete a high school education meant that students with a much wider range of interests, maturity levels, and intellectual capacities were enrolling in high school geometry courses. The difficulties experienced by many of these students, especially in their attempts to understand what geometry was all about, became much more in evidence and the demand for change increased.

Allendoerfer(1957), in advocating change, pointed to another criticism of Euclidean geometry, when he stated:

> As mathematics has developed today, there seems to be no compelling reason for using Euclidean geometry as the principal example of the deductive method of logical reasoning. As a matter of fact, there are reasons to believe that Euclidean geometry is even an unfortunate example to use with beginning students. Euclidean geometry is quite a complicated mathematical system, and as presented in most textbooks is not even completely logical (p.66).

In the same context, many mathematics educators pointed out that since students were being introduced to deductive methods only in Euclidean geometry, many students considered geometry to be the only area of mathematics where such methods were being used. Hence it was argued that deductive methods should and could be introduced in other areas of the school mathematics curriculum.

Allendoerfer (1957), among many others, echoed the comments of Bell (1936) in suggesting that deductive methods be introduced in high school algebra. Bell argued that "Geometry is considerably more complicated structurally than algebra (p.145)." Allendoerfer (1957) suggests:

> The reason that Euclidean geometry has been used traditionally as the prime example of logical reasoning is that until recent times there was no other example to which teachers could turn. There are now available, however, a number of other mathematical examples of deductive systems which lend themselves much more effectively to this purpose (p.66).

The suggestions of Allendoerfer and many others were sanctioned in the far reaching 1959 Report of the Commission on Mathematics of the College Entrance Examination Board. One recommendation that was most influential in much of the curriculum reform to follow, suggested that students be introduced to the nature and role of deductive reasoning in algebra as well as in geometry. As a result, many of the groups involved in the movement to reform the school mathematics curriculum in the late nineteen fifties, notably the School Mathematics Study Group, began writing and publishing textbooks which presented an axiomatic approach to the study of the number systems and which attempted to introduce deductive methods at an early level in the algebra curriculum.

Kline (1973), while being very critical of many of the reforms of the school mathematics curriculum, summarized the results of these efforts by stating that "the major innovation of the new mathematics is the deductive approach to traditional subject matter (p.40)." Allen (1966) went even further when he stated:

> The emphasis on structure and proof in algebra is the fundamental component of a change that has taken place in school mathematics in the United States at the secondary level during the past ten years. This change is so profound and far reaching that it only can be described as a revolution (p.3).

It was argued by many that the introduction of deductive methods in algebra would mean that students would no longer be forced to memorize and rely upon mechanical drill but would be forced to think about and understand the basic algebraic concepts being studied. In the process, it was envisioned that students would gain a greater appreciation of the nature of deductive methods in mathematics generally and in particular would understand better the nature and role of mathematical proof.

The result is that during the past two decades, many students have been introduced to deductive methods much earlier and over a longer period of time. At the same time, however, many schools have reduced or even eliminated altogether the teaching of formal Euclidean geometry, perhaps in response to the famous statement of Dieudonne (1961) when he argued that "Euclid must go (p.35)."

In any event, the time would appear ripe to attempt to assess whether or not students who have been exposed to the modern mathematics curricula have any better idea of the concept of proof than was claimed for students who studied proof only in traditional Euclidean geometry.

II STATEMENT OF THE PROBLEM

The general objective of this study was to investigate and attempt to determine the extent to which grade eleven students, who have studied a modern high school curriculum designed to prepare them for the study of post-secondary mathematics, understand the nature and role of mathematical proof. Specifically, the aim was to identify and categorize some of the subjective thinking processes used by these students in attempting to justify a variety of mathematical generalizations and conclusions presented to them. In this context, answers were sought to the following questions:

1. To what extent or in what contexts do high school students subjectively see the need for proof in mathematics?
2. To what extent do high school students subjectively recognize that induction is inadequate to support mathematical generalizations?
3. To what extent do high school students subjectively realize that the objects studied in mathematics have only those properties ascribed to them by definition or by postulate?

4. To what extent do high school students subjectively understand the indirect method of proof in mathematics?
5. To what extent do high school students subjectively realize that a single counter-example to a stated proposition is sufficient to reject that proposition?
6. To what extent do high school students subjectively realize that a mathematical statement and its contrapositive are logically equivalent?
7. To what extent do high school students subjectively recognize that a mathematical proposition and its converse are not logically equivalent?

III DEFINITIONS

Deductive Proof. A chain of reasoning based upon accepted assumptions (called axioms or postulates), definitions and/or previously proven propositions, which, provided the accepted rules of logic are followed, demonstrates that a conclusion is necessarily true if the postulates on which the argument is based are accepted as true.

Proposition. A mathematical statement about two or more terms defined or undefined.

Axiom (Postulate). A mathematical proposition assumed to be true.

Theorem. A mathematical proposition which is not an axiom and for which a proof has been given.

Definition. A mathematical statement which is not a proposition and is stated in terms of previously defined terms or previously undefined terms or any combination of these.

Conjecture. An unproven proposition.

Counter-example. If P(x) is a mathematical proposition where x is any element in some domain D, then for any constant c in D, P(c) is a counter-example to P(x) if P(c) is known to be false.

Induction. The process of reasoning by which a generalization is reached from a study of particular facts.

Contrapositive. The contrapositive of the statement p implies q where p and q are statements, is the statement not-q implies not-p.

Converse. The converse of the statement p implies q is the statement q implies p.

Indirect proof. A mathematical proposition P is proven true indirectly if one can deduce a contradictory statement from the negation of P.

IV SIGNIFICANCE OF THE STUDY

As was indicated in the introduction, traditional attempts to teach the nature and role of mathematical proof at the high school level do not appear to have been very successful. Many teachers encountered great difficulties in their attempts to teach the nature of proof and blamed the traditional Euclidean geometry curriculum as the main culprit responsible for this state of affairs. Although there has been an increased emphasis on presenting a deductive approach to much of the traditional curriculum during the past two decades, very little independent research has been carried out to assess whether or not students have any better understanding of the concepts being studied, particularly in the area of mathematical proof. From a curriculum point of view, it would be instructive to know whether or not students face the same problems today as they did thirty years ago, or are there new difficulties being encountered.

These questions are particularly relevant in view of many of the criticisms aimed at the mathematics curriculum in recent years, notably by Kline (1973). The distinguished French mathematician René Thom has also taken issue with recent curricular innovations in school mathematics. In particular Thom (1971) states:

> the contemporary trend to replace geometry with algebra is educationally baneful and should be reversed. There is a simple reason for this: while there are geometry problems, there are no algebra problems. A so-called algebra problem can only be a simple exercise requiring the blind application of arithmetical rules and of a pre-established procedure. With rare exceptions, one cannot ask a student to prove an algebra theorem; either the requested answer is almost obvious and can be arrived at by direct substitution of definitions, or the problem falls into the category of theoretical algebra and its solution exceeds the capacities of even the most gifted student. Exaggerating only slightly, one can say that any question in algebra is either trivial or impossible to solve. By contrast, the classic problems of geometry present a wide range of challenges (p.696).

Another aspect of this study deemed significant concerns the issue of testing. As pointed out in the introduction, the obvious lack of good items to assess learning outcomes in the proof domain may have been one of the main reasons for the lack of success in teaching these concepts. There is very little evidence to suggest that the testing situation, at least in the area of proof, has changed very much. Hence a secondary objective of this study was to determine whether or not items of the type developed for and used in this study are useful as a means of diagnosing and evaluating the extent to which students do understand many of the concepts related to proof.

Finally, the majority of recent studies in the area of mathematical proof have been experimental in nature. It is apparent that many of these studies have not been particularly helpful in the area of curriculum development. A number of related studies, many of which are summarized by Roberge (1972) and Jannson (1974), have investigated the development of deductive reasoning in children. Some of these studies have attempted to determine and classify fallacies in logical reasoning. However, the main focus of this study was not on logical reasoning abilities and inference patterns. The attempt is not to determine the extent to which students are good at proof-making but rather to determine the extent to which they understand what proof in mathematics really means. Hence, while this study is not experimental in nature, it was hoped that some useful information that could benefit teachers would result.

V DESCRIPTION OF THE STUDY

The first and perhaps most important stage of this study was the development of items to provide the kind of information sought. Initially, a number of items were devised which collectively were designed to assess student understanding of a wide variety of concepts in the proof domain. A panel of three judges examined the original set of items and made suggestions. As a result, some of the original items were deleted and others were modified.

The revised set of items were used to construct six versions of an instrument. Each version, with one exception, consisted of four distinct items. Each item appeared on a separate page with the item at the top and space at the bottom for the student to respond in writing.

Each version of the instrument was piloted in randomly selected mathematics classes of grade eleven students enrolled in the academic programs of the Edmonton Public and Separate school systems. The instruments were administered during the first thirty minutes of a regularly scheduled mathematics period.

The responses of each student to each item was examined by three independent judges. This examination was carried out in two stages. First, for each item, the responses of all students responding to that item were summarized. As a result, a set of mutually disjoint response categories was devised for each item. Secondly, the three judges were asked to independently place each response to a given item in exactly one of the response categories for that item. This procedure was used initially to determine the adequacy of the categories and later to determine the reliability of assigning each response to the appropriate category.

The objectives of this type of analysis were firstly to determine the feasibility of using such a procedure on a much wider scale. Secondly, while the six versions of the pilot instrument contained essentially different items, in several instances, the same item was included in different formats on different versions of the pilot instrument. The objective was to determine whether or not the way in which the items were presented or worded made any difference in the type of responses given by students.

As a result of this analysis, each item used in the pilot study was judged to be either suitable as is, suitable with modifications or unsuitable. It was determined that the format in which an item was presented did make a significant difference in the type of responses students gave and that one format was superior to the others.

As a result of the pilot study, twelve items were eventually selected and used in this study. All items were similar to or modifications of items used in the pilot study. As in the pilot study, each item was presented on a separate page with sufficient space on the page for the students to respond to the item.

It was decided that the instrument be administered during a regular class session of approximately eighty minutes. In order to determine whether students could reasonably respond to the items in the time available, the proposed instrument was given a trial administration to one class of students. In addition, the responses of these students were analyzed in the same manner as in the pilot study.

As a result, several minor changes were made in the wording of some of the items. Also, it was decided that students would be requested to respond to eight of the twelve items selected.

The items were then administered to two hundred fifty-five grade eleven students in ten randomly selected classes of nine senior high schools in the Edmonton Public and Separate School systems. Seven of the items were administered to all students and the remaining five items were administered at random so that in any given class each of these items was administered to approximately twenty percent of the students.

The responses of students to each of the items were then summarized and response categories for each item developed using the same procedure as in the pilot study. The responses of each student to each item were categorized by three independent judges.

An analysis was carried out to investigate whether or not there was any relationship between student response on each of the items and achievement. The main analysis, however, consisted of an examination of student responses to each item with particular emphasis on the type of thinking exhibited for each category of response.

VI HYPOTHESES TESTED

In order to analyze the response patterns of students statistically, a scoring scheme was devised based upon the hierarchial ordering of the response categories for each item. Scores of 0, 1, 2 or 3 were assigned to each student's response to each item. The scoring scheme was used to test the following hypotheses:

1. The mean total scores for each group of students who were administered the instrument are not significantly different.

2. For each group of students who were administered the instrument, there are no significant correlations between achievement in Mathematics 20 and the total score obtained by each student.

3. There are no significant correlations between achievement in Mathematics 20 and
   (a) the score obtained on each situation
   (b) the total score obtained on the eight situations presented to students.

4. There are no significant differences between
   (a) the mean scores of males and females on each of situations one through seven
   (b) the mean total scores of males and females.

VII LIMITATIONS

The results of this study are perhaps only as valid as the items used to obtain those results. Although extensive pilot testing was carried out in an attempt to generate items to which students would respond in a subjective manner, it was not possible to objectively determine the degree of reliability of these items other than to observe that there was a high degree of consistency in the response patterns of students.

The fact that students could only respond in writing and were required to read and understand the situation presented in each item in order to respond meaningfully, could have affected the type of response given. Although the situation presented in each item was considered to be such that most students would have no difficulty in relating to and understanding what was written and what was expected of him or her, it is possible that the kind of responses given by students depended upon which aspect of the situation the students happened to focus on in his or her initial reading of it. Hence, if it were possible to determine in any objective or consistent way which aspect of the situation, if any, a student zeroed in on, then perhaps the responses of that student would have been categorized differently.

The response categorization scheme for each item was not pre-determined but constructed on the basis of the various kinds of responses students actually gave for each of the items. Hence, while the categorization scheme was considered adequate by three independent judges, the fact remains that the categories were developed from the perspective of the three judges which may be entirely different from that of, for example, the students in making these responses. In this sense, the categorization scheme developed for each item together with the assignment of student responses to these categories may be open to question. Three independent judges were used, however, to minimize such possibilities and increase the realibility of the procedures employed.

Every effort was made to ensure that students did not consider the instrument as a test of their knowledge. For example, the students did not know in advance that they were participating in this study. It is still possible, however, that some students perceived it as such, especially since the instrument was administered during regular school hours in a regularly scheduled mathematics class. If so, then this could have had an effect upon the type of responses students gave. Also the pressure of time may have affected student responses although there was no evidence of this.

Finally, the results obtained and the conclusions which are suggested by these results cannot be generalized beyond the population sampled, namely, grade eleven students enrolled in an academic mathematics program who have studied a modern mathematics curriculum including some work in formal Euclidean geometry.

VIII OUTLINE OF THE REPORT

Chapter II contains a review of selected relevant literature. A detailed account of the design of the study including the pilot study, the items used and their rationale, the testing procedures used, the response categorization scheme employed and the hypotheses tested relative to the response patterns of students is reported in Chapter III. Chapter IV reports the results of the study including a breakdown of student responses for each of the items together with examples and a discussion of student responses in each category of each item. Also the results of the statistical analysis of the response patterns of students are presented in this chapter. Chapter V includes a summary and discussion of the findings with reference to the specific questions posed in the statement of the problem. Also a discussion of some of the educational implications of the findings and suggestions for further research are given in this chapter.

CHAPTER II

REVIEW OF RELATED LITERATURE

Numerous studies that might be classified as having implications for the teaching, learning and assessment of various concepts related to mathematical proof have been reported in the literature. Some of the categories under which these studies fall include The Teaching of Proof, The Development of the Concept of Proof, Deductive Reasoning Studies, Critical Thinking Studies, Piagetian Studies, Studies on the Teaching of Inference Patterns and Logic and so on. It is not the purpose of this chapter to review studies in all of the above categories but rather to review and summarize some of the investigations which appear to be directly related to this study.

One of the classic studies in the area of mathematical proof was that reported by Fawcett (1938) in the Thirteenth Yearbook of the National Council of Teachers of Mathematics. The main purpose of his study was "to describe classroom procedures by which geometric proof may be used as a means for cultivating critical and reflective thought and to evaluate the effect of such experiences on the thinking of the pupils (p.1)". Fawcett developed an experimental course in plane geometry which he taught over a two year period. The basic purpose of this course was to teach students the principles of logical deduction and determine the extent to which students could use these principles in non-mathematical situations.

In his study, Fawcett assumed that a student understands the nature of deductive proof when he understands:

> (a) the place and significance of undefined concepts in proving any conclusion;
>
> (b) the necessity for defined terms and their effect upon the conclusion;
>
> (c) the necessity for assumptions and unproved propositions;
>
> (d) that no demonstration proves anything that is not implied by the assumptions (p. 10).

Hence, throughout his course, Fawcett attempted to teach his students the role of undefined terms, definitions and assumptions in making deductive inferences, not only in geometry but in a variety of non-mathematical situations as well. He attempted to show that with explicit instruction, students could be taught to apply the principles of deductive reasoning used in plane geometry to many situations occuring in everyday life.

Fawcett concluded that:

> 1. Mathematical method illustrated by a small number of theorems yields a control of the subject matter of geometry at least equal to that obtained from the usual formal course.
>
> 2. By following the procedures outlined, it is possible to improve the reflective thinking of secondary school pupils.
>
> 3. This improvement in the pupil's ability for reflective thinking is general in character and transfers to a wide variety of situations.

> 4. The usual formal course in demonstrative geometry does not improve the reflective thinking of the pupils (p.119).

Rolland Smith (1940) investigated the difficulties encountered by students in writing proofs in plane geometry. Students studying units on congruence and parallel lines were tested almost daily over a fifty day period. The written tests were used to determine the type and prevalence of errors made by the students. Smith found that most of these errors could be classified in three categories, one of which was "those (errors) due to a meager understanding of the meaning of proof (p.2)". His study showed that errors in each category persisted throughout the study but that errors in the category mentioned above were more prevalent. Smith's results were similar to those obtained earlier by Touton (1924).

There are a number of interesting studies of more recent vintage. Many of these were carried out at the time when efforts to reform the traditional mathematics curriculum were at a peak. A recommendation to place more emphasis on providing college capable students with an "understanding of the nature and role of deductive reasoning in algebra as well as in geometry (p.iii)" was contained in the 1959 Report of the Commission on Mathematics of the College Entrance Examination Board, referred to earlier in this report. As a result, the attention of many mathematics educators was drawn to the problem of how best to incorporate this objective in the school mathematics curriculum.

E. P. Smith (1959) suggested that except in plane geometry, there had up to that point been very little research into methods by which concepts in the proof domain could be made an integral part of the school mathematics curriculum. Hence Smith suggested ways of introducing the basic ideas of proof into the school curriculum at both the elementary and high school levels. The results of his work are contained in the Twenty-fourth Yearbook of the National Council of Teachers of Mathematics.

The 1959 Report of the Commission on Mathematics together with the work of Smith mentioned above led to a study by Robinson (1964), which perhaps influenced the present study more than any other. Robinson attempted to assess the extent to which "the nature and role of deductive reasoning in mathematics can be meaningfully taught at an earlier level than that which has been customary in the schools (p.2)", as implied by the recommendations of the Commission on Mathematics. The general aims of her research were to explore (a) "what students themselves regard as proof in mathematical situations (p.3)" and (b) the extent to which "junior high school students understand the need for proof in mathematics (p.98)".

Robinson sought answers to four specific questions. The first was: "As students grow older, do they, without direct instruction, come to realize that generalizations about infinite sets of mathematical objects cannot be adequately supported by examining a finite number of cases (p.30)?" Robinson concluded that children do have this understanding at least by grade seven. She states that "whether or not seventh grade students specifically recognize that induction is adequate to support generalizations in mathematics, their responses to situations of free choice are characterized by other forms of reasoning, proof being one of these (p.100). Robinson points out in this context, however, that "no claim is made that these children understand the meaning of the phrase, 'mathematical proof' (p.99)."

Robinson's second question was: "As children grow older, do they naturally come to realize that the objects studied in mathematics have only those properties ascribed to them by definition or by postulate (p.30)?" Robinson concluded that the answer to this question was no, at least for students at the junior high school level. She states that "most seventh and ninth-grade students will, when given free choice, justify a mathematical generalization by deductive reasoning from a set of premises if and only if these premises agree with their intuitions; that is, if and only if these premises are believable to the children (p.101)."

The third question posed by Robinson was: "Does formal instruction in the use of strategies of proof contribute to the understandings described in (the first two questions)(p.30)?" Again Robinson concluded that the answer to this question was no. She states that "there is no evidence from this study to suggest that instruction in plane geometry contributes to a student's understanding of the need for proof or to his understanding of the nature of an axiomatic (mathematical) system (p.102)."

It should be noted, perhaps, that Robinson's sample consisted of forty-eight students, half of whom were enrolled in grade seven and the other half in grade nine. Only eight of the ninth grade students had completed a year of plane geometry. The third conclusion above was based upon a comparison of the performance of these eight students with the others. It is doubtful whether any meaningful conclusions can be drawn from such a small sample. Furthermore, there is the implication that the instruction in plane geometry received by these students included a study of various strategies of proof. However, whether or not this was the case is not verified in the study. In fact, no information at all is given on the nature of the geometry course studied by these students.

The final question asked by Robinson was "If students do see the need for proof, is this understanding more related to mental maturity or to achievement in mathematics (p.30)?" Robinson obtained no conclusive evidence on this very important question. She suggests that to adequately answer this question, "a further study is needed in which reliable measures are available for some of the traits which were not sampled (p.104)."

The study of Robinson has been reported in some detail because it appears more closely related to the present study than any other and also because the present study is in some respects an extension of Robinson's work to the high school situation.

A study by Reynolds (1967) as reported by Lovell (1971) also has significance for the present study. The aim of this study was "to investigate the development of the understanding of mathematical proof in pupils in British selective (grammar and technical) secondary schools and to see how well this development is explained by the framework provided by Piaget's genetic psychology (p.66)."

Reynolds constructed tests involving the following aspects of proof: generalizations, symbols, assumptions and methods of proof. These tests were administered to four groups of students in the age categories 12-13 (First Form in Britain), 13-14 (Third Form), 15-16 (Fifth Form) and 17-18 (Sixth Form). Lovell states that "by means of the common questions to every age group. it was possible to get some idea of the development with age of the understanding and use of the aspects of proof considered (p.68)." Reynolds conclusions were based upon an analysis of the written responses of the students to the various test items.

In summarizing some of the main conclusions, Lovell (1971) states that while "Piaget's formulations regarding stages of thinking account for a good deal in the nature of the replies (p.77)", yet "there were common approaches to the questions in all age groups (p.77)", and "to some questions the answers of the fifth and sixth-form pupils showed only gradual improvement over those in the first and third forms (p.77)."

One of the more important aspects of Reynolds' study is the type of questions he used, some of which, especially those dealing with generalizations, attempt to assess understanding not usually tested with traditional kinds of items. However, most of the items used by Reynolds appear to be knowledge oriented. Hence students may have responded to these items by relying heavily upon what they had learnt in school which may not be indicative of their stage of development with respect to the concepts being investigated. Hence it is possible that what Reynolds assessed is not the development of an understanding of proof, but rather how well concepts in the proof domain were taught at the various levels. If so, then it is possible that entirely different conclusions could have been reached.

The type of research that Reynolds was engaged in can be classified as developmental or Piagetian-oriented. There have been a number of studies of a similar nature investigating the development of deductive reasoning and critical thinking. In particular, the work of Ennis and Paulus (1962) is probably the most significant. Since most of these studies are concerned with the verbal form of arguments and are not considered to be directly related to this study, no attempt will be made to review them here. However, an excellent summary of the results of many of these studies has been given by Roberge (1972).

Several studies have investigated the role of proof-related logic. Morgan (1970) found that mathematical experience was not a sufficient condition for learning all of the patterns of conditional reasoning that he investigated, namely recognition of (a) equivalence of a conditional statement and its contrapositive; (b) invalidity of the inverse and converse of a conditional statement, and (c) the starting assumption for a direct proof, contrapositive proof and a proof by contradiction.

Some educators have argued that the basic concepts of mathematical logic should be taught as a necessary pre-requisite to the teaching of concepts in the proof domain. The work of Suppes (1962), Dienes (1964) and Scandura (1971) suggest that the teaching of basic inference patterns should begin in the elementary grades. On the other hand, studies by Phillips (1968), Roy (1970) and Mueller (1975) suggest that formal instruction in logic does not significantly aid students in the writing of proofs.

Phillips (1968) attempted to determine whether or not there were any differences in the abilities of two groups of high school students to analyze and construct simple mathematical proofs. One group had studied the formal rules of deduction, while the other had studied these rules informally in their regular mathematics instruction. Both groups were tested on their ability to (a) complete information in partial informal proofs; (b) give original informal proofs; and (c) analyze a given proof for correctness or incorrectness. Phillips concluded that there were no significant differences in the ability of the two groups to deal with the three aspects of informal mathematical proofs referred to above.

In a similar study, Roy (1970) investigated the extent that it is necessary to teach explicitly those basic logical principles that are essential to deductive reasoning. His primary aim was to determine whether or not the formal study of mathematical logic improves the ability of secondary school mathematics students to (a) judge the validity of arguments and (b) prove theorems by using the principle of mathematical induction. Roy concluded that the study of mathematical logic has little or no effect on how students perform in carrying out any of the two tasks studied. Mueller (1975) obtained similar results.

Bostic (1970) assumed that the formal study of the basic principles of mathematical logic should precede the introduction of proof. As a result, he attempted to develop materials suitable for an introduction to the concept of proof at the tenth grade level and to establish their suitability for presentation at this level. The materials consisted of twenty-two lessons, the first eight of which contained an informal introduction to the basic concepts of logic. Perhaps the significance of this study lies in the fact that the assumptions that underly this study and the type of materials developed seem to concur with the thinking of many mathematics educators in the late nineteen sixties. Essentially what the study suggested was that an appropriate introduction to deductive mathematics could consist of a study of the basic principles of formal logic followed by the study of a relatively small number of theorems compared to the rather extensive body of theorems studied in traditional Euclidean geometry. The study suffers, however, from the fact that there was no objective assessment of the extent to which the materials developed are useful as a viable alternative to introduce and teach students an understanding of the nature and role of mathematical proof.

Two other related studies are reported by Baker (1969) and Byham (1969). Baker attempted to produce a textbook with proof as its central theme for use at the high school level. However, his proposed textbook does not appear to contain anything significantly different from materials already available. Byham investigated how the concept of indirect proof is treated in current textbooks. Specifically, he compared the presentation of indirect proofs in modern curricula with that of older textbooks and even with how the topic was presented by Euclid himself. However, no new insights for the teaching and learning of this seemingly difficult topic were evident in this study.

In this chapter, a review of those studies which were considered to be most closely related to this study has been presented and, in general, an attempt has been made to indicate the nature of some of the more significant studies dealing with the teaching, learning and assessment of concepts related to mathematical proof.

CHAPTER III

RESEARCH PROCEDURES

I. THE PILOT STUDY

The aim of this study was to identify and categorize some of the subjective thinking processes used by eleventh grade high school students in attempting to justify a variety of mathematical generalizations and conclusions presented to them. The first task, therefore, was to devise a means of presenting to students a variety of conclusions to which they could meaningfully react and which would elicit responses indicative of their subjective thinking processes. In order to determine an effective and efficient method for obtaining the desired information, a pilot study was undertaken. Since the information obtained from the pilot study was of considerable importance, the procedures used in the pilot study will be reported in some detail.

It was decided initially to construct items which could be administered to students on a class basis during regular school hours, and to which students could respond in writing. Hence a number of items were constructed in each of which the basic idea was to describe or present a mathematical situation, generalization or conclusion which in effect required each student to either accept or reject the stated conclusion with supporting reasons. Also, in selecting the content and wording of each item, it was hoped that most students could respond to the items in a meaningful way without having to rely upon specific knowledge that may have been acquired in, for example, the study of plane geometry. In particular, it was considered important that students not perceive the items as a test of their knowledge but rather that they respond in a spontaneous manner using the first thoughts that came to mind, thus maximizing the chances of obtaining student responses indicative of their subjective thought processes.

A panel of four judges were asked to evaluate the original set of items with respect to content, wording and overall appropriateness. This assessment produced a number of constructive suggestions which were incorporated in a revised set of items. These items were used to construct six pilot instruments which were administered to students in six different grade eleven mathematics classes of the Edmonton Public and Separate school systems. The instruments were administered during the first thirty minutes of a regularly scheduled mathematics class containing approximately thirty students.

Each pilot instrument, with one exception consisted of four distinct items. Each item was presented at the top of a page with space at the bottom for the student to respond in writing.

Some examples of the type of items used in the pilot study were:

1. Fill in the blank:

   (i) n is a factor of $a + b$

   (ii) Is n always a factor of a and of b? _____

   Prove it.

2. Fill in the blank:

(i) n is a factor of a + b

(ii) Is n always a factor of a and of b?

Convince a friend who disagrees with your last answer.

3. Note the following:

$8^2 - 7^2 = 15$ $3^2 - 2^2 = 5$

$5^2 - 4^2 = 9$ $9^2 - 8^2 = 17$

$11^2 - 10^2 = 21$ $14^2 - 13^2 = 27$

(i) What conclusion can you make about differences like those above? Give examples if you like.

(ii) Do you think your conclusion in (i) is always true? Why?

4. ABCD is a rectangle. E is the mid-point of AB.

(i) The area of triangle ECD is what fractional part of the area of the whole rectangle?

(ii) If your friend disagrees with you, how are you going to convince him.

5. Tom says, "Hey Joe, I've found a formula that will always produce prime numbers for me. It is $n^2 + n + 11$. When $n = 1$, then its value is 13. For $n = 2$, the value is 17, and for $n = 3$, the value is 23. It just keeps giving me prime numbers." Joe replies, "What about $n = 11$, then $n^2 + n + 11$ equals 143 which is not a prime number because 143 equals 13 times 11." Tom replied, "Well, I'm still going to say that $n^2 + n + 11$ is a formula that always produces prime numbers when n is a positive

integer."

(i) Whose side would you be on in this argument?

(ii) Why?

6. Joe and Tom are discussing the inequality $20 - n > 2n - 1$.

Joe: "It is easy to show that this inequality is true when n is assigned the values 1, 2, 3 and 4."

Tom: "It's my guess that the inequality is true when n is assigned any natural number."

Joe: "I don't agree with you Tom."

If you were Tom, what would you reply?

The objectives in administering the pilot instruments were to determine (a) whether or not the kind of items constructed would elicit responses from students which could be analyzed to interpret the subjective thinking processes used by students; (b) whether or not the wording and format of the items made any difference in the type of responses students gave; (c) the extent to which the instruments could be administered successfully under group conditions and (d) the amount of time needed by students to comfortably respond to the various items.

Excellent cooperation was received from the selected schools in conducting the pilot investigations. Teachers willingly agreed to permit their classes to participate after a brief explanation of the nature of the project.

The following procedures were used in administering the pilot instruments. Firstly, students were not forewarned that they would be participating in the pilot study. The teacher simply introduced the investigator at the beginning of a regularly scheduled mathematics class stating that during the next thirty minutes, the investigator wished them to respond to a few mathematical questions and that their responses would assist the investigator in attempting to devise a better mathematics curriculum. The teacher requested the full cooperation of the students and then left the room.

The investigator then distributed a copy of the instrument to each student. He asked them to read the instructions on the first page which were:

> On the following pages you will be asked some questions related to Mathematics. These questions are in no way intended as a test or examination. You are requested to answer each question as best you can using the first thoughts that enter your mind.

The investigator repeated verbally that the questions were not to be considered as a test of their mathematical knowledge since for most of the questions there were no right or wrong answers. The students were requested to read each item carefully and to raise their hands for assistance if there was anything that they did not understand. The importance of responding to each item as quickly as possible with the first thoughts that came to mind was emphasized. The students were given as much time as they needed to complete their responses, which varied from twenty to forty minutes.

Surprisingly, very few questions were asked by students and no difficulties were encountered in administering the instruments. At the end of the sessions, the students were thanked for their cooperation and all copies of the instrument were returned to the investigator.

An analysis of the responses to each of the items was then undertaken. Firstly, the responses of all students responding to each item were summarized. This summary was used to construct a set of mutually disjoint response categories for each item. It is to be noted that the resulting set of response categories for each item were based entirely upon the actual responses students gave and were not pre-determined. In fact, the process of analyzing the responses and devising appropriate categories to correspond to the various kinds of responses obtained for each item was a challenging and time consuming activity.

Three judges were then requested to independently place each response to a given item in exactly one of the response categories for that item. This procedure was used initially to determine the adequacy of the categories and later to determine the realibility of assigning each response to an appropriate category.

As a result of this analysis, it was determined that a relatively small number of response categories for each item was adequate to account for all of the responses given for that item. Also, the procedure of constructing response categories and assigning student responses to appropriate categories was considered feasible with larger numbers of students than that used in the pilot study. More importantly, however, these procedures were considered to be an effective way of summarizing student response patterns for further analysis and also of classifying the kinds of thinking exhibited by students in responding to many of the items.

It was also determined from the pilot study that the nature of student responses to some of the items varied greatly depending upon the format or way in which the items were worded and presented. As an illustration, consider the following item:

Fill in the blanks:

(i) Three is a factor of 24.

(ii) Three is a factor of 33.

(iii) Is three a factor of 24 + 33? _____

(i) Six is a factor of 72.

(ii) Six is a factor of 132.

(iii) Is six a factor of 72 + 132? _____

(i) n is a factor of a.

(ii) n is a factor of b.

(iii) Is n a factor of a + b? _____

This item was included on two versions of the pilot instrument. Space was provided beneath the question "Is n a factor of a + b?" for the student to justify his conclusion. On one of the pilot instruments, the words "prove it" appeared. On the other, the words "How would you convince a friend that you are right if he disagrees with your last answer?" Several other items were presented on different versions of the pilot instruments in exactly the same way.

Not surprisingly, perhaps, the words "How would you convince a friend, etc.?" produced many more meaningful responses than did the words "prove it!". In fact almost all of the students responded to the words "How would you convince a friend, etc.?" whereas only half the students responded to the words "prove it!". Also, the responses of both groups of students were in several respects quite different in nature. It is possible that many students did not understand what was meant by the words "prove it!". It appeared, however, that most students reacted to the words "prove it" in a very formal manner and perhaps responded by relying to a large extent upon specific knowledge gained from their study of mathematics. This may explain why almost half of the students did not respond at all. Many of the responses that were given appeared to be based upon what the students perceived to be expected by someone in authority rather than upon their immediate subjective thoughts.

On the other hand, students appeared to respond to the question "How would you convince a friend that you are right if he disagrees with you?" in a much more informal and open manner. Not only did a much higher proportion of the students respond to each item, but their responses in this case were considered to be much more indicative of their subjective thought processes than was evident in the "prove it" responses.

The importance of item format and wording was also indicated in another context. Several of the items used in the pilot study presented a mathematical situation from which students were asked to draw their own conclusions and then provide justification for it. However, it was found that in such cases, more often than not, students either did not respond at all or if they did, their conclusion was either irrelevant or meaningless with no attempt to justify it. Hence, in terms of determining the kinds of thinking students use in attempting to justify various mathematical conclusions and also in maximizing the number of meaningful student responses, it was considered necessary that students be asked to respond to items which contained an explicitely stated conclusion.

As a result of these findings, it was decided that a more extensive study could be undertaken to seek answers to the specific questions posed in the statement of the problem. It was felt that the kind of information sought could be obtained by administering carefully selected items designed with the findings of the pilot study in mind.

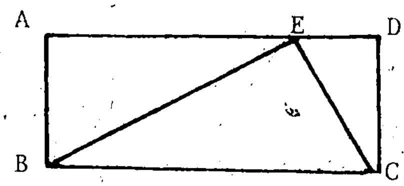
II. INSTRUMENTATION

As a result of the pilot investigation and in particular the observations relative to item format, some of the items constructed initially were considered inappropriate, others were modified and new items were constructed. It was decided to write all items using a similar format. As a result, most of the items were written in the form of a dialogue between two hypothetical mathematics students. The dialogue was used as a means of presenting to students a mathematical situation in a manner that, it was hoped, they could easily relate to and understand. The dialogue was normally written in such a way that one of the hypothetical students presents or states a conclusion or generalization based, in some instances, upon empirical evidence. The second hypothetical student, however, either does not agree with the stated conclusion or else does not readily accept that the conclusion is a true statement and hence requests that the conclusion be justified.

The students to whom the items were administered were normally requested to indicate which of the two hypothetical students they agree with or which in their opinion is mathematically correct and to justify the position they take. In other words, the students who were administered the instrument were normally given a choice of either accepting or rejecting the stated conclusion and were requested to provide their reasons for doing so. It was hoped that the students would respond freely to the dialogue in the same manner as those who were asked to convince a friend in the pilot study.

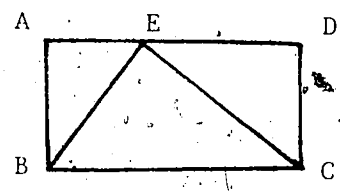
Twelve items, referred to as situations, were selected and administered to students. The items selected are presented below together with a brief rationale.

SITUATION ONE



Joe's Diagram

Joe: "In my diagram, the altitude of triangle BEC is $\overline{CD}$. Therefore the area of triangle BEC is $(\overline{BC}.\overline{CD})/2$. But the area of rectangle ABCD is $\overline{BC}.\overline{CD}$. Therefore, the area of triangle BEC is ½ the area of rectangle ABCD. The same is true in your diagram Tom."



Tom's Diagram

Tom: "I disagree with you Joe. My diagram is different from yours and I cannot say that something is true in my diagram just because it is true in yours."

Joe: "Of couse you can. Because your diagram is not the same as mine doesn't matter."

Tom: "I don't agree. Our two diagrams are different and what's true in your diagram Joe, has nothing to do with what is true in my diagram."

QUESTION: Whose side would you be on in the above discussion?

Joe's ________ Tom's __________ or Neither ________

WHY?

The main purpose of situation one was to determine whether or not students understand the application of the generalizat principle at least in geometric proofs. Symbolically, the generalization principle, as used in elementary mathematics, can be stated as follows: If P(x) is a statement function which is proven for any arbitrary but fixed value of the variable x belonging to some domain D, then P(x) is proven for all x in D.

Rosskopf and Exner (1955) illustrate the use of the generalization principle in proving a trigonometric identity such as $\text{Sec}^2 x = 1 + \text{Tan}^2 x$:

> We choose some unknown or representative, but fixed, value of x and show that the equality is satisfied. Our next step is to state, either to ourselves or in words, since the equality is satisfied for a representative value of x, then it must be satisfied for all values of x (p.294).

Hence the substance of the principle is that one can substitute a variable for a fixed unknown in a proved statement.

As an illustration of its use in plane geometry, consider the theorem "If a triangle is isosceles, then the angles opposite the equal sides are equal." This is obviously a statement about all triangles that are isosceles. To prove this theorem, however, the normal procedure is to start with an arbitrary triangle "ABC" in which "AC = BC" and deduce that "angle A = angle B". In the proof, "ABC" is considered to be an arbitrary but fixed isosceles triangle. It would be difficult if not impossible to permit "ABC" to vary from step to step in the proof. However, while "ABC" is considered fixed during the proof, it is perfectly arbitrary and hence no further proof is required to show that the theorem is true for all isosceles triangles. In other words, the same proof will apply regardless of which isosceles triangle "ABC" is considered to be.

Since the proofs of most theorems in plane geometry are presented with reference to a fixed geometric figure, in order for students to fully understand the nature of such proofs, it is necessary that they understand the generalization principle. The same argument applies in many non-geometric proofs as well. However, it is not clear that students actually do acquire the necessary understanding nor is it clear that teachers emphasize the importance of the generalization principle either in their teaching or their testing.

SITUATION TWO

Joe has observed the following interesting pattern:

$4 - 1 = 3$ is divisible by 3. $4^4 - 1 = 255$ is divisible by 3.
$4^2 - 1 = 15$ is divisible by 3. $4^5 - 1 = 1023$ is divisible by 3.
$4^3 - 1 = 63$ is divisible by 3. $4^6 - 1 = 4095$ is divisible by 3.

Joe borrowed a calculator and found out that $4^n - 1$ is divisible by 3 regardless of what value of n he tried. Therefore he came to the following conclusion:

> $4^n - 1$ is divisible by 3 for all positive whole numbers n.

While Joe was working on this problem, Tom walked into the room. Tom looked at the conclusion and immediately stated that he was not convinced that Joe's conclusion was always true. Tom felt that while the conclusion was true for n = 1, 2, 3, 4, 5 and 6, etc., this did not rule out the possibility of there being some number for which the conclusion was not true.

But Joe disagreed with Tom. Whatever value of n he had tried on the calculator confirmed the truth of his conclusion and therefore as far as Joe was concerned, it was always true.

QUESTIONS: (a) Whose side are you on?

Joe's ____ or Tom's ____.

Why?

The conclusion presented in the above item is an inductive generalization of the numerical observations. It is a valid generalization of the specific numeric examples presented but without proof it is not necessarily a true statement. Hence this item was intended to assess the extent to which students either accept the empirical evidence supporting the generalization as constituting proof of the generalization or alternatively see the need for a deductive proof of the stated conclusion. In other words, what type of reasoning is considered adequate by students to justify the stated conclusion?

SITUATION THREE

Joe and Tom are discussing factors:

Joe: "Since 3 is a factor of both 24 and 33, therefore 3 is a factor of 24 + 33."

Tom: "That's obvious and you can say even more than that Joe. If 3 is a factor of any two different numbers, then 3 is a factor of the sum of those two numbers."

Joe: "Well, if you put it that way, we can go even further and say that if any integer n is a factor of two different numbers, then n is a factor of the sum of those two numbers."

Tom: "Not too fast, Joe. If n is a factor of some number p and n is a factor of another number q, can we always conclude that n is a factor of their sum $p + q$?"

Joe: "Of course."

Tom: "I'm not convinced. In fact I don't think that your conclusion is always true."

QUESTIONS: (a) Who do you agree with? Joe ____ or Tom ____.

(b) If you agree with Joe, how would you convince Tom?

(c) If you agree with Tom, how would you convince Joe?

Situation three can be considered similar both in content and purpose to situation two. However, there was one major difference. In situation two, it was not expected that students would attempt to provide a formal proof of the conclusion presented there since to do so would require a knowledge of either the binomial theorem or the principle of mathematical induction. In fact it is doubtful whether many of the students were sufficiently knowledgeable to provide such a proof and this was one of the reasons for selecting the type of conclusion presented in situation two. The conslusion presented in situation three, however, was considered sufficiently elementary that most high school students should have no difficulty in providing a formal proof to justify it. Hence the objective was to determine whether or not the responses of students to situation three were any different in nature from the responses given in situation two.

SITUATION FOUR

Joe has written a number of interesting equations, some of which are as follows:

$8^2 - 7^2 = 15$	$3^2 - 2^2 = 5$
$5^2 - 4^2 = 9$	$9^2 - 8^2 = 17$
$11^2 - 10^2 = 21$	$14^2 - 13^2 = 27$

From these equations, Joe concludes that for all integers a and b,

$$a^2 - b^2 = a + b$$

QUESTIONS:

(a) Do you think that Joe's conclusion is true ______ or false ______ ?

(b) If you think that Joe's conclusion is true, then state why.

(c) If you think that Joe's conclusion is false, then state why.

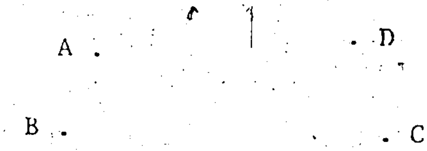
(d) If you think that Joe's conclusion is false, then what do you think might be a correct conclusion?

Situation four, though similar in some respects to situations two and three, differs in that the stated conclusion is incorrect. Hence the objective of this item was to investigate the extent to which students will either recognize this fact and attempt to provide a counter-example or accept the stated conclusion based upon the six numeric examples given. While it was intended that the items not force students to rely upon specific knowledge, in this case it was recognized that the algebraic identity $a^2 - b^2 = (a-b)(a+b)$ would be familiar to most students. Hence, to some extent, the acceptance or rejection of the conclusion presented in situation four may depend upon how well students can transfer their knowledge.

SITUATION FIVE

Joe and Tom are discussing parallel lines.

Joe: "Suppose there are only four distinct points on this sheet of paper instead of an infinite number."



Tom: "So you are imagining that the four points A, B, C and D above are the only points on this sheet of paper."

Joe: "Right. Now, since two distinct points determine a unique line, these four points determine 6 distinct lines. Each of these six lines contains only two points. Do you see what the six lines are Tom?"

Tom: "Yes."

Joe: "Now remember that if two lines have no point in common, then they are parallel."

Tom: "I agree."

Joe: "I claim that there are 3 pairs of parallel lines determined by the four given points."

Tom: "That's nonsense, Joe. None of the lines determined by the four given points can possibly be parallel."

Joe: "So you would say that the line determined by the points A and B (for example) is not parallel to the line determined by the points C and D."

Tom: "Of course these lines are not parallel."

Joe: "O.K. if these two lines are not parallel, then they must intersect in some point. Since there are only four points on the sheet of paper, the lines must intersect in either A, B, C or D. But clearly the line determined by A and B does not intersect the line determined by C and D in either C or D and vice versa. Therefore, these two lines must be parallel."

Tom: "I don't care what you say. The four given points do not determine any parallel lines."

Joe: "But I've shown otherwise Tom, and in fact I can show that there are three pairs of parallel lines determined by the four given points."

QUESTIONS: (a) Whose side would you be on in the above discussion? Joe's ____ Tom's ____. Neither ____.

(b) Why?

In their study of mathematics, students are often required to accept and reason from axioms and definitions that are not only unfamiliar but perhaps are contrary to their intuition or previous experience. It is normally assumed that students at the high school level are mature enough in the Piagetian sense, to reason, for example, from an assumption stated for the sake of argument or to understand the importance of definitions in the context of mathematical arguments. The simple non-Euclidean geometry presented in situation five introduces students to hypotheses and definitions which are perhaps contrary to their intuition, in an attempt to assess the extent to which students are willing to accept these statements and understand their use in the context of the argument presented to justify the stated conclusion.

SITUATION SIX

Joe: "Tom, did you know that there is a way to show that $1 \neq 0$?

Tom: "No, how?"

Joe: "Well, it goes like this:

1. Suppose that $1 = 0$

2. Let a and b be any numbers such that $a \neq b$.

3. Since $1 = 0$, therefore $a = a.1 = a.0 = 0$.

4. Similarly, since $1 = 0$, $b = b.1 = b.0 = 0$.

5. Therefore $a = b$.

6. But $a = b$ is false and so $1 \neq 0$."

Tom: "But you started out by supposing that $1 = 0$. How can you say something that isn't true. To me, it doesn't make sense to suppose that $1 = 0$ in order to show just the opposite."

Joe: "I don't agree with you, Tom."

QUESTIONS: (a) Whose side would you be on in the above discussion? Joe's ______ or Tom's ______

(b) If you are on Joe's side, how would you show Tom why you disagreed with him?

(c) If you are on Tom's side, how would you show Joe why you disagree with him?

The main objective of situation six was to determine whether or not students understand the logic of an indirect argument. Specifically, to what extent do students understand and accept a chain of reasoning which starts by assuming the negation of the statement to be proven. Alternatively, to what extent do students reject the given argument because it starts with an assumption known to be false.

SITUATION SEVEN

Joe has shown that the following statement is true for all real numbers x and y.

STATEMENT: Suppose $x.y = 0$. If $y \neq 0$, then $x = 0$.

Joe's argument is as follows:

(a) Given $x.y = 0$ and $y \neq 0$
(b) To show that $x = 0$

1. Either $x = 0$ or $x \neq 0$

2. For the sake of argument, suppose that $x = 3$.

3. Since $x.y = 0$, therefore $3.y = 0$ and therefore $1/3.(3.y) = 0$.

4. But, since $1/3 . (3) = 1$, this means that $y = 0$.

5. But $y = 0$ is false and so the supposition that $x \neq 0$ must be false.

6. Therefore $x = 0$.

Although Joe's argument is correct, Tom does not understand it.

QUESTIONS: (a) How would you explain Step 2 of Joe's argument to Tom?

(b) How would you explain Step 5 of Joe's argument to Tom?

(c) How would you explain Step 6 of Joe's argument to Tom?

Situation seven is also concerned with the indirect form of proof. In contrast to situation six, however, the objective of this item was to assess the extent to which students understand the important steps in an indirect argument. In particular, will students more readily accept the statement "suppose that x = 3" in step two of situation seven than the statement "suppose that 1 = 0" in step one of situation six. Also, to what extent are students convinced that the statement in situation seven is proven as a result of the contradiction deduced in step five of the argument.

SITUATION EIGHT

Joe: "Tom, in situation 7 on the previous page, I showed that for all real numbers, x and y, if $x.y = 0$ and if $y \neq 0$, then $x = 0$.

Tom: "Yes, I see. However, I feel that your argument is completely unnecessary. Look, everybody knows that if $x.y = 0$ and $y \neq 0$, then x must be equal to zero. There is no need to show it."

Joe: "I agree that everybody knows that this proposition it true, but I disagree that my argument is unnecessary."

Tom: "Look, if $3x = 0$, then $x = 0$; if $7x = 0$, then $x = 0$, and so on. You don't have to give me any argument to show me that the proposition is true."

QUESTIONS: (a) Whose side would you be on in the above discussion? Joe's _____ or Tom's _____

(b) Why?

The purpose of this item was to investigate the extent to which students see a need to prove a proposition which they may perceive as being intuitively obvious.

SITUATION NINE

STATEMENT A: Let f be any factor of some number n. If n is an odd number, then f is an odd number.

Joe: "I think that Statement A is true Tom."

Tom: "Let me see. If n = 21, then the factors of n are 1, 3, 7 and 21. n is odd and all of its factors are odd. If n = 45, then the factors of n are 1, 3, 5, 9, 15 and 45. Again n is odd and all of its factors are odd. So Statement A seems to be true Joe."

Joe: "I also think that Statement B is true Tom."

STATEMENT B: Let f be any factor of some number n. If f is an even number, then n is an even number.

Tom: "Why?"

Joe: "1. Since f is a factor of n, therefore $n = f.m$, where m is some integer.

2. If f is even, then $f = 2.k$, where k is some integer.

3. Therefore $n = f.m = (2.k).m = 2.(k.m)$.

4. Therefore 2 is a factor of n and n is even.

Tom: "But Joe, this only shows that Statement B is true. It doesn't show that Statement A is true."

Joe: "Yes it does."

Tom: "I don't agree. Statement B has nothing to do with Statement A."

QUESTIONS: (a) Whose side would you be on in the above discussion? Joe's _____ or Tom's _____?

(b) Why?

The purpose of this item was to determine whether or not students recognize that a mathematical statement and its contrapositive statement are logically equivalent.

SITUATION TEN

Joe and Tom are discussing prime numbers. Recall that a prime number is a positive whole number other than one which is divisible only by one and itself.

Joe: "I've been trying to find a formula which will always give me a prime number and I've finally succeeded Tom."

Tom: "What is your formula, Joe?"

Joe: "$n^2 - n + 17$.
When $n = 1$, $n^2 - n + 17 = 1^2 - 1 + 17 = 17$.

When $n = 2$, $n^2 - n + 17 = 2^2 - 2 + 17 = 19$

When $n = 3$, $n^2 - n + 17 = 9 - 3 + 17 = 23$.

It just keeps giving me prime numbers."

Tom: "What about when $n = 17$? Then $n^2 - n + 17 = 17^2 - 17 + 17 = 17^2$."

Joe: "Well, that's only one exception and we can ignore that."

QUESTIONS: (a) Whose side would you be on in the above discussion?
Joe's ______ or Tom's ______ ?

(b) Why?

The purpose of situation ten was to determine the extent to which students recognize that a single counter-example to a stated proposition is sufficient to disprove the proposition.

SITUATION ELEVEN

Suppose you are given as true the following facts:

1) Tokyo is larger than Los Angeles.
2) Toronto is smaller than Los Angeles.
3) Toronto is larger than Edmonton.

Joe and Tom wish to show that Edmonton is smaller than Tokyo, using the above facts:

Joe argues as follows: Since Toronto is larger than Edmonton, therefore Edmonton is smaller than Toronto. Since Edmonton is smaller than Toronto and Toronto is smaller than Los Angeles, therefore, Edmonton is smaller than Los Angeles. Since Tokyo is larger than Los Angeles, therefore Los Angeles is smaller than Tokyo. Since Edmonton is smaller than Los Angeles and since Los Angeles is smaller than Tokyo, therefore, Edmonton is smaller than Tokyo.

Tom argues as follows: Either Edmonton is smaller than Tokyo or it is larger than Tokyo. Suppose Edmonton is larger than Tokyo. Then since Tokyo is larger than Los Angeles, therefore Edmonton would be larger than Los Angeles. Since Los Angeles is larger than Toronto, therefore Edmonton would be larger than Toronto. But this would contradict the fact that Toronto is larger than Edmonton. Therefore, the original hypothesis that Edmonton is larger than Tokyo must be false. Therefore, Edmonton is smaller than Tokyo.

Both Joe and Tom have shown that Edmonton is smaller than Tokyo.

QUESTIONS: (a) Which argument would you have used? Joe's ____ or Tom's ____

(b) Why would you have used this argument?

The objective of this item was to present both a direct and an indirect argument based upon given hypotheses and to determine the extent to which students prefer one argument over the other and why.

SITUATION TWELVE

Joe: "All odd numbers greater than 627 are prime numbers."

Tom: "Show me."

Joe: "1. Suppose x is a prime number greater than 627.

2. It follows from the definition of a prime number that the only exact divisors of x are 1 and x itself.

3. Therefore 2 cannot be an exact divisor of x.

4. Therefore x cannot be even.

5. Therefore x is odd. So all odd numbers greater than 627 are prime numbers."

QUESTION: How would you reply if you were Tom?

The purpose of this item was to determine whether or not students distinguish between a mathematical statement and its converse and recognize that the two are not logically equivalent. Specifically, will students realize that the argument presented in situation twelve shows that the converse of the stated proposition is true whereas the proposition itself is not true.

III. THE SAMPLE

The instrument was administered to two hundred fifty-five grade eleven students in ten randomly selected mathematics classes from nine senior high schools of the Edmonton Public and Separate School systems. All students were enrolled in an academic program designed to prepare them for the study of post-secondary mathematics. Table I gives a distribution of the sample by sex, age and class.

Students at the grade eleven level were selected because the mathematics program taken by these students at both the tenth and eleventh grade levels were similar to programs offered at that level in many schools across Canada, whereas many of the grade twelve programs differ significantly.

Since the instrument was administered near the end of the school year, each teacher was asked to rate each student on a low, medium, high scale with respect to mathematics achievement. The results of this rating are summarized in Table II.

IV TESTING PROCEDURES

Initially, a request was made to both the Edmonton Public and Separate School Boards to administer the instrument in randomly selected grade eleven mathematics classes. As a result, six composite high schools in the public system and three in the separate system agreed to cooperate. The principal of each of these schools was contacted and a meeting arranged with each of the cooperating teachers. At each meeting, the nature of the project was explained and it was proposed that the instrument be administered to all students of a randomly selected class during a regularly scheduled mathematics period of approximately eighty minutes duration. As a result, times were agreed

TABLE I

THE SAMPLE

GROUP	SCHOOL	MALES	FEMALES	MEAN AGE (MONTHS)	NUMBER
1	A	19	10	189.8	29
2	B	13	11	190.0	24
3	C	23	10	188.5	33
4	C	18	7	188.9	25
5	D	12	11	189.5	23
6	E	17	14	189.4	31
7	F	7	12	191.9	19
8	G	5	6	192.3	11
9	H	15	17	188.1	32
10	J	21	7	188.9	28
	TOTAL	150	105	189.4	255
	PERCENT	58.8	41.2		

TABLE II

MATHEMATICS ACHIEVEMENT LEVELS BY GROUP

GROUP	LOW	MEDIUM	HIGH	TOTAL
1	5	18	6	29
2	4	15	5	24
3	6	21	6	33
4	3	15	7	25
5	4	14	5	23
6	7	16	8	31
7	2	15	2	19
8	3	5	3	11
9	5	19	8	32
10	5	19	4	28
TOTAL	44	157	54	255
PERCENT	17.2	61.6	21.2	

upon for the administration of the instrument and teachers were requested not to inform the students in advance that they would be participating in this project.

The procedures followed during the administration of the instruments were as follows: At the beginning of each class, the cooperating teacher introduced the investigator and requested the students to cooperate fully. Once the teacher had left the classroom, the investigator explained his reason for being there as follows:

> "My purpose in being here today is to present you with some mathematical situations which you can read and think about. In general, each situation consists of a dialogue between two hypothetical mathematics students called Joe and Tom. After reading the discussion in each situation, you will find several questions related to that particular situation. I wish to emphasize that these questions are not to be considered as some sort of test or examination. In fact, your responses to these items cannot be classified as right or wrong. The whole idea is to find out how you relate to the situations presented and to let you present your own ideas and thoughts, not those of someone else. As a result, you will be assisting me in my attempt to devise and propose some alternative ways that mathematics can be presented and studied at the high school level."

A copy of the instrument was then distributed to each student. The first page of the instrument contained the following instructions:

> On the following pages you will be asked some questions related to mathematics. These questions are in no way intended as a test or examination. You are requested to read the discussion on each page carefully and answer each question in the space provided using the first thoughts or opinions that enter your mind.

Students were requested to provide the following information:

Name of School ____________________
Student's name ____________________
Male __________ Female __________
Date of birth ____________________
Teacher's name ____________________

It should be noted that on three of the pilot instruments, students were requested to provide similar information, while on the other three they were not. It was determined, however, that the response patterns of students who were asked to write their name on the instrument were no different from those who were not requested to do so.

After the instruments were distributed, students were asked to read the instructions and provide the requested information. They were also asked to raise their hands for assistance if there was anything they did not understand or when they had finished responding. As was the case in the pilot study, very few questions were asked by students indicating that most students felt comfortable with the instrument and understood what was expected of them. Also students appeared to have ample time to complete responses to each of the items administered since most did not require the full eighty minutes.

It should be noted here that as a result of the pilot study, it was determined that ten minutes for each item was sufficient time for students to respond. Hence, students were requested to respond to eight of the twelve situations. Situations one to seven were administered to all students and situations eight to twelve were administered at random so that approximately twenty percent of the sample responded to each of the last five items.

V. THE CATEGORIZATION SCHEME

As was the case in the pilot study, the responses of each student to each of the items were analyzed by three independent judges. Initially, one of the judges, namely the investigator, examined the responses to each of the items. As a result of this examination and the experience gained in the pilot study, a proposed classification of responses was devised. This classification consisted of a set of mutually disjoint response categories for each item. As with the pilot study, the response categories were not pre-determined but were based upon the kinds of responses that students gave to each of the items. When the proposed classification scheme was completed, the three judges were asked to independently place each response to a given item in exactly one of the response categories for that item. This procedure was used initially to determine the adequacy of the proposed categorization scheme. As a result modifications to the proposed scheme were suggested. The modifications consisted mainly of adding new categories for some of the items or defining existing categories more sharply.

The resulting categorization scheme which was used to summarize student response patterns and also to classify the kinds of thinking exhibited by students, is presented below for each situation.

SITUATION ONE

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	Joe's conclusion that the area of triangle BEC is one-half the area of rectangle ABCD is rejected and reasons given supporting the rejection.
3.	Joe's conclusion is accepted but it does not apply to both diagrams for reasons such as: (a) the two diagrams are not the same, or have different measurements or are not congruent; or (b) the conclusion is not generalizable to all other such diagrams. (All responses in this category indicate a lack of understanding of the generalization principle.)
4.	Joe's conclusion is accepted but proof is required in order to show that it applies to both diagrams. (Responses in this category include incomplete or incorrect attempts to provide such a proof.)
5.	Joe's conclusion is true for both diagrams but the reasons given make it difficult to acertain whether or not the student understands the generalization principle.
6.	Joe's initial conclusion is true for both diagrams because the area of a triangle is always one-half the base times the altitude. Included in this category are all responses which show some understanding of the generalization principle.
0.	The response does not appear to fall in any of the above categories.

SITUATION TWO

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	Joe's generalization is considered invalid on the basis of a "false" counter-example such as: 4^n-1 is not divisible by 3 when n = 0.
3.	Joe has made a correct generalization based on the examples given and this generalization is always true because: (a) Joe has given a sufficient number of examples to prove it, or (b) it is impossible for Joe to verify it for all positive integers individually, and since it is true in all the cases he tried, therefore it must be true always.
4.	Joe has made a correct generalization, but it may not be true for all positive integers because Joe has not tested it enough or tried all of the possibilities. The responses in this (and the previous) category are characterized by thinking considered to be inductive in nature.
5.	Joe has made a correct generalization, based on the evidence presented, which is always true because Tom has not provided a counter-example or any reason why it is not true. Included in this category are those responses which indicate that any proposition in mathematics is true unless or until it is proven false.

Category:	Description of Category:
6.	Joe has made a correct generalization based on the evidence presented but the generalization may not be true for all positive integers because it is possible that a counter-example exists.
7.	Joe has made a correct generalization based upon the evidence presented but this evidence does not constitute proof of Joe's generalization. Responses in this category include complete, incomplete or incorrect attempts to prove Joe's generalization.
0.	The response does not appear to fall in any of the above categories.

SITUATION THREE

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	Joe's proposition is untrue: (a) because 3 is "not" a factor of 57 as indicated in the first statement of this item; (b) because a counter-example to the converse of Joe's proposition was provided; or (c) because some other "invalid" counterexample was given.
3.	No definite position is taken with respect to the truth of Joe's proposition. Included in this category are those responses which consider and present arguments (sometimes contradictory in nature) supporting the points of view of both Joe and Tom.

Category:	Description of Category:
4.	Joe's proposition is always true and Tom can be convinced by providing him with a "sufficient" number of confirming instances of the proposition.
5.	Joe's proposition may or may not be true because not enough confirming instances of the proposition have been provided.
6.	Joe's proposition is always true because no counter-example to the proposition has been or can be provided by Tom. Tom can be convinced of this fact simply by asking him to produce at least one counter-example.
7.	Joe's proposition may not be true (a) as can be demonstrated by providing a counter-example; or (b) because it is possible that a counter-example exists.
8.	Joe's proposition is always true and Tom can be convinced by providing him with a proof of the proposition.
9.	Joe's proposition is always true and a complete, incomplete or incorrect proof provided.
0.	Responses which do not appear to fall into any of the above categories.

SITUATION FOUR

Category:	Description of Category:
1.	(a) No response (b) The response is clearly based upon meaningless or irrelevant considerations.

Category:	Description of Category:
2.	Joe has made a correct generalization based upon the evidence presented.
3.	The evidence given is insufficient to support any conclusion.
4.	Joe's conclusion is rejected but a valid reason is not given.
5.	Joe's conclusion is rejected (a) by providing an explicit counter-example; (b) by appealing to the identity $a^2-b^2 = (a-b)(a+b)$, or (c) for other valid reasons.
6.	Joe's conclusion is rejected and a correct generalization given based upon the evidence presented.
0.	The response does not appear to fall in any of the above categories.

SITUATION FIVE

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	No definite position is taken with respect to the truth of Joe's conclusion. Included in this category are those responses which consider and present arguments (often contradictory in nature) supporting the points of view of both Joe and Tom.
3.	Joe's conclusion is rejected because of an apparent refusal to accept and use the given definition of a line,

Category: Description of Category:

in this geometry. Normally, responses in this category refuse to accept any definition other than the Euclidean, thereby ignoring or refusing to accept the hypothesis of only four distinct points.

4. Joe's conclusion is partially accepted. Responses in this category, among others, accept the proposition that two pairs of parallel lines exist but not three.

5. Joe's conclusion is accepted because Tom has not provided any justification for the position he takes.

6. Joe's conclusion that there exists three pairs of parallel lines in this geometry is accepted. Responses in this category justify the truth of Joe's conclusion by stating that the argument given follows from the definitions and assumptions given.

7. Joe's conclusion is accepted and justified by showing that one or both of the other two pairs of lines are parallel.

0. Responses which do not appear to fall into any of the above categories.

SITUATION SIX

Category: Description of Category:

1. (a) No response

(b) No meaningful response.

2. Joe's argument is valid and yet it isn't. Responses in

Category:	Description of Category:
	this category agree with both Joe and Tom and present arguments for both sides without realizing the contradictory nature of such a response.
3.	Joe's argument is rejected because it is "invalid" to "suppose that 1 = 0". Responses in this category do not appear to accept the reasoning from a hypothesis which is considered to be "false".
4.	Joe's argument is rejected because of an apparent lack of understanding of the role that the hypothesis in step one plays in the suceeding argument.
5.	Joe's argument is rejected because the reasoning in steps 3, 4, and/or 5 of the argument is considered invalid.
6.	Joe's argument is rejected because of an apparent lack of understanding of what the contradiction in step 5 really means.
7.	Joe's argument is accepted but the reasons given indicate an acceptance of authority rather than an understanding of the methods used.
8.	Joe's argument is accepted and the justification given is considered to be based on a peripheral understanding of the methods used. Responses in this category include, where appropriate, those which suggest that one can show that a statement is valid by demonstrating that its negation is false or contradictory.

Category:	Description of Category:
9.	Responses which reflect a mature view of the method of indirect proof.
0.	Responses which do not appear to fall in any of the above categories.

SITUATION SEVEN

Category:	Description of Category:
1.	No response.
2.	Response displays a complete lack of understanding of the methods used.
3.	The role of the hypothesis in step 2 of the argument is not adequately explained.
4.	The hypothesis in step 2 is adequately explained but the contradiction in step 5 is not. Responses in this category include those which argue in a circle by assuming the validity of the proposition being proven in order to explain step 5 of the argument.
5.	The explanations given are judged to be based on an acceptance of authority rather than an understanding of the methods used.
6.	A mature understanding of the argument given is indicated by adequate explanations of steps 2, 5, and 6.
0.	Responses which do not appear to fall in any of the above categories.

SITUATION EIGHT

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	It is intuitively obvious that Joe's proposition is valid for all real numbers x and y and hence no more convincing evidence (proof) is required.
3.	Joe's proposition is intuitively valid for all real numbers x and y, but a proof is required. Responses in this category are judged to be based upon an acceptance of authority rather than a mature understanding of the necessity for proof.
4.	Even though Joe's proposition may be considered to be intuitively valid for all real numbers x and y, a proof is required to show its validity beyond all doubt. Responses in this category are judged to possess a mature understanding of the need for proof.
0.	Responses which do not appear to fall in any of the above categories.

SITUATION NINE

Category:	Description of Category:
1.	(a) No response. (b) No meaningful response.
2.	Statement A or B is shown to be "invalid" by providing a counter-example to its converse.

Category:	Description of Category:
3.	Response indicates that there is no relationship between statements A and B.
4.	Response indicates that there is a relationship between A and B, but the nature of this relationship is not clearly explained or the explanation given is invalid.
5.	Response indicates that statements A and B are equivalent because one is the contrapositive of the other.
0.	Responses which do not fall in any of the above categories.

SITUATION TEN

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	Joe's claim is true because it is true in the majority of instances.
3.	Joe's claim is true because $n = 17$ might be the only exception.
4.	Joe's claim is untrue because $n = 17$ is a counter-example and this is sufficient reason for the claim to be rejected.
0.	Responses which do not fall in any of the above categories.

SITUATION ELEVEN

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.

Category:	Description of Category:
2.	The direct argument is selected because it is easier to follow whereas Tom is arguing backwards.
3.	The indirect argument is selected because it is less confusing.
4.	Responses which attempt to justify both arguments.
0.	None of the above.

SITUATION TWELVE

Category:	Description of Category:
1.	(a) No response; (b) No meaningful response.
2.	Joe's proposition is true because he has proved it.
3.	Joe's proposition is true only if he can prove it for any specific number greater than 627 that is selected.
4.	Joe's proposition is rejected and a counter-example given.
5.	Joe's proposition is rejected because his argument is that of the converse of the stated proposition.
0.	None of the above.

The above categorization scheme was considered by all three judges to be adequate to account for the responses of all students to each of the items. Once this had been agreed upon, the procedure of having each judge independently place each response to a given item in exactly one of the response categories was repeated, so that the reliability of placing a given response in the appropriate category could be determined.

Cohen's coefficient of agreement for nominal scales (Cohen, 1960) was used to determine inter-judge agreement. The coefficient of agreement k was determined for the first seven situations with the following results:

SITUATION:	1	2	3	4	5	6	7
k:	0.70	0.88	0.86	0.79	0.79	0.67	0.63

It should be noted that the coefficient k is simply the proportion of chance-expected disagreements which do not occur. In other words, k is the proportion of agreement after chance agreement is removed from consideration. When obtained agreement equals chance agreement, k = 0. Greater than chance agreement leads to positive values of k, and less than chance agreement leads to negative values. The upper limit of k is +1, occuring when there is perfect agreement between the judges.

VI. THE SCORING AND ANALYSIS OF RESPONSES

The response categories for each item, given in the previous section, were numbered so that responses placed in a low numbered category were indicative of a lower level of thinking than responses placed in a higher numbered category. This attempt to order the response categories in a hierarchy was carried out so that a common scoring scheme could be devised for each item.

Consequently, the categories for each item were divided into four distinct classes which were assigned scores of 0, 1, 2 and 3 respectively. The categories assigned a score of zero were those for which no meaningful response was given. Scores of 1, 2 and 3 were assigned to categories which were considered to represent, respectively, low, medium and high level responses. The scoring scheme devised for each situation is presented in Table III.

This scoring scheme was devised in order to carry out a statistical analysis of the response patterns of students and in particular to test the hypotheses stated in Chapter I. In this respect, hypothesis one was tested using a one-way analysis of variance. Hypotheses two and three were tested using Pearson correlation coefficients. Hypothesis four was tested using a two-tailed t-test.

TABLE III

RESPONSE SCORING SCHEME

SITUATION	CATEGORY 0	1	2	3	4	5	6	7	8	9
1	0	0	0	1	2	3	3			
2	0	0	0	1	1	2	2	3		
3	0	0	0	0	1	1	2	2	3	3
4	0	0	1	2	3	3	3			
5	0	0	0	1	2	2	3	3		
6	0	0	0	1	2	2	2	3	3	3
7	0	0	0	1	2	3	3			
8	0	0	1	2	3					
9	0	0	0	1	2	3				
10	0	0	1	2	3					
11	0	0	1	2	3					
12	0	0	1	2	3	3				

CHAPTER IV

RESULTS OF THE STUDY

I. DISTRIBUTION OF RESPONSES BY CATEGORY

The distribution of student responses by category for each situation is reported in this section. In addition, in order to indicate the nature of student responses to each item, examples of typical responses assigned to each of the response categories are presented and discussed. A total of two hundred fifty five students responded to situations one through seven and approximately fifty students responded to each of the remaining five items.

A. SITUATION ONE

The distribution of responses for Situation one is presented in Table IV. As indicated, nine percent of the students did not respond at all or gave no meaningful response. Exactly nineteen students disagreed with Joe's original conclusion. A typical response assigned to category two was:

(a) Whose side would you be on in the above discussion?

Response: "Neither."

(b) Why? Response: "In order for a triangle to have half the area of a rectangle, one side of the triangle must be a diagonal of the rectangle"; or "The area of the triangle is not $1/2$ BC.DC; it is $1/2$ CE.BE. Therefore, both statements are false."

TABLE IV

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION ONE.

GROUP	CATEGORY 1	2	3	4	5	6	0	TOTAL
1	2	1	9	2	7	8	0	29
2	3	0	2	0	8	11	0	24
3	1	4	12	1	9	6	0	33
4	0	4	8	3	5	5	0	25
5	3	1	3	2	5	9	0	23
6	8	1	6	5	7	4	0	31
7	1	3	6	0	3	6	0	19
8	1	2	3	2	1	2	0	11
9	1	1	1	4	7	18	0	32
10	3	2	1	3	10	9	0	28
TOTALS	23	19	51	22	62	78	0	255
PERCENT	9.0	7.5	20.0	8.6	24.3	30.6	0	

Since the students to whom the instrument was administered were in the highest mathematics stream of each school, it was surprising to find that over sixteen percent of these students either did not respond at all or else considered the proof given in situation one to be invalid.

Twenty percent of the students responded in category three. These students appeared to accept the argument which shows that in Joe's diagram, the area of the triangle is one-half the area of the rectangle. However, they disagreed that this argument is valid in the context of Tom's diagram. A typical response was:

(a) Whose side would you be on in the above discussion?

Response: "Tom's."

(b) Why? Response: "Because you cannot assume that something is true by just one try. Just because one case is true doesn't mean that all others are."

Evidently, students responding in category three do not realize that the two diagrams presented in situation one illustrate two special cases of the same general proposition, or if they do, then they appear to be suggesting that the one instance where proof is provided does not constitute proof of the general proposition. As a result, these students were considered not to understand the essence of the generalization principle.

It should be noted, however, that the reasons given by these students for not accepting Joe's argument in the context of Tom's diagram are similar to the reasons given by students who argued that a deductive proof was necessary to prove the proposition presented in situation two. In fact, as will be discussed in the next chapter, many of the students who stated correctly that an inductive argument does not constitute formal proof of the conclusion presented in situation two, appeared to use the same kind of thinking incorrectly in responding to situation one. Evidently, many of these students did not recognize the subtle distinction between the conclusions presented in situations one and two.

Student responses assigned to category four differed slightly from those assigned to category three. Students responding in category four appeared to realize that the two diagrams presented in situation one do represent two special cases of the same general proposition. In other words, these students accepted without question the conclusion that the area of the triangle was one-half the area of the rectangle in both diagrams. However, these students also did not feel that the argument presented was sufficient to justify this conclusion, and therefore suggested explicitly that it was necessary to prove that the conclusion holds in the case of Tom's diagram. A typical response in category four was:

(a) Whose side would you be on in the above discussion?

Response: "Tom's."

(b) Why? "Joe should have to prove that the same is true in Tom's diagram. Just saying so cannot be accepted by everyone, especially those not familiar with the principles involved."

Approximately fifty-five percent of the students responded in categories five and six. The responses assigned to these categories indicated that these students did understand that the argument presented was valid for both diagrams and that no further proof is necessary. However, responses assigned to category five were less convincing than those assigned to category six. A typical response in category five was:

(a) Whose side would you be on in the above discussion?

Response: "Joe's."

(b) Why? "It doesn't matter how you draw the diagram; just be sure the letters correspond with each other; the lines don't have to be exact."

This response may be compared with a typical response assigned to category six:

(a) "Joe's."

(b) "Because Joe is dealing with the mathematics of, or the principles behind the areas of the rectangles and triangles. Diagrams are just illustrations of these ideas and cannot be considered as graphically correct in their construction in comparison to what is mathematically correct."

B. SITUATION TWO

The distribution of responses for situation two is presented in Table V. As in situation one, approximately fifteen

TABLE V

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION TWO

GROUP	CATEGORY								TOTAL
	1	2	3	4	5	6	7	0	
1	3	1	15	1	2	5	2	0	29
2	5	3	7	2	5	0	1	1	24
3	2	1	20	0	2	4	3	1	33
4	0	1	8	5	7	2	1	0	25
5	2	4	8	0	5	3	1	0	23
6	0	1	16	5	6	3	0	0	31
7	1	2	8	4	2	1	1	0	19
8	1	1	5	2	1	0	1	0	11
9	3	2	13	3	6	1	4	0	32
10	0	5	13	1	1	5	2	1	28
TOTALS	17	21	113	23	37	24	16	4	255
PERCENT	6.7	8.2	44.3	9.0	14.4	9.4	6.4	1.6	

percent of the students either did not respond at all or attempted to show that Joe's conclusion was untrue. All of the remaining students, however, did not appear to doubt the validity of Joe's generalization based upon the empirical evidence presented. However, less than ten percent of these students saw any need for a formal deductive proof of the conclusion. In fact, it was evident that most students do not realize that induction is inadequate to justify a mathematical proposition of the type presented in situation two.

More than fifty percent of the students responded in categories three and four. Those responding in category three argued that since Joe's conclusion was true in the special cases enumerated, therefore it must be true for all positive integers. Typical responses were as follows:

(a) Whose side are you on? Response: "Joe's."

(b) Why? Response: "There is little chance of there being a number that doesn't work if you get past ten." or "After a certain amount of tries, a conclusion has been reached. It would be impossible to prove every number, so you have to take it at face value that if a theorem is true for numbers one to twenty, it must be true for twenty-one."

Students responding in category four also argued inductively stating that it is impossible to prove Joe's conclusion true for all positive integers since there are an infinite number of cases to consider. Typical responses in category four were:

(a) Whose side are you on? Response: "Tom's."

(b) Why? "Because in order to prove it true, Joe would have to go through every number in the system and since the positive number system is infinite, this is not possible." or "Always is a lot of times. To be absolutely sure that it was always true, Joe would have to try every positive number. This is impossible, however, so Joe cannot say always. He may say, however, that it is reasonable to assume that it may always be true."

Students responding in category five also believed that Joe had provided adequate proof of his conclusion. However, the reasons given by these students were different from those given by the other students. They suggested that any proposition in Mathmatics is true unless or until it is proven false. As a result, they appear to be arguing that since no counter-example can be found to contradict Joe's conclusion, it therefore is automatically a true statement. A typical response in category five was:

(a) Whose side are you on? Response: "Joe's."

(b) Why? Response: "Joe had obviously put some serious thought and effort into his observations. By using both his head and calculator, he could not prove his conclusion wrong. So his conclusion is true until he or someone else proves it wrong." or "Because Tom didn't show an example that disproved Joe's conclusion."

Responses assigned to category six indicated that Joe's conclusion may not be true since, without formal proof it is possible that a counter-example exists. A typical response assigned to this category was:

(a) Whose side are you on? Response: "Tom's."

(b) Why? Response: "Although Joe's conclusion was right and all numbers he tried worked out correctly, one cannot rule out the fact that there is always the possibility that one number may not work."

Very few students indicated that the only adequate justification for Joe's proposition was some form of formal proof. In fact, only sixteen students responded in category seven. A typical response was:

(a) Whose side are you on? Response: "Tom's."

(b) Why? Response: "Joe may be right but I'd agree with Tom because in the examples tried, $4^n - 1$ is divisible by three but one cannot draw the conclusion that because it works in some cases, it will work for all. Numbers go on to infinity and I think more proof would be necessary."

C. SITUATION THREE

The distribution of responses for situation three is presented in Table VI. As was indicated in the rationale for situation three, the objective of this situation was similar to that of situation two. However, since the conclusion presented in three was not supported by the same amount of empirical evidence, and also since it was believed that most students would have no

TABLE VI

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION THREE.

GROUP	CATEGORY 1	2	3	4	5	6	7	8	9	0	TOTAL
1	4	2	3	8	0	6	1	2	3	0	29
2	5	1	4	6	0	6	0	2	0	0	24
3	2	4	1	17	0	3	1	3	1	1	33
4	1	3	0	10	0	5	2	2	2	0	25
5	4	0	2	8	1	2	1	3	1	1	23
6	5	2	2	7	2	4	4	2	3	0	31
7	0	1	0	10	0	4	4	0	0	0	19
8	0	3	3	0	0	3	0	0	0	2	11
9	0	1	1	13	1	7	3	4	2	0	32
10	4	1	0	11	0	3	3	5	1	0	28
TOTALS	25	18	16	90	4	43	19	23	13	4	255
PERCENT	9.8	7.1	6.4	35.3	1.6	17.0	7.5	9.0	5.1	1.6	

difficulty in presenting a formal proof of this conclusion, an attempt was made to determine whether or not the distribution of responses for situation three was any different from that of situation two.

In order to compare the distribution of responses of situation three with that of situation two, a four by four contingency table summarizing the distribution of scores for both situations was constructed as reported in Table VII.

TABLE VII

DISTRIBUTION OF SCORES ON SITUATIONS TWO AND THREE

		SITUATION THREE SCORES 0	1	2	3	TOTAL
SITUATION TWO SCORES:	0	12	15	8	3	38
	1	31	51	39	15	136
	2	12	16	19	13	60
	3	4	11	0	6	21
TOTAL		59	93	66	37	255

Chi-square = 18.98 with nine degrees of freedom.
Probability = 0.025

To test the null-hypothesis of independence between the scores of situations two and three, a chi-square test was performed which resulted in rejection of the null hypothesis at the 0.05 level. Hence, the distribution of scores for situation three was not significantly different from that of situation two.

An examination of Table VI shows that approximately eighteen percent of the students either did not respond or else concluded that Joe's conclusions were untrue. A number of students, surprisingly, did not agree that three was a factor of fifty-seven since it was claimed that fifty-seven is a prime number.

Students responding in category three could not decide whether Joe's conclusion was true or false. As an illustration, consider the following response:

(a) Who do you agree with? Response: "Joe and Tom."

(b) If you agree with Joe, how would you convince Tom?
Response: "I would challenge Tom to find two numbers for which Joe's conclusion does not hold."

(c) If you agree with Tom , how would you convince Joe?
Response: "I would ask Tom to find a number that does not agree with Joe's conclusion."

This response is interesting in that it incorporates two distinct lines of thinking. On the one hand, the student is suggesting in (a) that since no counter-example can be found, therefore, the conclusion must be a true statement. However, in (b), the student suggests quite correctly that since a counter-example may exist, Joe's conclusion is not necessarily ture. Approximately twenty-five percent of the students used the same kind of reasoning. However, most did so to justify that Joe's conclusion was either true or false but not both. These responses were assigned to categories six and seven and will be discussed later.

Student responses assigned to category four were identical to those assigned to category three of situation two. These students again did not realize that induction was inadequate to prove Joe's conclusion and contrary to expectations very few students saw a need or attempted to provide any kind of formal proof of the conclusion.

Students responding in category six concluded that Joe's proposition is true because no counter-example has been or can be provided by Tom. A typical response in this category was:

(a) Who do you agree with? Response: "Joe."

(b) How would you convince Tom? Response: "I would try out quite a few numbers and if Tom still wasn't convinced, I would ask him to show me one instance that wouldn't work."

It is possible that students use this kind of argument when the proposition under consideration is one which they consider to be intuitively obvious.

As in situation two, some students felt that Joe's conclusion may not be true since, without formal proof, it is possible that a counter-example exists. It was evident that these students do realize that induction is inadequate, as the following response in cagegory seven indicates:

(a) Who do you agree with? Response: "Joe."

(b) How would you convince Tom? "By using different examples.

There would, however, be the remote possibility of one number or numbers not working. Therefore, the conclusion cannot be totally proven using examples."

Only fourteen percent of the students suggested that some form of deductive proof was necessary to justify Joe's conclusion and only thirteen students attempted to give a formal proof. A number of these students obviously had the right idea but could not symbolize the proof correctly, as illustrated in this example:

(a) Who do you agree with? Response: "Joe."

(b) How would you convince Tom? Response: "I would tell him that if n was the same in both cases in p and q and if p was divisible by n and if n was again the same number and q was also divisible by n, it follows that the sum would also be divisible by n."

Contrary to expectations, at most three or four students gave what might be considered an acceptable proof of Joe's conclusion.

D. SITUATION FOUR

The distribution of responses for situation four is presented in Table VIII. Almost all students responded in a meaningful way to this item. However, approximately thirty percent of the students considered Joe's conclusion to be a true statement for which no additional proof was necessary. Typical responses assigned to category two were:

(a) Do you think that Joe's conclusion is true or false?
Response: "True."

TABLE VIII

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION FOUR

GROUP	CATEGORY 1	2	3	4	5	6	0	TOTAL
1	2	6	0	6	6	9	0	29
2	0	9	0	2	4	9	0	24
3	0	16	0	0	3	14	0	33
4	0	7	0	2	9	7	0	25
5	0	9	0	1	7	6	0	23
6	0	6	0	4	10	11	0	31
7	0	3	1	1	7	7	0	19
8	0	3	0	3	3	2	0	11
9	0	8	1	0	11	12	0	32
10	2	8	0	3	9	6	0	28
TOTALS	4	75	2	22	69	83	0	255
PERCENT	1.6	29.4	0.8	8.6	27.1	32.5		

(b) Why? "The examples shown and the examples that one could add give conclusive proof of his equation." or "From what I see of Joe's equations, all of them agree with the conclusion and I would continue to believe in it until it was proven false."

Almost seventy percent of the students rejected Joe's conclusion, some without any stated reason. Most of these students, however, provided a counter-example as illustrated in this response assigned to category five:

(a) Do you think that Joe's conclusion is true or false?

Response: "False."

(b) Why? "$4^2 - 2^2 = 16 - 4 = 12$, whereas $4 + 2 = 6$." or "It just shows that it is true when you use consecutive numbers with the smaller being subtracted from the larger."

(c) If you think that Joe's conclusion is false, then what do you think might be a correct conclusion? Response:

"$a^2 - b^2 = (a + b)^{a-b}$. If the difference between the two numbers is one, then it is true. If the difference is two, then you square the added answer; if three, you cube it and so on."

This last response to question (c) exhibits one of the few cases where an incorrect generalization was given.

Approximately thirty percent of the students rejected Joe's conclusion and, based upon the empirical evidence presented, indicated what the correct conclusion should be. Surprisingly, however, very few students stated their conclusion in the form

$a^2 - b^2 = (a - b)(a + b)$, which, it was assumed, must have been familiar to all students.

E. SITUATION FIVE

The distribution of responses for situation five is presented in Table IX. This item is different both in nature and purpose than the previous three items and the responses of students to this item are therefore not only different but also provide a great deal of additional information about how high school students react to mathematical concepts and ideas that may be unfamiliar to them.

As indicated in Table IX, less than ten percent of the students either did not respond or responded in an unintelligible manner. In response to the question "Whose side would you be on in the above discussion?", approximately seven percent of the students responded "neither." Several of the reasons given in response to the question "Why?" were as follows:

(i) "When the definition of parallel lines was made up, it referred to situations where there were an infinite number of points on a plane and for a line in any other kind of situation this definition is void."

(ii) "Parallel lines were first introduced on a plane with an infinite number of points. Therefore, Joe must, if he is seeing a plane with only four points, start over and make up a whole new set of ideas, definitions and theorems to

TABLE IX

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION FIVE

GROUP	CATEGORY								TOTAL
	1	[illegible]	3	4	5	6	7	0	
1	4	6	13	1	0	5	0	0	29
2	3	3	14	1	0	2	0	1	24
3	0	3	19	4	0	5	0	2	33
4	2	0	15	4	1	3	0	0	25
5	2	2	10	3	0	5	0	1	23
6	2	4	20	0	0	5	0	0	31
7	0	0	14	1	0	3	0	1	19
8	3	0	5	2	0	1	0	0	11
9	1	0	25	2	0	3	0	1	32
10	3	0	18	2	0	5	0	0	28
TOTALS	20	18	153	20	1	37	0	6	255
PERCENT	7.8	7.1	60.0	7.8	0.4	14.4	0	2.4	

identify with his plane."

(iii). "The definitions which they have of parallel lines refer to a true situation in which this whole sheet of paper is full of points. Thus, the lines would not be parallel. In this case, they are imagining that there are only four points but then they are using the definition from another situation requiring an infinite number of points. I don't agree with either of them for this reason."

In order to understand the conclusion presented in situation five, students must first understand and accept the hypothesis that the simple geometry under discussion contains only four distinct points. They must also understand, accept and use the fact that a line in this geometry, by definition, consists of only two distinct points. Finally, they must be willing to apply, in the context of this geometry, the definition that two lines are parallel if both have no point in common, a definition which most of these students should be familiar with from their study of Euclidean geometry.

In each of the three responses presented above, the students appear to have had difficulty in accepting and understanding the hypotheses and definitions presented, particularly the definition of parallelism. They appear to reject the notion that the definition of parallel lines can be applied in this geometry suggesting that parallel only has meaning in the context of the familiar Euclidean geometry. In fact, in response (iii) above, the student suggests that the definition of parallel only has meaning in "a true situation", implying perhaps that the definitions and hypotheses introduced in situation five are in some sense "untrue." If these students consider that the hypotheses, definitions and conclusions presented in this item are "untrue" because these ideas do not concur with their intuition or previous experience, then it would appear that these students do not understand that all of the concepts introduced and studied in mathematics have only those properties that are ascribed to them either by definition or by postulate. In this sense, it is doubtful that these students really understand the role of axioms and definitions and more generally, the nature of a mathematical system. In fact, if students are unwilling to accept the reasoning used in the simple geometry presented in situation five, then perhaps they lack a basic understanding of one of the most important aspects of modern mathematics.

Exactly sixty percent of the students responded in category three. Most of these students rejected Joe's conclusion because of an apparent refusal to accept the definition of a line in this geometry. Typical responses assigned to category three were:

(a) Whose side would you be on in the above discussion?
Response: "Tom's."

(b) Why? Response: "If Joe can say that the two lines must intersect at A, B, C or D because there are no other points but these on the paper, then Joe cannot say that points A and B can create a line. Joe is contradicting himself." or "These lines, even if they have no point in common, are not parallel because they are not an equal distance apart throughout. Also, Joe is only supposing that only four points exist, but in actual fact there are an infinite number, so the line would have no endpoints and would eventually meet."

The fact that many of these students argued that the lines in situation five could not be parallel because they were not the same distance apart in the diagram, also suggests that these students did not understand or accept the definition of a line as presented in this item. Evidently, these students continued to think of a line as consisting of an infinite number of points as in Euclidean geometry. Presumably, if the points A, B, C and D had been placed in a rectangular arrangement in the diagram, then it is possible that many of the students would have readily agreed that there were two pairs of parallel lines but obviously, they would not agree that there were three such pairs of lines. In spite of this, exactly twenty students who responded in category four did agree that there were two pairs of parallel lines but they evidently did not understand why there were three such pairs.

Approximately fourteen percent of the students responded in category six. The conclusion that three pairs of parallel lines are determined by the four distinct points was considered a true statement by these students. However, it was evident that none of these students saw any need to justify this conclusion since not a single student in the sample attempted to do so, as indicated in category seven of Table IX. It should be noted that even though no one responded in category seven, this category was included because it was expected that some of the students would formally prove that the pairs of lines AD, BC and AC, BD are parallel by using essentially the same argument given in this item to show that the lines AB and CD are parallel.

F. SITUATION SIX

The distribution of responses for situation six is presented in Table X. The responses to situation six were perhaps the most disappointing in that almost forty percent of the students either did not respond or their responses could not be meaningfully categorized. Two examples of responses assigned to category one are:

(a) Whose side would you be on in the above discussion?

Response: "Joe's."

(b) How would you show Tom why you disagreed with him? Response: "Repeat Joe's technique but substitute whole numbers instead of a and b. The a and b way is abstract and difficult to understand."

Alternatively:

(a) Whose side would you be on in the above discussion? Response: "Tom's."

(b) How would your show Joe why you disagree with him. Response: "It's simply because $1 \neq 0$ is universally true on this earth."

It seems clear that the second student responding above,

TABLE X

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION SIX

GROUP	CATEGORY 1	2	3	4	5	6	7	8	9	0	TOTAL
1	17	1	4	0	2	0	3	2	0	0	29
2	9	1	4	1	2	0	3	4	0	0	24
3	9	1	8	4	1	0	4	5	1	0	33
4	8	0	4	2	0	0	3	7	0	1	25
5	13	0	3	0	0	0	4	3	0	0	23
6	12	0	5	4	2	0	6	1	0	1	31
7	4	0	0	3	1	0	2	5	0	4	19
8	3	2	0	0	0	0	4	2	0	0	11
9	9	1	2	0	0	1	8	9	0	2	32
10	10	0	4	2	1	0	5	6	0	0	28
TOTALS	94	6	34	16	9	1	42	44	1	8	255
PERCENT	37.0	2.4	13.3	6.3	3.5	0.4	16.6	17.4	0.4	3.1	

whether he understands the argument or not, does not see any need to prove a proposition which he considers to be intuitively obvious.

Responses assigned to categories three to six inclusive, which constitute approximately one quarter of all responses, consider Joe's argument to be invalid. The majority of these students argued that since the initial statement in Joe's proof, namely "suppose that $1 = 0$", is false and therefore meaningless, this invalidates the whole argument. Typical responses assigned to category three were:

(a) Whose side would you be on in the above discussion?

Response: "Tom's."

(b) How would you show Joe why you disagree with him? Response: "To start with, one can't be equal to zero. He should use true statements all the time. You can't prove one thing and take the opposite to say another thing, to prove it is true" or "It's common sense to know that $1 \neq 0$. You can't prove something is right by using false information. You are contradicting yourself. First, you say $1 = 0$ and then you say $1 \neq 0$."

Some of the students realized that the initial assumption was made for the sake of argument but still did not understand the role of this assumption in the remainder of the argument presented in situation six. For example, a typical response in category four was:

(a) "Tom's."

(b) "If you are supposing that $1 = 0$, then you can't turn around and say that $1 \neq 0$. $1 = 0$ is something that you accept as being true."

Approximately one third of the students responded in categories seven, eight and nine. While all these students agree with Joe and presumably consider the argument presented to be valid, it is not always clear from the reasons they give, that they really understand the indirect argument presented in situation seven. Responses assigned to category seven in particular, were considered to be based upon an acceptance of authority in the sense that proofs given in textbooks are in many cases accepted and learnt by students without any real understanding of the proof itself. A typical response assigned to category seven was:

(a) Whose side would you be on in the above discussion?
Response: "Joe's."

(b) How would you show Tom why you disagree with him? Response: "One way of eliminating a theorem is to find one single case where it is false. By disposing one theorem, one is sometimes able to prove the opposite or at least that a statement is not true."

This response does not indicate that the students understand the method of indirect proof, at least in the context of the argument presented in situation six. It would appear that the student is simply using general ideas which seems appropriate based upon his previous mathematical experience.

Approximately seventeen percent of the students responded in category eight. Most of these responses simply asserted that one can demonstrate the truth of a proposition by showing that its negation implies a contradictory statement and is therefore false.

Only one student , however, displayed what was considered to be a mature understanding of the method of indirect proof based upon the responses obtained for this item. While this was unexpected, a more significant finding that emerged from these responses is the fact that most students do not appear to be willing to reason for the sake of argument from a hypothesis which they consider to be untrue. This, in fact, would appear to be one of the major difficulties students have with the method of indirect proof.

G. SITUATION SEVEN

The distribution of responses for situation seven is presented in Table XI. Although situation seven is similar in nature to situation six, it differed in that students were requested to explain specific steps in the argument presented. However, the distribution of responses for both items was very similar. Almost forty percent of the responses to situation seven indicated that these students either did not understand the argument or misinterpreted the questions. For example, a typical response assigned to category two was the following:

(a) How would you explain step two of Joe's argument to Tom?

Response: "A real number is easier to work with and to understand. Any number could have been used."

(b) How would you explain step five of Joe's argument to Tom?

Response: "If $y = 0$ is false, then $y \neq 0$ is true. Looking at the statement 'if $y \neq 0$, then $x = 0$;' and if $x = 0$,

TABLE XI

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION SEVEN

GROUP	CATEGORY 1	2	3	4	5	6	0	TOTAL
1	1	16	3	7	1	1	0	29
2	0	5	1	13	1	4	0	24
3	0	5	3	18	3	5	0	33
4	2	9	1	9	0	4	0	25
5	3	12	1	2	2	3	0	23
6	4	15	1	4	2	5	0	31
7	1	6	6	3	0	3	0	19
8	0	4	2	2	0	3	0	11
9	0	12	0	8	0	12	0	32
10	2	13	4	4	1	4	0	28
TOTALS	13	97	21	70	10	44	0	255
PERCENT	5.1	38.0	8.2	27.5	3.9	17.3	0	

then $x \neq 0$ must be false."

(c) How would you explain step six of Joe's argument to Tom?

Response: "If $y \neq 0$, then $x = 0$ and if $x \neq 0$ is false, then $x = 0$ is true."

It is apparent that the students responding above does not really understand the argument but attempts to justify the crucial steps by reasoning in a circle. A number of students responded only to question (a), typically as follows:

(a) How would you explain step two of Joe's argument to Tom?

Response: "I wouldn't be able to because if the statement says $x = 0$ (at the very top), how can it equal three?"

As in situation six, it is evident that these students do not understand the logic of assuming, for the sake of argument, that the negation of a statement is true in order to show that the negation is, in fact, false. These students apparently consider the negation to be false de facto and therefore, perhaps, believe that the argument proves nothing since, in their view, the statement to be proven is assumed and used in a circular fashion.

Almost half of the student demonstrated some understanding of the role of the hypothesis in step two of the argument. However, less than twenty percent gave a satisfactory explanation of why this statement must be false. A typical response in category four was:

(a) How would you explain step five of Joe's argument to Tom?

Response: "$y \neq 0$, so y must be some number. If y is a number, then that number times x must be equal to zero. No two numbers (excluding zero) can be equal to zero so the statement $x \neq 0$ must be false since y is a number."

This student has attempted to construct his own argument but shows no appreciation of the argument given in situation seven. Almost thirty percent of the students responded in a similar manner.

Approximately eighteen percent of the students responded in category six. Whether or not these students recognized that the proof presented was in the form of an indirect argument was, in many cases, difficult to determine. However, these students did appear to have the right idea even though it was not always clearly expressed, as illustrated in the following response:

(a) How would you explain step two of Joe's argument to Tom?

Response: "Joe took an obviously false supposition in order to show Tom that it worked out wrong. Therefore, the opposite must be right."

(b) How would you explain step five of Joe's argument to Tom?

Response: "If the answer is wrong, then at some point the work must be wrong. The only place it can be is in the supposition. Therefore $x \neq 0$ is false."

(c) How would you explain step six of Joe's argument to Tom?

Response: "If an answer can only be one of two things, and it is not one of them, then it is the other."

As a general observation, it can be argued that items six and seven do not assess in a valid way student understanding of the indirect method of proof. It is possible that entirely different types of responses might have been obtained if different items or different specific arguments had been used. However, based upon the responses of students, which were fairly consistent to both of these items, there would appear to be a great deal of evidence supporting the conclusion that in excess of eighty percent of the students do not understand the logic of an indirect proof.

H. SITUATION EIGHT

The distribution of responses for situation eight are presented in Table XII. It should be noted, as pointed out earlier, that each of situations eight through twelve were administered at random to approximately twenty percent of the sample. Hence, because of the relatively small numbers of students responding from each class, it is possible that the responses are not entirely representative of student thinking.

Situation eight was administered to fifty-three students. Based upon the responses obtained, approximately forty percent of these students agree that the proposition presented in situation seven is intuitively obvious and do not see any need to formally it. A typical response assigned to category two was:

(a) Whose side would you be on in the above discussion?

Response: "Tom's."

(b) Why? Response: "Because he makes a lot of sense and he is very straight forward in his approach. He also doesn't use a lot of funny Jargon to polish up his argument."

Responses assigned to category three indicated that the stated proposition cannot be accepted as true without formal proof. However, the reasons given by these students to justify their position,

TABLE XII

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION EIGHT

GROUP	CATEGORY 1	2	3	4	0	TOTAL
1	2	2	2	0	0	6
2	0	4	1	2	0	7
3	1	1	2	2	0	6
4	0	2	0	3	0	5
5	2	1	0	0	0	3
6	2	3	0	2	0	7
7	0	3	0	2	0	5
8	0	0	1	1	0	2
9	1	0	2	3	1	7
10	0	5	0	0	0	5
TOTALS	8	21	8	15	1	53
PERCENT	15.1	39.6	15.1	28.3	1.8	

were less convincing than those given in responses assigned to category four. In this sense, it was questionable whether students were expressing their true feelings, i.e. whether they really saw a need for proof in this situation, or whether they simply responded in a manner that they considered would be more acceptable to the investigator. For example:

(a) Whose side would you be on in the above discussion?

Response: "Joe's."

(b) Why? "Because it is so obvious that the statement is true but a proof helps a person understand how that answer is really obtained. Sometimes a proof helps a person understand the situation better." or "Because Tom cannot say that everybody knows [that the proposition is true] because Tom hasn't asked everybody if they know. Tom is generalizing too much."

The reason given in this last response is significant in that the student appears to suggest quite correctly that what is intuitively obvious to one student may not be intuitively obvious to another. Hence, the need for proof is based upon the premise that the stated proposition may not be intuitively obvious to all students. It is possible that this student would not see any need to prove a proposition considered intuitively obvious by all students.

Less than thirty percent of the student responses were assigned to category four and these were considered to indicate a mature understanding of the need to prove the stated proposition. A typical response assigned to category four was:

(a) Whose side would you be on in the above discussion?

Response: "Joe's."

(b) Why? "Tom's explanation is simpler and easily understood. However, Joe proves the proposition for all values of y whereas Tom is only dealing with a few examples."

I. SITUATION NINE

The distribution of responses for situation nine is presented in Table XIII. This item was devised to assess whether or not students realize that a mathematical statement and its contrapositive are logically equivalent. It is possible, however, that this item may have contained too much information for students to assimilate in a short time. Hence, the validity of their responses is, perhaps, open to question. In any event, not a single student indicated that statements A and B were logically equivalent. In fact, the majority of students suggested that the two statements were unrelated.

Twenty percent of the students claimed that either statement A or statement B was false and tried to justify their claim with a "counter-example": For example:

(a) Whose side would you be on in the above discussion?

Response: "Tom's."

(b) Why? "Because statement B is not necessarily true. Three is a factor of twenty-four and 3 is an odd number. When Joe proved statement B, it wasn't very complete. Two is a factor but many odd numbers are factors of even numbers."

TABLE XIII

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION NINE.

GROUP	CATEGORY 1	2	3	4	5	0	TOTAL
1	0	1	3	2	0	0	6
2	0	0	3	0	0	0	3
3	0	2	3	2	0	0	7
4	1	0	2	2	0	0	5
5	0	1	3	0	0	0	4
6	2	2	0	2	0	0	6
7	0	0	2	1	0	0	3
8	1	0	0	2	0	0	3
9	0	2	4	1	0	0	7
10	0	1	2	1	0	0	4
TOTALS	4	9	22	13	0	0	48
PERCENT	8.3	18.8	45.8	27.1			

Almost fifty percent of the students responded in category three, typically as follows:

(a) "Tom's."

(b) "I think that Joe's proof only shows statement B is true. You would need a new proof to prove statement A is true."

Some students agreed with Joe but their reasons were not very explicit. For example:

(a) "Joe's."

(b) "Joe is saying the same thing, but he is just using a different example."

As stated earlier, it is possible that the example presented in this item is too difficult for most high school students to comprehend. However, based upon the limited number of responses to this item, there was no evidence to suggest that these students are familiar with the concept of logically equivalent statements.

J. SITUATION TEN

The distribution of responses for situation ten is presented in Table XIV. As was evident in the responses to several of the earlier items, the majority of students do appear to realize that a single counter-example to a mathematical proposition is sufficient to disprove that proposition. Almost eighty percent of those responding to item ten agreed that $n = 17$ is a valid counter-example to the stated proposition with reasons such as:

TABLE XIV

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION TEN

GROUP	CATEGORY 1	2	3	4	0	TOTAL
1	1	1	0	4	0	6
2	0	0	0	4	0	4
3	0	0	1	6	1	8
4	0	1	0	4	0	5
5	0	2	0	4	0	6
6	0	0	2	3	0	5
7	0	1	1	2	0	4
8	0	0	0	2	0	2
9	0	0	0	6	0	6
10	1	0	0	5	0	6
TOTALS	2	5	4	40	1	52
PERCENT	3.8	9.6	7.7	76.9	1.9	

"When one exception has been found, you cannot ignore it. There will more than likely be many more exceptions, so the statement is wrong. As soon as one exception is found, a hypothesis is invalid."

K. SITUATION ELEVEN

The distribution of responses for situation eleven is presented in Table XV. Almost eighty percent of the students selected Joe's direct argument over Tom's indirect argument. Surprisingly, the reasons given for selecting the direct argument were very consistent. For example:

(a) Which argument would you have used? Response: "Joe's."

(b) Why would you have used this argument? Response: "It's more straight forward; also, I never would have thought of the second one!" or "More fact, easier to think through; put together better; not as much supposing."

If the responses to this item indicate the student's subjective preferences, then it would appear that a direct argument is perceived as being more "natural" than an indirect argument. It may not be meaningful, however, to draw any conclusions from the student responses to this item. It is possible that the selection of one argument over the other was influenced by unknown factors such as the fact that Joe's direct argument was presented before Tom's on the sheet of paper. Unfortunately, the order of presentation was the same on all instruments.

TABLE XV

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION ELEVEN

GROUP	CATEGORY 1	2	3	4	0	TOTAL
1	1	2	2	0	0	5
2	0	3	1	1	0	5
3	0	3	3	0	0	6
4	0	4	1	0	0	5
5	0	6	0	0	0	6
6	0	6	0	0	0	6
7	0	3	1	0	0	4
8	0	2	0	0	0	2
9	0	4	1	1	0	6
10	0	7	0	0	0	7
TOTALS	1	40	9	2	0	52
PERCENT	1.9	76.9	17.3	3.8		

L. SITUATION TWELVE

The distribution of responses for situation twelve, presented in Table XVI, is similar to that for situation nine. The fact that the argument presented in this item did not prove the stated proposition, which is clearly false, but rather proved the converse of the proposition, was not recognized by anyone. Over forty percent of the students realized that the proposition was false and provided a counter-example. However, just as many students did not even do this. Some of these students recognized that the number 627 was of no particular significance but surprisingly this did not lead them to conclude that the given proof must be incorrect. For example, one student responded:

"What about the odd numbers less than 627? Why start at 627? For example, 13 is a prime number; so is 17; so is 19."

It was evident that many students were confused by this item and that none could really understand why the given proof was incorrect, even though the stated proposition was clearly false. Hence, while it may have been unrealistic to expect most students to determine what was wrong in this item, it was anticipated that some would recognize that a mathematical proposition and its converse are not logically equivalent.

TABLE XVI

DISTRIBUTION OF RESPONSES BY CATEGORY FOR SITUATION TWELVE

GROUP	CATEGORY 1	2	3	4	5	0	TOTAL
1	1	2	1	2	0	0	6
2	2	0	1	1	0	1	5
3	5	0	0	1	0	0	6
4	0	3	0	1	0	1	5
5	1	0	0	3	0	0	4
6	3	1	0	2	0	1	7
7	0	0	1	2	0	0	3
8	1	0	0	1	0	0	2
9	1	1	0	4	0	0	6
10	0	1	0	4	0	1	6
TOTALS	14	8	3	21	0	4	50
PERCENT	28	16	6	42	0	8	

II. RESULTS OF TESTING THE HYPOTHESES

In order to carry out a statistical analysis of the response patterns of students, each response to each item was assigned a score of zero, one, two or three, as outlined in the previous chapter. Since eight items were administered to each student, the maximum total score obtainable was twenty-four. Also, since this study was carried out near the end of the school year, each teacher was requested to provide achievement data for those students who participated in the study. As a result, a mathematics achievement mark was obtained for each student in nine of the ten classes that participated.

This data was used to investigate (a) whether or not the mean total scores of each class who were administered the instrument are significantly different; (b) whether or not there are any significant correlations between total score and achievement in each class; (c) whether or not over the entire sample there are any significant correlations between the score obtained on each situation and achievement, and also between total score and achievement, (d) whether or not there are any significant inter-correlations between the scores obtained on each of situations one through seven; and (e) whether or not the mean scores of males and females differed significantly on each situation and also on total score.

The mean achievement score, the mean scores obtained on situations one through twelve and the mean total score for all students are presented in Table XVII. Also, the distribution of scores for each situation is presented in Table XVIII.

TABLE XVII

MEAN ACHIEVEMENT AND SITUATION SCORES

VARIABLE	N	MEAN SCORE	STANDARD DEVIATION
ACHIEVEMENT	232	62.37	14.53
SITUATION ONE	255	2.01	1.19
SITUATION TWO	255	1.25	0.81
SITUATION THREE	255	1.32	0.99
SITUATION FOUR	255	2.36	0.96
SITUATION FIVE	255	1.25	0.88
SITUATION SIX	255	1.43	1.31
SITUATION SEVEN	255	1.27	1.22
SITUATION EIGHT	53	1.58	1.06
SITUATION NINE	48	0.98	0.73
SITUATION TEN	52	2.63	0.77
SITUATION ELEVEN	52	1.21	0.54
SITUATION TWELVE	50	1.76	1.27
SITUATION TOTAL	255	12.49	3.93

TABLE XVIII

DISTRIBUTION OF SCORES FOR EACH SITUATION
(Percentage in brackets).

SITUATION	SCORES 0		1		2		3		N
1	42	(16.5)	51	(20.0)	22	(8.6)	140	(54.9)	255
2	38	(14.9)	136	(53.3)	60	(23.5)	21	(8.2)	255
3	59	(23.1)	93	(36.5)	66	(25.9)	37	(14.5)	255
4	4	(1.6)	75	(29.4)	2	(0.8)	174	(68.2)	255
5	38	(14.9)	153	(60.0)	27	(10.6)	37	(14.5)	255
6	100	(39.2)	34	(13.3)	33	(12.9)	88	(34.5)	255
7	110	(43.1)	21	(8.2)	70	(27.5)	54	(21.2)	255
8	8	(15.1)	21	(39.6)	9	(17.0)	15	(28.3)	53
9	13	(27.1)	22	(45.8)	13	(27.1)	0	(0.0)	48
10	2	(3.8)	5	(9.6)	5	(9.6)	40	(76.9)	52
11	1	(1.9)	40	(76.9)	9	(17.3)	2	(3.8)	52
12	14	(28.0)	8	(16.0)	7	(14.0)	21	(42.0)	50

A. Hypothesis One

The mean total scores for each group of students who were administered the instrument are not significantly different.

Results

The means and standard derivations of the scores for each group who were administered the instrument are reported in Table XIX. Hypothesis one was tested using a one way analysis of variance, the results of which are also presented in Table XIX. The resulting F-ratio of 2.42 is non-significant at the 0.05 level. Hence, hypothesis one was not rejected.

B. Hypothesis Two

For each group of students who were administered the instruments, there are no significant correlations between achievement in mathematics and the total score obtained by each student.

Results

The mean situation scores for each group are reported in Table XX together with the mean achievement scores for each group except group five for which achievement scores were not obtained. The last column of Table XX presents the probability of a non-significant correlation between mathematics achievement and total score for each group. As a result of these findings, hypothesis two was rejected for all groups with the exception of groups five, seven and marginally for group ten. It was concluded, therefore, that while the standards under which achievement marks were awarded in each

TABLE XIX

ANALYSIS OF VARIANCE ON MEAN TOTAL SCORE FOR EACH CLASS

CLASS	N	MEAN	VARIANCE
1	29	10.93	15.64
2	24	12.58	18.43
3	33	12.73	16.20
4	25	13.00	13.67
5	23	11.35	12.60
6	31	11.39	15.78
7	19	13.21	11.06
8	11	11.91	12.69
9	32	14.72	13.43
10	28	12.68	15.11

SOURCE	df.	ss.	ms	F	p
CLASSES	9	320.27	35.59	2.42	0.012
ERROR	245	3609.47	14.73		

TABLE XX

CORRELATIONS BETWEEN TOTAL SCORE AND ACHIEVEMENT FOR EACH GROUP

GROUP	SCHOOL	N	MEAN SITUATION SCORE	MEAN ACHIEVEMENT SCORE	CORRELATION	p.
1	A	29	10.93	61.21	0.54	0.003
2	B	24	12.58	59.12	0.44	0.030
3	C	33	12.73	61.67	0.47	0.006
4	C	25	13.0	57.20	0.51	0.009
5	D	23	11.35			
6	E	31	11.39	65.16	0.44	0.013
7	F	19	13.21	65.68	-0.10	0.680
8	G	11	11.91	58.0	0.68	0.022
9	H	32	14.72	61.25	0.52	0.002
10	I	28	12.68	69.46	0.37	0.054

class may differ, the overall level of responses obtained from students in each class correlates significantly with their level of achievement in mathematics.

C. Hypothesis Three

There are no significant correlations between achievement in mathematics and (a) the score obtained on each situation; (b) the total score obtained.

Results

Table XXI presents all of the computed intercorrelations between achievement, the scores obtained on situations one through twelve and the total score obtained on the eight situations presented to students. The intercorrelations between the scores of situations eight through twelve could not be computed since these items were administered to disjoint subsets of the sample. Consequently, only the intercorrelations between the scores of situations one through seven were considered meaningful since these were the only items administered to the entire sample.

With respect to hypothesis 3(a), significant correlations, at the 0.05 level, were found to exist between achievement and the scores obtained on situations one, four, six, seven, eight, eleven and twelve. Hence, hypothesis 3(a) was not rejected for situations two, three, five, nine and ten, but was rejected for the other items.

A correlation significant at the 0.001 level was found to exist, however, between achievement and total score. Hence, hypothesis

TABLE XXI

INTERCORRELATIONS BETWEEN ACHIEVEMENT AND SITUATION SCORES

VARIABLE	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0 ACHIEVEMENT	1.000	0.242*	0.074	0.082	0.244*	-0.022	0.201*	0.308*	0.237*	-0.055	0.103	0.257*	0.509*	0.368*
1 SCORE SIT ONE		1.000	0.147*	0.089	-0.055	0.033	0.108*	0.206*	0.033	0.062	0.146	0.121	0.199	0.480*
2 SCORE SIT TWO			1.000	0.132*	-0.060	0.023	0.132*	0.131*	0.114	0.202	0.145	0.258*	0.155	0.389*
3 SCORE SIT THREE				1.000	0.067	0.077	0.168*	0.093	0.060	-0.170	-0.123	0.031	0.037	0.415*
4 SCORE SIT FOUR					1.000	0.049	0.001	0.043	-0.287*	-0.260*	-0.017	0.108	-0.036	0.241*
5 SCORE SIT FIVE						1.000	0.153*	0.070	-0.005	0.009	0.067	-0.014	0.163	0.348*
6 SCORE SIT SIX							1.000	0.462*	0.375*	-0.171	0.040	0.296*	0.180	0.654*
7 SCORE SIT SEVEN								1.000	0.299*	0.144	0.234*	0.321*	0.034	0.659*
8 SCORE SIT EIGHT									1.000	-	-	-	-	0.487*
9 SCORE SIT NINE										1.000	-	-	-	0.182
10 SCORE SIT TEN											1.000	-	-	0.363*
11 SCORE SIT ELEVEN												1.000	-	0.446*
12 SCORE SIT TWELVE													1.000	0.443*
13 TOTAL SIT SCORE														1.000

$* p < .05$

3(b) was rejected.

In addition to the above findings, the intercorrelations between the scores on situations one through seven were examined. This examination together with the results of hypothesis 3(a) resulted in the following observations:

(a) The scores on situations one and six, one and seven, and six and seven correlate significantly with each other and the scores of each of these situations correlates significantly with achievement;

(b) The scores on situations two and three do not correlate significantly with each other;

(c) The scores on situation four does not correlate significantly with the scores on any of situations one, two, three, five, six or seven but does correlate significantly with achievement;

(d) The score on situation five correlates significantly only with the scores on situation six. These were the only observations that appeared to have any significance and were made in the context of examining what might be termed uni-dimensional relationships between the performance of students on situations one to seven. However, it is possible that only a multi dimensional interpretation can account for all of the relationships evident in these observations.

D. Hypothesis Four

There are no significant differences between (a) the mean scores of males and females on each of situations one through seven; (b) the mean total scores of males and females

Results

Hypothesis four was tested using a two-tailed t-test. The appropriate means and standard deviations together with t-values and probabilities are reported in Table XXII. As a result of the findings, Hypothesis 4(a) was not rejected for all situations investigated except situation six where the mean score of females was significantly higher than that of males.

The mean total scores of males and females was not significantly different, so that hypothesis 4(b) was also not rejected.

III. SUMMARY

A variety of responses were obtained for each item administered in this study. As a result, response categories were devised for each item and each response assigned to one of the categories. In this chapter, the distribution of responses by category for each item has been presented together with a discussion of the categories and typical responses assigned to them.

The response categories for each item were further classified as representing either low, medium or high levels of thinking and a score of one, two or three, respectively, was assigned to each level. Students who did not respond were assigned a score of zero. The scoring scheme was used to analyze the response patterns of students statistically, and the results of the hypotheses tested are also presented in this chapter.

TABLE XXII

COMPARISON BETWEEN THE SCORES OF MALES AND FEMALES

VARIABLE	MALE MEAN SCORE	STANDARD DEVIATION	FEMALE MEAN SCORE	STANDARD DEVIATION	DF	t-value	Probability
SITUATION ONE	1.97	1.27	2.22	1.11	253	-1.623	0.10
SITUATION TWO	1.31	0.85	1.29	0.78	253	0.183	0.85
SITUATION THREE	1.37	1.07	1.36	0.92	253	0.019	0.98
SITUATION FOUR	2.49	0.94	2.32	0.98	253	1.386	0.17
SITUATION FIVE	1.28	0.89	1.20	0.87	253	0.712	0.48
SITUATION SIX	1.27	1.28	1.65	1.33	253	-2.257	0.02
SITUATION SEVEN	1.24	1.24	1.30	1.19	253	-0.417	0.68
TOTAL SCORE	12.18	4.07	12.94	3.71	253	-1.528	0.13

CHAPTER V

SUMMARY, CONCLUSIONS, DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

I. SUMMARY OF THE STUDY

The present study was designed primarily to investigate, identify and categorize some of the subjective thinking processes used by students when they are faced with the problem of either accepting or rejecting various mathematical propositions and are asked to justify the position they take. The main objective of this undertaking was to try to assess the extent to which students understand the nature and role of mathematical proof with particular emphasis on (a) the circumstances where students do or do not see a need for proof, (b) the kinds of misunderstandings students appear to have and (c) the kinds of errors in thinking that they appear to exhibit when attempting to justify the propositions presented to them.

A secondary objective of this study was to evaluate whether or not items of the type used in this study have any potential for classroom use as a means of diagnosing student understanding of, or even teaching, some of the fundamental concepts necessary for a basic understanding of what proof in mathematics really means.

A. The Instrument

In order to carry out the study, a number of items were constructed which were intended to be administered to students on a class basis during regular school hours and to which students would respond in writing. These items were subjected to extensive pilot testing and subsequent revision resulting in a common format for their presentation to students and also the development of procedures for the analysis of student responses.

Twelve items were selected and administered to students. Most of these items were presented in the form of a dialogue between two hypothetical mathematics students in which one of the students makes a statement or states a conclusion which may or may not be true. The second student normally disagrees with the first. He sometimes rejects the stated conclusion completely or else claims that he does not understand it. In each item, one of the hypothetical students requests the other to essentially prove the statement he makes.

The first seven items were administered to all students in the sample. The remaining five items were administered at random to disjoint subsets of the sample so that approximately one fifth of the sample responded to each of these items.

In the first item, a simple geometric proposition was presented, the proof of which was given with reference to an appropriate geometrical figure. The objective of this item was to determine whether or not students see a need to prove the same proposition when it is illustrated by a different diagram. It was assumed that students who understand how the generalization principle is applied to propositions of this type, would agree that no additional proof is necessary.

Items two, three and four are algebraic or number-theoretic in nature and the objective of each was to determine what kind of evidence students consider adequate to prove a proposition which generalizes a number of numerical examples. Specifically, do students see a need to formally prove such a proposition or is the numerical evidence itself considered adequate justification?

In item five, a simple non-Euclidean geometry was presented. The objective of this item was to assess the extent to which students understand that the objects studied in mathematics have only those properties that are ascribed by definition or by postulate.

Items six and seven deal specifically with the indirect form of proof. The objective was to determine some of the factors that cause students to have difficulty with this form of argument.

Item eight was intended to assess whether or not students see a need to prove a mathematics proposition, the truth of which is considered to be intuitively obvious.

Items nine through twelve were intended to assess student understanding of (a) the logical equivalence of a mathematical statement and its contrapositive, (b) the role of a counter-example to a stated proposition and (c) the non-equivalence of a mathematical statement and its converse.

B. Procedures

The high school students who were administered the items were asked to respond to each item by deciding which of the two hypothetical students was, in their view, correct and to give reasons for their choice. The responses of students to each of the items were carefully examined by three independent judges. As a result, response categories for each item were developed and each response to each item was assigned to an appropriate category. Each category was considered to represent a distinct line of thinking so that the categorization scheme was used to summarize not only the kinds of thinking exhibited by students but also the distribution of responses for each of the items.

In order to analyze the response patterns of students statistically, each category was assigned a score. The scoring scheme was used mainly to determine whether or not there was any relationship between student achievement in mathematics and how they responded to the various items.

C. The Sample

The instrument was administered to two hundred fifty five grade eleven mathematics students in ten randomly selected classes of the Edmonton Public and Separate School Systems. All students were enrolled in the course Mathematics 20 which was designed as a preparation for the study of post-secondary mathematics.

II. DISCUSSION OF SOME OF THE FINDINGS

The main conclusions resulting from this study are presented in this section and will be discussed in terms of the seven specific questions which this study was designed to answer.

A. Question One

To what extent or in what contexts do high school students subjectively see the need for proof in mathematics?

The responses of students to each of the items incorporated a number of distinct lines of thinking, each of which was represented by an appropriate response category. Based upon the distribution of responses in the various categories for each item and also upon the level of response each category was considered to represent, it was concluded that in general the student population sampled can be sub-divided into three distinct groups, designated as high, medium and low.

Students in the high group, which contained approximately thirty percent of the students sampled, do appear to have a basic understanding of the nature and role of mathematical proof. In particular, these students appear to understand that it is always necessary to prove any mathematical proposition before calling it a theorem, in the sense of establishing a chain of reasoning, based upon accepted assumptions or postulates, undefined terms, definitions and previously proven propositions, which, provided the accepted rules of logic are followed, demonstrates that a conclusion is necessarily true if the postulates on which the argument is based are accepted as true.

Approximately twenty percent of the students belong to the medium group. These students show some understanding of the importance of proof in mathematics although it was evident that in some cases this understanding was minimal. While most of these students appear to see a need for proof, it was not clear that they really understand the meaning of proof. In particular, these students more often than not failed to distinguish between a valid and an invalid argument. In other words, they did not always appear to understand what did and what did not constitute proof of some of the conclusions presented to them.

The most interesting group is the low group, which contains approximately fifty percent of the students sampled. It was concluded that this group of students do not, in general, understand the importance or meaning of proof in mathematics. Most of these students do not, for example, distinguish between an inductive and a deductive argument. In fact, most of these students appear to understand very little other than the fact that every mathematical proposition is either true or false. In general, students in this group appear to adopt the attitude that some form of justification, not necessarily formal proof, may be necessary to verify certain propositions. However, whether or not these students see a need to justify a particular proposition appears to depend upon the extent to which students subjectively believe that the proposition is true, based upon their intuition or numeric evidence or maybe even because the proposition was stated by some one in authority. In fact, the extent to which not only students, but also teachers and professional mathematicians believe that a particular proposition is true is perhaps of primary importance in terms of whether individuals either see a need to prove a proposition or will in fact attempt to prove it.

To illustrate, it is perhaps safe to assume that a mathematician will not attempt to prove a proposition unless he or she believes it to be true. Many of the great unsolved problems or conjectures in mathematics have attracted the efforts of many able mathematicians mainly because there was an inherent belief that these problems could be solved and proofs established. It is clear that if mathematicians had no faith that, for example, the four color conjecture would eventually be proven, then it is doubtful that the vast literature which resulted from attempts to prove this conjecture would have been witnessed. In general, if a mathematician does not believe that a proposition is true, then it is doubtful that any attempt will be made to prove it, although, obviously, attempts might be made to disprove it.

It was evident, however, from this study, that most high school students, especially those in the low group, think quite differently. In fact, it would appear that the thinking of these students is almost the exact opposite to that of the mathematician. Specifically, students in the low group do not appear to see any need to prove any mathematical proposition which they believe to be true or consider intuitively obvious.

On the other hand, there was some evidence to suggest that many of these students do see a need for some form of justification of propositions which are not intuitively obvious or which are difficult to understand. However, in instances where students in the low group do see a need for proof, the actual justification given by these students was very subjective in nature, amounting in most cases to nothing more than an opinion.

It is perhaps significant to note that most students do appear to understand that a single counter-example to a stated proposition disproves the proposition. In fact, a surprising number of students in the low group argued that any proposition which is intuitively obvious or which is supported by numerical evidence is true unless it can be proven false.

In summary, it would appear that less than thirty percent of the students sampled really understand the meaning of proof and see the need to deductively prove any proposition that is not an axiom or definition.

Approximately half the students show very little understanding of the meaning of proof in mathematics or of what constitutes proof and do not see any need to prove any statement which is considered to be intuitively obvious.

B. Question Two

To what extent to high school students subjectively realize that induction is inadequate to support mathematical generalization?

The responses of students to situations two and three in particular suggest that at least seventy percent of the students sampled do not realize that induction is inadequate to support the type of mathematical generalizations presented in these items. In fact, less than ten percent of the students sampled explicitely stated that a deductive proof was necessary to justify the validity of the conclusion presented in situation two. Approximately eighty percent of the students appeared to be convinced that the conclusion was true on the basis of the numerical evidence presented.

The inductive nature of some students thinking is clearly evident from the responses of those students who argued that the conclusion presented in situation two cannot be proven for all positive integers since there are an infinite of cases to be considered. Most students, however, argued inductively in the other direction, stating that the conclusion could be proven true by simply enumerating a number of special cases in addition to those presented. The classic response was one which suggested that there was very little chance that the generalization was false since it could easily be shown true for all positive integers up to ten.

The apparent failure to distinguish between or understand the difference between an inductive and a deductive argument and hence between a conjecture and a theorem, is one of the most widespread errors evident in the subjective thinking exhibited by those students sampled. In fact, it is this characteristic perhaps, more than any other, that distinguishes student thought from more mathematically mature forms of thought.

It should be noted, perhaps, that when the distribution of responses for situations two and three were compared with the distribution of responses for situation one, it was apparent that many students who were classified as inductive thinkers as a result of their responses to situations two and three, were also classified as understanding the application of the generalization principle in situation one. In fact, slightly over thirty percent of the students who scored low on situations two and three, also scored high on situation one. Since students were generally consistent in the level of responses they gave to the various items, this observation suggests that many of the students, whose responses were considered mathematically correct in situation one, argued inductively and hence did not really understand why the proof given in situation one applied to both diagrams as implied by the generalization principle. If this explanation is, in fact, valid, then perhaps it is not surprising that many students do not realize that induction is inadequate to support mathematical generalizations especially if students do not make any distinction between the type of propositions presented in situation one and those presented in situations two and three.

C. Question Three

To what extent do high school students subjectively realize that the objects studied in mathematics have only those properties ascribed to them by definition or by postulate?

The distribution of responses to situation five, in particular, suggests that most high school students do not always realize the significance of assumptions and definitions, especially those that do not agree with their intuition or are used in an unfamiliar context. This conclusion was based upon the fact that approximately eighty-five percent of the students sampled believed that the proposition presented in situation five was untrue, even though a partial proof was provided.

The proposition in question, namely, that three distinct pairs of parallel lines are determined by the four distinct points, was rejected by these students for several related reasons. First, it was evident that many students considered the entire discussion presented in situation five to be meaningless because it was based, in their view, upon a false hypothesis. These students argued that since the plane consists of an infinite number of points, it is meaningless, or in the words of many, "impossible", to assume that only four distinct points exist. It would appear, therefore, that these students are unwilling to reason from or to accept any argument based upon an assumption which they believe is false.

Approximately sixty percent of the students rejected the conclusion stated in situation five because they apparently refused to accept that a line, by definition, was determined by and consisted of exactly two distinct points. These students did not appear to understand the significance of this definition and, judging from their responses, continued to think of a line in the usual Euclidean sense. This was particularly evident in the responses of those students who suggested that the lines in the diagram of situation five obviously intersect and therefore cannot possibly be parallel. These students evidently ignored completely the hypothesis that only four distinct points exist and consequently did not realize that the definition of parallel lines presented in situation five was perfectly valid.

In summary, the reasons given for not accepting the conclusion presented in situation five indicate that the vast majority of students do not realize that this conclusion and the argument supporting it can only be interpreted in terms of the hypothesis and definitions on which it is based.

D. Question Four

To what extent do high school students subjectively understand the indirect method of proof in mathematics?

On the basis of this study, it was concluded that less than twenty percent of the students sampled really understand the method of indirect proof. It was evident from the responses given to situations six and seven, that most students are confused by an indirect argument. In fact, many believe that an indirect proof is not a valid form of reasoning.

There are a number of possible explanations for the apparent difficulty that students have with an indirect argument. Many students indicated that they would never be able to prove a proposition on their own using this form of argument. Others suggested that the indirect method of proof is an "unnatural" form of reasoning. It would appear, therefore, that most students do not consider the indirect method of proof to be a convincing form of argument.

The objective of an indirect argument is to show that the negation of a stated proposition is false. If the negation can be shown false, then the stated proposition must be true. Hence, the central idea of the method of indirect proof is to assume, for the sake of argument, that the negation of a stated proposition is a true statement. It if can be demonstrated that this assumption implies another statement which is inconsistent with the stated proposition or other known theorems, then the negation must also be contradictory and therefore false. As a result of this contradiction, the stated proposition must therefore be true.

It would appear, however, that most students are not convinced by an indirect argument because they do not understand the logic of such arguments, as outlined above. Specifically, many students appear to view an indirect argument as if it were a direct argument. Hence many are confused because, in their view, the argument does not prove the stated proposition but rather proves the "opposite". In other words, these students do not realize that the objective of the argument is to show that the negation is false but believe that it, in fact, shows the negation to be true and is thus an invalid and contradictory form of reasoning. This was particularly evident from statements such as: "You can't prove something is true by showing that the opposite is true" or "You can't prove something is true by showing that something else is false".

Another reason why students have difficulty with an indirect argument results from the fact that many do not distinguish between the negation of a stated proposition and the proposition itself. These students therefore claim that in such arguments, the statement to be proven is assumed in order to prove it. Hence the argument is considered invalid because it employs circular reasoning and effectively proves nothing.

Perhaps the most significant reason why students do not understand an indirect argument is because many believe that the negation of any stated proposition is always a false statement. As was evident from the responses of students to situation five, most students consider any conclusion or any argument to be invalid if it is based upon an assumption or hypothesis which they believe is false. This kind of reasoning was clearly evident from the responses of students to situations six and seven also. Hence many students considered the indirect arguments presented to be invalid because, in their words "you can't assume that a false statement is true".

In summary, it is possible that to fully understand the indirect method of proof requires a greater degree of mathematical maturity than that demonstrated by most of the students sampled. Based upon the kinds of reasons students gave, it is evident that it may be unrealistic to expect most students at the high school level to understand and use what is essentially a very sophisticated form of argument greatly misunderstood not only by students but by many others as well.

E. Question Five

To what extent do high school students subjectively realize that a single counter-example to a stated proposition is sufficient to reject that proposition?

On the basis of this study, it would appear that high school students do understand that a single counter-example to a stated proposition is sufficient to disprove that proposition. Evidence for this conclusion was obtained from the response of students to situations two, three, four, ten and twelve. In fact, it was evident that the counter-example idea was understood by more students than any of the other concepts that were investigated.

Situation ten was designed specifically to assess whether or not students understood the role of a counter-example. A false generalization, together with a specific counter-example was presented in this item. Almost eighty percent of the students readily agreed that this counter-example was sufficient to disprove the stated conclusion.

The conclusions presented in situations four and twelve were also untrue. In both of these items, however, no counter-example was suggested nor was there any hint that the stated conclusions were incorrect. In spite of this approximately sixty percent of the students provided a counter-example to reject the conclusion presented in situation four and over forty percent did the same in situation twelve.

As discussed earlier, a significant number of students agreed that the conclusions presented in situations two and three were true statements because they could not find any counter-example to disprove these statements or because they were convinced that no such counter-examples exist. These students apparently reasoned that since any proposition can only be disproven by exhibiting at least one counter-example, therefore a statement is proven if no counter-example can be found. It is possible, however, that many of these students used this line of reasoning only because they considered the conclusions presented in situations two and three to be intuitively obvious and hence either did not see the need for a formal proof or could not provide such a proof. In any event, it was evident that while most students do appear to realize that a single counter-example to a stated proposition is sufficient to reject that proposition, a significant number of students incorrectly believe that the notion of counter-example can be used in some situations to prove certain propositions, a finding which was totally unexpected.

F. Question Six

To what extent do high school students subjectively realize that a mathematical statement and its contrapositive are logically equivalent?

On the basis of student responses to situation nine, there was no evidence to suggest that high school students understood the logical equivalence of contrapositive mathematical statements. As was indicated in the previous chapter, not a single student recognized that Statements A and B presented in situation nine were contrapositive statements and hence logically equivalent. In fact, it was stated explicitely in situation nine that the proof given for Statement B also proves Statement A. In spite of this, most students felt that the two statements were unrelated. Almost half of the students, in fact, suggested that in order to show Statement A true, a different proof was necessary than that given for Statement B.

C. Question Seven

To what extent do high school students subjectively recognize that a mathematical proposition and its converse are not logically equivalent?

The evidence obtained relative to the above question, as with the previous question, is perhaps insufficient to justify any positive conclusions. Only one item, namely situation twelve, was designed with question seven in mind. While the majority of students realized that the proposition stated in situation twelve was false, none of the students sampled appeared to understand why the argument presented was invalid. In particular none of the students indicated that the argument proved the converse of the stated proposition. Almost half the students sampled, however, did correctly disprove the stated proposition with a suitable counter-example. It is possible that most students who realized that the proposition was untrue, simply ignored the argument and attempted to find a counter-example. Some of the students however, could not evidently reconcile what they considered to be a false proposition with a valid proof. In any event, there was no evidence to suggest that high school students recognize the fact that a mathematical statement and its converse are not logically equivalent. In fact, about the only conclusion that can possibly be drawn is that many students do not distinguish at all between a statement and its converse.

III. DISCUSSION OF THE RESULTS OF THE HYPOTHESIS TESTING

In order to carry out a statistical analysis of the response patterns of students, the response of each student to each item was assigned a score of zero, one, two or three, as outlined in Chapter Three. Since each student responded to eight of the items, the maximum possible score was twenty-four.

The statistical analysis was carried out primarily to investigate whether or not the level of response students gave to all eight items, as measured by their total score, was related to their achievement in mathematics. It was decided as well to use the scoring scheme to determine whether or not there were any significant differences between the mean total scores of each class who were administered the instrument, to look for inconsistencies in the response patterns of students to pairs of items and to determine whether or not there was any significant difference in the mean total scores of males and females.

The major findings from this analysis can be summarized relative to the hypotheses which was tested, as follows: (a) No significant differences were found to exist between the mean total scores of each group of students who were administered the instrument; (b) In seven of the ten classes which were administered the instrument, significant positive correlations were found to exist between mathematics achievement and the total score obtained on the instrument; (c) For all students for whom achievement data was obtained, a significant correlation was found to exist between mathematics achievement and total score. However, correlations between mathematics achievement and the scores obtained on situations two, three, five, nine and ten were not significant, and (d) For the entire sample, no significant difference was found to exist between the mean total scores of males and females. Furthermore, no significant differences were found to exist between the mean scores of males and females on each of situations one through seven, with the exception of situation six, where the mean score of females was found to be significantly higher than that of males.

With respect to the first of the above findings, it will be recalled that the instrument was administered to ten intact classes in nine different high schools. It was recognized that these schools differed not only in size, but also in purpose and philosophy. Some of the schools were large composite high schools with a great deal of emphasis on extra curricular activities; others were smaller and more intimate with a greater emphasis on academics. In addition, perhaps the most important variable of all, in each class, was the teacher whose demands and expectations of students may have resulted in entirely different attitudes in different classes. These factors could have had an effect upon how students responded to the various items thus producing variations in the response patterns from different classes, in spite of the fact that every effort was made both in the construction of the items and in their administration to prevent students from perceiving the instrument as a test of their mathematical knowledge. In fact, the items which were administered to students were selected on the basis that each would be unfamiliar to most students so that hopefully the effect of any previous instruction would be minimized. However, there was no way of insuring this in advance.

The fact that the mean scores of the ten classes were not significantly different suggests that the factors mentioned above did not significantly influence how students responded to the various items. Also, the fact that the response patterns of students in different classes were consistent would appear to enhance the reliability of the items used not only from the point of view of this particular study but also in terms of their potential usefulness as a means by which teachers can diagnose student understanding of many concepts related to the notion of mathematical proof.

The findings which show a significant correlation between mathematics achievement and total score suggest that previous instruction in mathematics did in fact influence how students responded to the various items. Perhaps it would be unrealistic to assume otherwise, at least for some of the items. However, the significance of this finding lies, perhaps, in the fact that only the very best students, both mathematically and perhaps intellectually, have displayed a mature understanding of what proof in mathematics really means. On the other side of the coin, judging from the type of responses obtained at the lower end of the scale, it would appear that the vast majority of high school students, all except the very high achievers, do not possess anything approaching a clear understanding of many of the concepts related to mathematical proof that were investigated in this study.

The final major finding shows that contrary to popular belief, as far as understanding the nature and role of mathematical proof is concerned, the level of understanding displayed by girls is no different from that displayed by boys. It would appear, therefore, that the main variable is not sex but brainpower.

In addition to the findings discussed above, and in spite of the fact that a correlation significant at the 0.001 level was found to exist between mathematics achievement and total score, correlations were obtained between achievement and the score on each situation. It was found that non-significant correlations exist between achievement and the scores on situations two, three, five, nine and ten.

Situations two and three were designed to assess whether or not students use inductive or deductive reasoning to justify the generalizations presented in these items. Hence, on the basis of the previous discussion, a non-significant correlation between achievement and the scores obtained on these situations was interpreted to suggest that even some of the better students are prone to use induction to support mathematical generalizations of the type presented in these items.

Situation five was designed to assess whether or not students are willing to accept and use definitions and hypotheses, which may be new and unfamiliar, in the context of a finite non-Euclidean geometry. Again, the fact that achievement did not correlate significantly with the score obtained on this item suggests that even the better students encountered difficulty with this item.

On situation nine, none of the students indicated that the two contrapositive statements were logically equivalent, thus explaining the non-significant correlation between the score obtained on this situation and achievement. Situation ten was designed to assess specifically student understanding of the concept of counter-example. The fact that there was no significant correlation between achievement and the score obtained on this situation is explained by the fact that the majority of students appear to understand the role of a counter-example including many of the low-achievers.

Significant correlations at the 0.05 level were found to exist between mathematics achievement and the scores obtained on situations one, four, six, seven, eight, eleven and twelve. These findings have been interpreted to suggest that only the high achievers displayed a mature understanding of the concepts being investigated in each item. For example, the fact that achievement correlated significantly with the scores obtained on situations six and seven suggests that only the very best students academically show any understanding of the method of indirect proof.

Finally, as reported in Table XXI, the incorrelations between the scores of each situation were obtained. An examination of these intercorrelations revealed what may be considered as an inconsistency in the response patterns of students on situations one, two and three. Specifically, a significant correlation was found to exist between the scores obtained on situations one and two and also between the scores obtained on situations two and three. However, the correlation between the scores obtained on situations one and three was non-significant. Furthermore, the score on situation one correlated significantly with achievement, whereas the scores on situations two and three did not. These results, while indicating apparent anomalies in the response patterns of students, are consistent with the observations, reported earlier in this Chapter, resulting from a comparison of the response patterns of students on situations one, two and three. To be specific, approximately one third of the students who were classified as understanding the use of the generalization principle in situation one, were also classified as inductive thinkers on situations two and three, while a significant number of students responded in just the opposite manner on these items.

If it is assumed that students responded consistently to items one, two and three, then there would appear to be several possible explanations for the apparent anomalies. Firstly, it is possible that some of these students considered the conclusions presented in all three items to be intuitively obvious and therefore did not see the need for any formal proof in each of the three situations. This would explain why some of these students were categorized as understanding the generalization principle in situation one. It would also explain, perhaps, why some of these students considered an inductive argument to be adequate justifications for the conclusions presented in situations two and three.

It is also possible, however, that some of these students responded to situation one not on the basis of an understanding of the generalization principle but rather by reasoning inductively consistent with their responses to situations two and three. To illustrate why this is possible, it should be noted that the propositions presented in situations one and two, for example, have the same general form, namely both are statements of the form P(x) where x belongs to an appropriate domain D. In general, to prove that a proposition of the form P(x) is true for all x in D, it is first assumed, without any loss of generality, that x is an arbitrary but fixed element of D. If it is then possible to prove the statement P(x) true for this arbitrary but fixed value of x, the generalization principle is applied, either implicitely or explicitely, to conclude that P(x) is true for all x in D.

In situation one, the domain D is the set of all geometric figures isomorphic to Joe's diagram. P(x) states that the area of the triangle is one-half the area of the rectangle for all such geometric figures. Since P(x) was proven for the arbitrary but fixed figure labelled Joe's diagram, no further proof is necessary to conclude that P(x) is true for all other isomorphic figures and, in particular, for Tom's diagram. It is possible, however, that some of the students who correctly stated that no further proof was necessary in situation one, did so on the basis of inductive rather that deductive reasoning. The reason for suggesting this is that essentially the same kind of argument must be used to prove the proposition presented in situation two. In this case, the domain D is the set of all positive integers. P(x) states that $4^x - 1$ is divisible by three. Again, if it can be shown that P(x) is true for any arbitrary but fixed positive integer x, which is not very difficult, then the generalization principle can be used to conclude that P(x) is true for all positive integers x. It was evident, however, that most students did not use this form of reasoning to justify the conclusion presented in situation two. Most students, in fact, believe that if P(x) can be shown true for one or more fixed or constant values of x, that is, for one or more specific positive integers, then this is sufficient to conclude that P(x) is true for all positive integers. It is possible of course that these students resorted to this inductive form of reasoning because they did not have any idea how to prove the stated proposition otherwise. However, based upon the fact that approximately one third of all students sampled, who justified the conclusions presented in situations two and three inductively, were classified as understanding the generalization principle on situation one, it is also possible that these students responded to situation one in the same way and for the same reasons as they did to situations two and three.

In summary, whether or not the responses of students to situations one, two and three are inconsistent, it would appear that students need a great deal of mathematical maturity to distinguish between showing a proposition of the form P(x) true for a fixed or constant value of x and proving P(x) true for an arbitrary but fixed value of x. Yet, this is essentially what students must be able to do if they are to understand not only the full generality of many propositions but also the difference between empirical evidence supporting a proposition and a valid proof of that proposition. It would appear that the majority of high school students are not mathematically mature enough to make this distinction without, perhaps, specific instruction and exposure to a variety of examples illustrating these ideas in an explicit manner.

IV GENERAL OBSERVATIONS

Smith and Henderson (1959) state:

> To the unsophisticated, 'proof' is practically synonymous with 'what convinces me'. A statement is 'proved' when a person is convinced that it is true. Such a concept makes proof a subjective and personal matter. The purpose of considering proof as a major idea in mathematics is to lead students from such a subjective concept to a more objective one, a concept based upon criteria impersonal in nature. (p. 178).

It was evident from this study that approximately half the students sampled were totally "unsophisticated" with respect to the concept of proof. These students do not see any need to prove any proposition that they consider to be self-evident or intuitively obvious. In this respect, the thinking of these high school students is no different from that exhibited by most students at the Junior high level. However, this kind of thinking at the Junior high school level is not surprising since many of the students are still at Piaget's concrete operational stage of development. Students at the Senior high school level are assumed to be at the formal operational stage, although the kind of thinking exhibited by many of the students sampled certainly does not confirm this.

On the basis of this study, at least seventy percent of the students in grade eleven do not appear to realize that generalizations about infinite sets of mathematical objects cannot be adequately supported by examining a finite number of cases. This finding contradicts a finding of Robinson (1964). She concluded that students do have this understanding, at least, by grade seven. It is possible that some grade seven students do have this understanding, but it is evident from this study that a significant number of students at the grade eleven level do not.

Robinson (1964) also concluded that:

> Most seventh and ninth-grade students will, when given free choice, justify a mathematical generalization by deductive reasoning from a set of premises if and only if those premises agree with their intuitions; that is, if and only if those premises are believable to the children (p. 101).

It was evident from this study that this same conclusion is also valid for over eighty percent of the grade eleven students sampled, most of whom, in fact, were not willing to reason, for the sake of argument, from any premise which they consider to be false. It was also determined that this is one of the main reasons why high school students experience difficulty with the indirect method of proof.

Again, in view of Piaget's findings, it is not surprising that many junior high school students would react to assumptions in this way. However, the fact that most senior high school students in the top mathematics stream of their respective schools did so was unexpected, although this finding is consistent with some of the findings of Reynold's (1967) which were discussed in Chapter II. He found that in responding to questions related to notion of mathematical proof, students in the age group seventeen to eighteen responded no differently to many of his items than students in the age groups twelve to thirteen.

This finding would appear to be significant because if most students at the senior high school level are only willing to reason from premises that they consider to be true, then it would appear that these students have very little idea about the nature of mathematical reasoning in general. In particular, such a finding suggests that these students do not understand that the very essence of a mathematical system is a set of assumptions called axioms or postulates which are assumed to be true whether they are self-evident or not. Furthermore these assumptions constitute the genesis of all mathematical structures and hence the core of all mathematical reasoning. If these students are unwilling to reason from a postulate because they consider it to be a false statement, or because it is not intuitively obvious, then the whole question of whether or not most high school students have benefitted from, or can benefit from, a study of formal axiomatics would appear to demand much greater study.

On the basis of this study, it was also evident that most high school students do not understand the logical principles underlying some of the arguments presented. This was particularly evident in the case of the method of indirect proof which does not appear to be understood by approximately eighty percent of the students sampled, mainly because they do not understand the logic behind such arguments. There was also no evidence to suggest that students understand the logical equivalence of a mathematical statement and its contrapositive or that a statement and its converse are not logically equivalent. As a result, many of these students were clearly handicapped since in general it is impossible to separate proof from logic.

The only finding of a positive nature was that most high school students do appear to understand that a single counterexample to a stated proposition is sufficient to disprove that proposition. It was evident that all students realize that a mathematical proposition is either true or false, but not both, and perhaps the idea of proving a statement false is much simpler to understand than the idea of proving a statement true.

V. EDUCATIONAL IMPLICATIONS

Based upon the findings of this study, it is apparent that the search for improved ways of introducing students to the deductive nature of modern mathematics must continue. Whether or not students go on to study higher mathematics, it is still desirable to give high school students some idea of what modern mathematics is all about and what mathematicians actually do. It should be possible, for example, to give students some understanding of the importance of famous conjectures in mathematics such as Goldback's conjecture, Fermat's conjecture or perhaps the famous Four Color Theorem. It should be possible for students to understand why these simply stated, and to many, intuitively obvious or self-evident, conjectures have attracted the efforts of many mathematicians over the years. Surely, the imagination of students can be challenged and they can be taught to appreciate why the mathematics community will not be satisfied until some of these conjectures are either proven or shown to be false.

To accomplish the kind of understanding referred to above does not, perhaps, require that students be exposed to a long and complicated axiomatic treatment of any part of mathematics. In fact, as stated earlier, it would appear that a deductive approach to the teaching of mathematics is appropriate, perhaps, only for a very select few students.

On the basis of this study, it would appear that if the teaching of deductive methods is to be very meaningful for most students, even at the senior high school level, then students must not be required to prove or even memorize the proofs of long strings of propositions that to them are intuitively obvious. On the other hand, students should be given the opportunity to become totally familiar with any axioms, hypotheses or definitions before being required to reason with them. In fact, the use of hypotheses and definitions that are intuitively acceptable to students is indicated.

It was also evident that it is necessary to teach explicitly the logical principles underlying many of the arguments used in mathematics. It would appear that the best time to do this is concurrent with the use of these principles in actual mathematical situations. Based upon previous research, it would appear to be insufficient to teach a course on logic prior to or separate from a course in mathematics and hope that students will automatically transfer and apply this knowledge to mathematical situations. Hence students should be exposed to many examples where the logic of an argument is explicitly discussed and revealed to them. This would appear to be particularly useful in the case of the indirect method of proof and also for some of the other ideas assessed in this study, such as the use of the generalization principle or the logical equivalence of a statement and its contrapositive.

On the basis of this study, however, it may be unrealistic to expect most students at the high school level to understand and use some of the forms of reasoning investigated in this study, particularly the method of indirect proof. It is possible that some students will never understand how the method of indirect proof proves statements in mathematics or will ever be convinced by such an argument.

Students must also be taught explicitly the difference between inductive and deductive reasoning. In general terms, it is perhaps essential that high school students understand the inductive nature of most scientific discovery. The fact that most conclusions that result from research in the fields of medicine or psychology or even mathematics are based upon inductive evidence whereas only conclusions in mathematics can be verified by deductive reasoning, does not appear to be a concept impossible to teach to most students. However, on the basis of this study, it would appear that most high school students do not have this understanding or at least do not make any distinction between both kinds of reasoning. Perhaps more attention should be paid to the suggestion of Allendoerfer (1957) when he stated:

> At some stage in the high school mathematics curriculum, there should be a serious discussion of deductive systems per-se, and later applications of this to mathematics and to non-mathematical situations should be used to reinforce the understanding of the students about deductive methods (p. 66).

Finally, and perhaps most significantly, on the basis of this study, it would appear that items of the type used in this study can be used by teachers (a) to initiate the instructional process relative to many of the concepts to be taught and studied in the area of mathematical proof; (b) to diagnose and identify erroneous student thinking relative to many of these concepts; and (c) to discuss with students, either on an individual or class basis, the kinds of errors in thinking exhibited in order to convince them of the nature of such errors and in the hope that students can learn from their errors the correct form of thinking expected of them.

VI. SUGGESTIONS FOR FURTHER RESEARCH

The motivation for this study resulted partially from the realization that many of the items used in the past to assess learning outcomes in the area of mathematical proof did not assess the extent to which some of the more important objectives of the high school mathematics curriculum were being attained by students. Fawcett (1938), Fehr (1966) and Wilson (1971), among others have suggested that many teachers tend to emphasize only the lower behaviour levels in their testing. This was particularly evident in the area of mathematical proof, where it would appear that the main emphasis was placed on factual knowledge. If, in fact, the objectives which are emphasized in the usual tests and examinations given to students reflect what students are actually taught in the classroom, then it is, perhaps, not surprising that most students lack a basic understanding of what proof in mathematics is all about.

A secondary objective of this study was to determine whether or not items of the type used in this study have any potential for use in the classroom as means of diagnosing or assessing student understanding of some of the more important objectives of the high school mathematics curriculum that relate to the notion of mathematical proof. This effort, perhaps, only scratches the surface, but it does suggest that a much greater effort is needed to develop and validate alternative methods by which learning outcomes relative to these objectives can be diagnosed and assessed in a manner meaningful not only to teachers but also to students.

There is, therefore, a need to devise and evaluate alternative teaching strategies for the presentation of concepts in the proof domain at a level consistent with mathematical maturity of most high school students. Many suggestions have been made along these lines, most recently by Sher (1978). He suggested that Boolean Algebra would be an appropriate means to introduce students to a mathematical system in which students can understand, appreciate and even prove theorems. However, this suggestion, and others like it, needs to be evaluated objectively in an actual classroom situation so that the effectiveness of these teaching strategies can be determined.

Another recommendation for further research resulting from this study results from the observation that if students are to really understand the nature of dedective mathematics, then the logical principles that underly some of the more basic forms of reasoning used in elementary mathematics must be made explicit to students. It would appear that many of those principles are left to incidental learning with the result that most students graduate from high school without any formal knowledge of some of the most basic rules of inference. Hence there is a need to develop ways of incorporating a study of the basic logical principles used in elementary mathematics directly into the mathematics curriculum in such a way that students can see immediately how these principles are applied to particular forms of reasoning. Previous researchers, such as Phillips (1968), Roy (1970) and Bostic (1970) have concluded that formal instruction in the basic principles of logic that are essential to deductive reasoning does not guarantee that students will automatically be able to understand the use of these principles in mathematical arguments, especially when these principles are studied prior to or separate from the mathematics in which they are employed. Hence, it would appear that most students need explicit and perhaps repeated instruction on the use of these principles in a mathematical context. This would appear to be especially important in the case of particular forms of reasoning such as the method of indirect proof and some of the other forms of reasoning investigated in this study.

VII. EPILOGUE

For the benefit of those who may be interested in continuing the line of research reported here, it would, perhaps, be useful to indicate, in a retrospective sense, some of the areas where potential improvements to this study might have produced more reliable results.

It is possible that the wording and nature of some of the items may have had a confounding effect upon how students responded to these items. For example, it is possible that the definitions of a line and of parallel lines in situation five were not made explicit enough.

It is also possible that students considered the proposition presented in situation six (i.e. proving $1 \neq 0$) to be intuitively obvious and hence did not bother to read the argument presented or else considered it unnecessary. Hence a more appropriate item might have been used.

Perhaps the main suggestion that results from these observations is that it may have been possible to determine student perceptions of these items more accurately by carrying out individual interviews with selected students after the items were administered. Such a procedure could have been used to determine whether or not there are alternative interpretations of the results of this study than those given in this report. In fact, it is entirely possible that reliable data could have been obtained solely by interview rather than by the procedures adopted in this study. It would appear useful, in fact, to determine, via a comparative study, whether or not the conclusions of this study would have been any different if the data had been obtained solely by interview or by a combination of written response and subsequent interview.

One final suggestion would be that whatever method is used to gather data, it would appear that the kind of research reported here could be carried out with students at all levels from the Junior High School through to beginning University and beyond, the objective of which could be to determine more precisely the characteristics of the various stages through which students appear to pass in gaining an understanding of the nature and role of proof in mathematics. The refinement of a developmental model relative to how students progress in their understanding of proof could add immensely to one's understanding of how best and when to introduce students to the various concepts so as to maximize student understanding at each stage of his or her development.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, Frank B. The Use of the Axiomatic Method in Teaching High School Mathematics. In: The Role of Axiomatics and Problem Solving in Mathematics. Begle, (ed.), Washington, D.C. The Conference Board of the Mathematical Sciences, Ginn and Company, 1966, pp. 1-12.

Allendoerfer, Carl B. Deductive Methods in Mathematics. In: Insights into Modern Mathematics. NCTM Twenty-Third Yearbook. Washington, D.C. The National Council of Teachers of Mathematics, 1957, pp. 65-99.

Baker, Truman D. A Study of the Nature of Proof in Algebraic Systems. Unpublished Ph.D. Thesis, University of Mississippi, 1969.

Bell, E.T. The Meaning of Mathematics. In: The Place of Mathematics in Modern Education. NCTM Eleventh Yearbook. New York: Bureau of Publications, Teachers College, Columbia University, 1936, pp. 136-158.

Bloom, Benjamin S. (ed.) Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York, David McKay Company, Inc., 1956.

Bostic, Doyle L. A Developmental Approach for Introducing the Concept of Proof Utilizing Informal Logic. Unpublished Ph.D. Thesis, The Florida State University, 1970. DAI 31B, 5468-5469, March 1971.

Byham, F.C. Indirect Proof in Geometry from Eculid to the Present. Unpublished Ph.D. Thesis, The Ohio State University, 1969. DAI 30A, 2889, January 1970.

Cohen, Jacob. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 20: 37-45, 1960.

Commission on Mathematics of the College Entrance Examination Board. Program for College Preparatory Mathematics. New York: College Entrance Examination Board, 1959.

Conover, W.J. Practical Nonparametric Statistics. New York: John Wiley and Sons, Inc., 1971.

Davis, R.B. Discovery in Mathematics. Reading, Mass.: Addison-Wesley, 1964.

Deep, G.W. The Effects of Teaching an Explicit Unit in Logic on Students' Ability to Prove Theorems in Geometry. Unpublished Ph.D. Thesis, Florida State University, 1969.

Dienes, Z.P. The Power of Mathematics. Reading, Mass.: Addison-Wesley, 1964.

Dieudonne, J. New Thinking in School Mathematics. Paris: Organization for European Economic Corporation Publications, 1961, pp. 31-46.

Ennis, R.H. and Paulus, D.H. Critical Thinking Readiness in Grades 1-12 (PhaseI: Deductive Reasoning in Adolescence). Ithaca, N.Y.: The Cornell Critical Thinking Project (ERIC ED. 040 437), 1969.

Exner, R.M. and Rosskopf, M.F. Logic in Elementary Mathematics, New York: McGraw-Hill Book Company, Inc., 1959.

Exner, R.M. and Rosskopf, M.F. Proof. In: The Teaching of Secondary School Mathematics, NCTM Thirty-Third Yearbook. Washington, D.C.: National Council of Teachers of Mathematics, 1971, pp. 196-240.

Fawcett, H.P. The Nature of Proof. NCTM Thirteenth Yearbook. New York: Bureau of Publications, Teachers College, Columbia University, 1938.

Fehr, H.A. Needed Research in Mathematics Education. New York: Bureau of Publications, Teachers College, Columbia University, 1966.

Fetisov, A.I. Proof in Geometry. Boston: D.C. Heath and Company, 1963.

Hanna, Gerald S. Testing Student's Ability to do Geometric Proofs: A Comparison of Three Objective Item Types. Journal for Research in Mathematics Education. 2: 213-217, May, 1971.

Henle, M. On the Relation between Logic and Thinking. Psychological Review. 69: 366-378, 1962.

Hiatt, A.A. Assessing Mathematical Thinking Abilities of Sixth, Ninth and Twelfth Grade Students. Unpublished Ph.D. Thesis, The University of Californa, Berkeley, 1970. DAI: 31B, 7427-7428.

Hrabi, J.S. An Experimental Study of a Unit on Logic and Proof Strategy in the Grade Seven and Nine Mathematics Curriculum. Unpublished Ed.D Thesis, The University of Colorado, 1967. DAI: 28A, 3528, March, 1968.

Inhelder, B. and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescense. New York: Haper and Roe, 1964.

Jansson, L.C. A Comparison of Two Approaches to the Assessment of Conditional Reasoning Abilities. Journal for Research in Mathematics Education. 9 175-188, 1978.

Jansson, L.C. The Development of an Instrument to Assess Critical Thinking Ability in Mathematics. Unpublished Ph.D. Thesis, Temple University, 1970. DAI: 32A, 1383, September, 1971.

Jansson, L.C. The Development of Deductive Reasoning: A Review of the Literature. Unpublished paper presented at the Annual Meeting, American Educational Research Association, Chicago, Illinois, April, 1974.

Kilpatrick, J. Problem Solving and Creative Behavior in Mathematics. In: Studies in Mathematics. Vol. XIX, Reviews of Recent Research in Mathematics Education, School Mathematics Study Group, Stanford University, 1969.

Kline, M. Why Johnny Can't Add: The Failure of the New Math. New York, Vintage Books, 1964.

Krygowska, Z. Treatment of the Axiomatic Method in Class. In: Teaching School Mathematics. Servais, W. and Varga T. (eds.), UNESCO: Penguin Books, 1971.

King, I.L. A Formative Development of a Unit on Proof for Use in the Elementary School. Unpublished Ph.D. thesis, University of Wisconsin, 1970. ERIC: ED 040 876.

Long, R.S., Meltzer, N.S. and Hilton, P.J. Research in Mathematics Education. Educational Studies in Mathematics. 2: 446-468, 1970.

Lovell, K. The Development of the Concept of Mathematical Proof in Abler Students. In: Piagetian Cognitive - Development Research and Mathematics Education. Washington, D.C.; National Council of Teachers of Mathematics, 1971, pp. 66-80.

Mueller, David J. Logic and the Ability to Prove Theorems in Geometry. Unpublished Ph.D. dissertation, The Florida State University, 1975. DAI: 36A, 851, August, 1975.

Phillips, H.P. A Comparative Study of the Effectiveness of Two Methods of Teaching Elementary Mathematical Proofs. Unpublished Ph.D. Thesis, The Ohio State University, 1968.

Polya, G. How to Solve it. Garden City, New York: Doubleday Anchor Books, 1957.

Reynolds, J. The Development of the Concept of Mathematical Proof in Grammar School Pupils. Unpublished Ph.D. Thesis, The University of Nottingham, 1967.

Roberge, J.J. Recent Research on the Development of Children's Comprehension of Deductive Reasoning Schemes. School Science and Mathematics. 72. 197-200, 1972.

Robinson, G.E. An Investigation of Junior High School Students' Spontaneous Use of Proof to Justify Mathematical Generalizations, Unpublished Ph.D. Thesis, The University of Wisconsin, Madison, 1964.

Rosskopf, M, and Exner, R. Some Concepts of Logic and their Application in Elementary Mathematics. Mathematics Teacher, 48, 290-298, May, 1955.

Roy, G.R. The Effect of the Study of Mathematical Logic on Student Performance in Proving Theorems by Mathematical Induction. Unpublished Ph.D. Thesis, SUNY at Buffalo, 1970. DAI 31A 4045, February, 1971.

Scandura, J.M. A Theory of Mathematical Knowledge: Can Rules Account for Creative Behaviour?. Journal for Research in Mathematics Education. 2: 183-196, May, 1971.

Scandura, J.M. Mathematics: Concrete Behavioural Foundations, New York, Harper and Roe , 1971

Sher, L. Boolean Algebra as a Proof Paradigm. The Two-Year College Mathematics Journal. 9: 186-190, 1978.

Shumway, R.J. Negative Instances and Mathematics Concept Formation: A preliminary study. Journal for Research in Mathematics Education. 2: 218-227, 1971.

Skemp, R.R. The Psychology of Learning Mathematics. London: Penquin Books, 1971.

Smith, E.P. A developmental Approach to Teaching the Concept of Proof in Elementary and Secondary Mathematics. Unpublished Ph.D. Thesis, Department of Education, The Ohio State University, 1959.

Smith, E.P. and Henderson, K.B. Proof. In: The Growth of Mathematical Ideas Grades K-12. NCTM Twenty-Fourth Yearbook. Washington, D.C.: The National Council of Teachers of Mathematics, 1959.

Smith, Rolland R. Three Major Difficulties in the Hearing of Demonstrative Geometry. Menasha, Wisconsin, George Banta, 1940.

Smith, Stanley A. What does a Proof really Prove? Mathematics Teacher, 61: 483-484, 1968.

Suppes, P. Mathematical Logic for the Schools. Arithmetic Teacher. 9: 1962.

Suppes, P. and Hill, S. First Course in Mathematical Logic. Waltham, Mass.: Blaisdell Publishing Co., 1964.

Taylor, H.E. and Wade, T.L. On the Meaning of Structure in Mathematics. In: Perspectives on Secondary Mathematics Education. McIntosh, J.A. (ed.), Englewood Cliffs, N.J.: Prentice-Hall Inc., 1971.

Thom, R. 'Modern' Mathematics: An Educational and Philosophic Error? American Scientist. 59: 695-699, 1971.

Wilder, Raymond L. Introduction to the Foundations of Mathematics. New York, Wiley, 1952.

Wilson, J. Evaluation of Learning in Secondary School Mathematics. In: Handbook on Formative and Summative Evaluation of Student Learning. Bloom, B., Hastings, J.T. and Madaus, G.F. (eds.), New York: McGraw-Hill, 1971.

Witter, G.E. Mathematics: The Study of Axiom Systems. New York: Blaisdell Publishing Company, 1964.

APPENDIX

APPENDIX I

ITEMS USED IN THE PILOT STUDY

1. Mark any of the following statements that accurately describe your thoughts about mathematical proof:

______1. A mathematical proof is necessary only when you want to convince someone of the truth of a statement.

______2. A mathematical proof can establish the absolute truth of a statement.

______3. To the hypothesis and the data of the problem, and previously accepted definitions and assumptions, the rules of logic are applied to arrive at a proof for a problem.

______4. A mathematical proof of a statement is not necessary if you can give any number of instances where the statement appears to be true.

2. Note the following pattern:

$$1 = 1 \times 1 = 1^2$$

$$1 + 3 = 4 = 2 \times 2 = 2^2$$

$$1 + 3 + 5 = 9 = 3 \times 3 = 3^2$$

$$1 + 3 + 5 + 7 = 16 = 4 \times 4 = 4^4$$

(a) What is the next statement in this pattern?

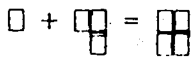
(b) Do you feel that this pattern could be continued indefinitely?

(c) Can you conceive of anyone requiring a proof in order to establish the pattern above? Why or why not?

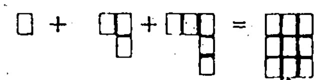
3. Note the following sequence of diagrams:

(i) 

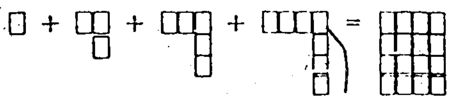
$1 = 1^2$

(ii) 

$1 + 3 = 2^2$

(iii) 

$1 + 3 + 5 = 3^2$

(iv) 

$1 + 3 + 5 + 7 = 4^2$

(a) Indicate what the next step would be in the above sequence.

(b) It seems that: The sum of the first n odd numbers is equal to n^2. For instance, $1 + 3 + 5 + 7 + 9 + 11 + 13 = 49 = 7^2$. Do you think that the above diagrams prove that this conclusion is true?

(c) Do you think that everybody would agree with you? Why or why not?

4. Note the following:

$7^2 - 6^2 = 49 - 36 = 13$ $\quad$ $3^2 - 2^2 = 5$

$5^2 - 4^2 = 9$ $\quad$ $4^2 - 3^2 = 7$

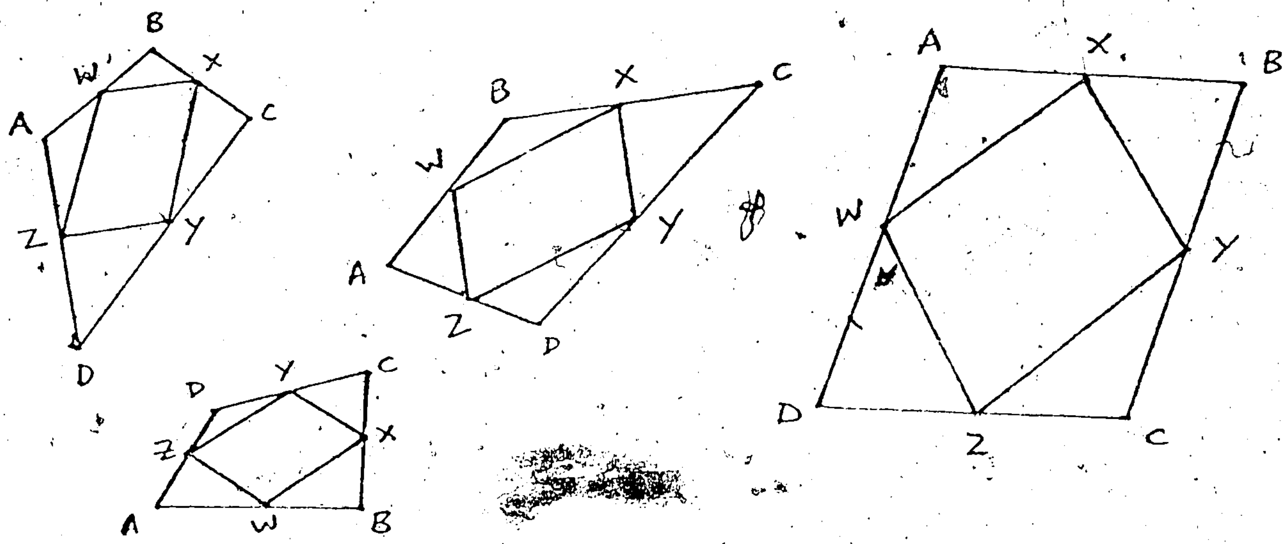
$8^2 - 7^2 = 15$ $\quad$ $11^2 - 10^2 = 21$

(a) What conclusion can you always make from these statements? Give examples if you like.

(b) Do you feel that the conclusion that you made in (a) is always valid?

4. (c) Can you conceive of anyone requiring a proof in order to establish the pattern above? Why or why not?

5. Note the following:



In all cases, ABCD is a four sided figure. The mid-points of the sides are W, X, Y and Z.

(a) What conclusion can you make about the four-sided polygon WXYZ from the diagrams above?

(b) If you made a conclusion in (a) do you think that it is always valid?

(c) Can you conceive of anyone requiring a proof in order to establish the pattern above? Why or why not?

6. Fill in the blanks:

(a) Three is a factor of 24.

(b) Three is a factor of 33.

(c) Is three a factor of 24 + 33? ________

6. (a) Six is a factor of 72.

(b) Six is a factor of 132.

(c) Is six a factor of 72 + 132? ________

(a) n is a factor of a.

(b) n is a factor of b.

(c) Is n a factor of a + b? ________

Prove it.

7. Fill in the blanks:

(a) n is a factor of a + b.

(b) Is n always a factor of a and of b? ________

Prove it.

8. Note the following:

$8^2 - 7^2 = 15$ $\quad$ $3^2 - 2^2 = 5$

$5^2 - 4^2 = 9$ $\quad$ $9^2 - 8^2 = 17$

$11^2 - 10^2 = 21$ $\quad$ $14^2 - 13^2 = 27$

(a) What conclusion can you always make about differences like those above? Give examples if you like.

(b) Do you think your conclusion in (a) is always true? Why?

9. ABCD is a rectangle. E is the mid-point of AB.

(a) The area of triangle ECD is what fractional part of the area of the whole rectangle?

(b) If your friend disagrees with you how are you going to convince him?

10. Fill in the blanks:

(a) Three is a factor of 24.

(b) Three is a factor of 33.

(c) Is three a factor of 24 + 33? ________

(a) Six is a factor of 72.

(b) Six is a factor of 132.

(c) Is six a factor of 72 + 132? ________

(a) n is a factor of a.

(b) n is a factor of b.

(c) Is n a factor of a + b? ________

How would you convince a friend that you are right if he disagrees with your answer?

11. Fill in the blanks:

(a) Six is a factor of 36 + 84.

(b) Is six a factor of 36? ________

(c) Is six a factor of 84? ________

(a) n is a factor of a + b.

(b) Is n always a factor of a and of b? ________

Convince a friend who disagrees with your last answer.

12. Triangle ABC is given such that sides AB and AC are of different length. Let p represent the measure of the angle at vertex B and q represent the measure of the angle at vertex C.

12. continued.

Suppose you wish to show that p and q have different values.

Complete the argument that starts:

"Either p and q are equal or they are different; if p and q are equal ..."

13. Two high school students, Sandra and Tom are discussing straight lines in a plane:

Sandra: "You know Tom, two distinct points in a plane determine one and only one straight line."

Tom: "Yes, of course Sandra. However, two distinct lines can intersect in two distinct points."

Sandra: "That's not true Tom. If you take two distinct lines which are not parallel, then anyone can see that they intersect in only one point."

Tom: "What you say is correct, but it does not show me that my last statement is incorrect."

If you were Sandra, what would you do at this point to show Tom that his statement is not true.

14. Sandra and Tom are discussing numbers:

Sandra: "I've been trying to show that $5 + \sqrt{2}$ is an irrational number."

Tom: "What's an irrational number?"

Sandra: "Well, do you know what a rational number is?"

Tom: "Yes, its a number which can be written as a fraction a/b, where a and b are integers and $b \neq 0$."

Sandra: "Well, an irrational number is a number which is not rational. For example, $\sqrt{2}$ is an irrational number."

14. continued.

Tom: "Oh yes, I remember. But what about $5 + \sqrt{2}$?"

Sandra: "All I know is that $5 + \sqrt{2}$ is either rational or it is irrational."

Tom: "Well, if it's irrational, your problem is solved, so suppose we say its rational."

Tom writes: Suppose $5 + \sqrt{2} = \frac{a}{b}$, where a and b are integers, and $b \neq 0$.

Then $\sqrt{2} = \frac{a}{b} - 5$

$$= \frac{a - 5b}{b}$$

Sandra: "But Tom, since a and b are integers, a − 5b is an integer and therefore $\frac{a-5b}{b}$ is a rational number. This cannot be true since $\sqrt{2}$ is an irrational number."

Tom: "You're right Sandra, but this is exactly the conclusion we must arrive at in order to show that $5 + \sqrt{2}$ is irrational."

Sandra: "Why?"

How would you reply if you were Tom?

15. Joe and Tom are discussing the inequality $20 - n > 2n - 1$.

Joe: "It is easy to show that this inequality is true when n is assigned the values 1, 2, 3 and 4."

Tom: "It's my guess that the inequality is true when n is assigned any natural number."

Joe: "I don't agree with you."

If you were Tom, what would you reply?

16. Joe and Tom are discussing straight lines in a plane. They agree that two points in a plane determine a line.

Joe: (In a prankish mood) "Two distinct lines can intersect in two distinct points."

Tom: "Joe, that's foolish."

Joe: "Well, if it is so foolish, show me why?"

If you were Tom, what would you do at this point?

17. Joe and tom are talking about divisors.

Joe: "If 3 divides some number n, then it divides any two numbers whose sum is n."

Tom: "That is not true. 3 divides 30, but it doesn't divide either 23 or 7 even though 23 + 7 = 30."

Joe: "Well, if 3 divides some number n, then three times another number, say a, equals n. That is what dividing n by 3 means." Joe then writes: 3 . a = n.

Joe: "Now, if 3 divides another number m, then 3 times some number, say b, equals m. Joe then writes: 3 . b = m.

Joe: "Now, 3 has to divide the sum of n and m."

Joe writes:

$$
\begin{aligned}
n &= 3 \cdot a \\
m &= 3 \cdot b \\
n + m &= (3 \cdot a) + (3 \cdot b) \\
&= 3(a + b).
\end{aligned}
$$

Joe: "Three divides the sum of n and m because n + m is equal to 3 times a + b. As I said before, if 3 divides some number n, then it divides any two numbers whose sum is n."

If you were Tom, how would you reply?

18. Tom: "Here is a problem for you Joe." Tom writes: Let $a + b\sqrt{2} = p/q$, where a, b, p and q are natural numbers and $q \neq 0$. The $b\sqrt{2} = p/q - a = \frac{p - aq}{q}$.

Therefore $\sqrt{2} = \frac{p - aq}{bq}$.

Now, since a, b, p and q are natural numbers, therefore both $p - aq$ and bq are natural numbers. Therefore $\sqrt{2}$ is the quotient of two natural numbers which means that $\sqrt{2}$ is not a rational number - in fact it is irrational. Therefore something must be wrong.

Tom: "What do you think is wrong Joe?"

If you were Joe, what would you reply?

19. Mathematicians have shown that all ellipses have equations of the form $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$ where A, B, C, D, E, F are constants. However, in finding equations of ellipses a mathematician found one that was not of the above form. Would you consider the first statment true or false now in view of this exception? Why?

20. Tom says. "Hey Joe, I've found a formula that will always produce prime numbers for me. It is $n^2 + n + 11$. When $n = 1$ then its value is 13. For $n = 2$ its value is 17; for $n = 3$ its value is 23. It just keeps giving me prime numbers." Joe says, "What about n = 11, then $n^2 + n + 11$ equals 143 which is not prime because 143 equals 13 times 11." Tom replies, "Well, I'm still going to say that $n^2 + n + 11$ is a formula that always produces prime numbers when n is a positive integer."

Whose side would you be on in this argument? Why.

21. According to wildlife experts Canada geese mate for life (i.e. have the same partner for life). However, a wildlife expert finds a Canada goose that left his mate for another. Would you consider the first statement true or false now in view of this exception? Why?

22. A fraction $\frac{a}{b}$ is said to be "reduced" if a and b have no common prime divisors. eg. $\frac{3}{4}$ is "reduced" since the only common divisor is 1. eg. $\frac{4}{6}$ is not "reduced" since it has 2 as a common prime divisor and $\frac{4}{6}$ and be reduced to $\frac{2}{3}$. Also $\frac{a^2}{b^2}$ is "reduced" if a^2 and b^2 have no prime common divisors. eg. $\frac{2^2}{3^2} = \frac{4}{9}$ is "reduced" however $\frac{4^2}{6^2} = \frac{16}{36}$ is not "reduced". O.K. Now, are you ready to read the following statement?

If the fraction $\frac{a^2}{b^2}$ is not reduced then the fraction $\frac{a}{b}$ is not reduced.

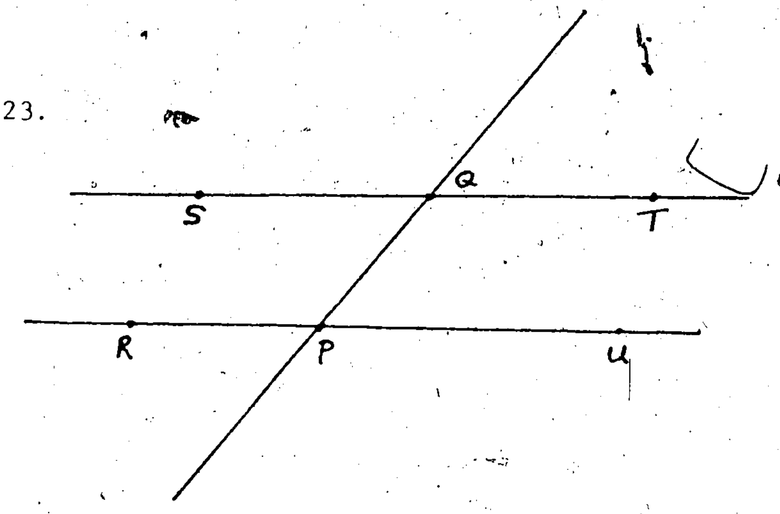
(a) Do you understand this statement?

(b) Do you believe this statement?

(c) Why do you or do you not believe this statement?

(d) Do you think a proof is necessary for the statement?

23.



Write down as many true statements as you can about the situation above.

Label your statements with one of the following:

A: An assumption that is accepted as true.

B: A definition.

P: A statement than can be proved from other statements.

24. Check any of the following statements about proof that you believe are accurate. Check as many as you wish.

_____ 1. Proofs are something found only in books.

_____ 2. If I study hard enough I can write proofs of given mathematical statements.

_____ 3. If you carefully study examples and try out values in statements, you don't need proof.

_____ 4. Nothing is true in mathematics without proof.

_____ 5. Proofs are where logic comes into a persons thinking about mathematics.

_____ 6. Proof is just a formal discussion of what you already know.

_____ 7. Proof in mathematics is simply, "that which convinces someone."

VITA

VITA

NAME: Edgar Roland Williams

PLACE OF BIRTH: Pouch Cove, Newfoundland

YEAR OF BIRTH: 1944

POST-SECONDARY EDUCATION AND DEGREES:

Memorial University of Newfoundland
St. John's, Newfoundland
1960-1965 B.Sc. (Hons.), B.Ed.
1966-1968 M.A. (Mathematics)

Dalhousie University
Halifax, Nova Scotia
Summer, 1967

HONOURS AND AWARDS

Newfoundland Government Employees Association Scholarship
Memorial University of Newfoundland
1960-1961

Undergraduate Studentships
Memorial University of Newfoundland
1963-1965

Canadian Mathematical Congress Bursary
Dalhousie University
Summer, 1967

Graduate Teaching Assistantship
University of Alberta
1971-1972

Graduate Student Assistantships
University of Alberta
1972-1974

RELATED WORK EXPERIENCE

Lecturer in Mathematics
Memorial University of Newfoundland
1965-1968

Assistant to Dean of Arts and Science
Memorial University of Newfoundland
Summer, 1968

Coordinator of Mathematics
Junior Division
Memorial University
1968-1971

Assistant Professor
Memorial University of Newfoundland
1968-1974

Mathematics Consultant (Part Time)
Newfoundland Department of Education
1969-1971

Associate Professor
Memorial University of Newfoundland
1974 -

MEMBERSHIP ON COMMITTEES

Chairman
High School Mathematics Curriculum Committee
Newfoundland Department of Education
1968-1971

Member
High School Mathematics Curriculum Committee
Newfoundland Department of Education
1974-1975

Metric Conversion Coordinator
Memorial University of Newfoundland
1975-

Program Committee
NCTM Name of Site Meeting
St. John's, Newfoundland
July, 1977

Canadian Mathematical Congress
Olympiad Committee
1977-

MEMBERSHIP IN LEARNED SOCIETIES

Canadian Mathematical Congress (1967)
Mathematical Association of America (1968)
National Council of Teachers of Mathematics (1970)
Canadian Association of Curriculum Studies (1971)

PUBLICATIONS

(1) "Mathematics Center", Newfoundland Teachers Association Journal, Vol. 58, Number 2, December, 1966.

(2) ________, Vol. 58, Number 3, February, 1967.

(3) ________, Vol. 58, Number 4, April, 1967.

(4) ________, Vol. 58, Number 5, June, 1967.

(5) ________, Vol. 59, Number 1, October, 1967.

(6) ________, Vol. 59, Number 2, December, 1967.

(7) ________, Vol. 60, Number 2, December, 1968 (with W. J. Blundon and R. H. Eddy)

(8) ________, Vol. 60, Number 3, February, 1969 (with W. J. Blundon and R. H. Eddy)

(9) ________, Vol. 60, Number 4, April, 1969 (with W. J. Blundon and R. H. Eddy)

(10) "A Teaching Guide for High School Mathematics, 1970-71", Newfoundland Department of Education, St. John's, September, 1970.

(11) "A Teaching Guide for Mathematics, Grades 7-11, 1971-72", Newfoundland Department of Education, St. John's, September, 1971.

(12) "Curriculum Evaluation", Alberta Teachers Association Magazine, Vol. 54, No. 1, September-October, 1973.

(13) "Some Lower Bounds for the Ramsay Numbers", Journal of Combinatorial Theory, Vol. 16, No. 1, January, 1974. (with H. L. Abbott).

(14) "Logic, Proof and the Role of Rigor in School Mathematics", Teaching Mathematics, Vol. 2, No. 2, December, 1974.

(15) "The High School Honours Mathematics Program", Teaching Mathematics, Vol. 4, No. 2

(16) "Honours Mathematics Update", Teaching Mathematics, Vol. 4 No. 3, February, 1977.

(17) "Update on 'Honours Mathematics Update'", Teaching Mathematics, Vol. 5, No. 2, January, 1978.

(18) "Report of the Tenth Canadian Mathematical Society Olympiad", Canadian Mathematical Society Olympiad Committee, Department of Mathematics, Memorial University of Newfoundland, September, 1978.