AN INVESTIGATION INTO THE LEVEL OF MATHEMATICAL PERFORMANCE OF STUDENTS ENROLLED IN THEIR FIRST SEMESTER AT MEMORIAL UNIVERSITY OF NEWFOUNDLAND

DAVID LANGDON GILL

AN INVESTIGATION INTO THE LEVEL OF MATHEMATICAL PERFORMANCE OF STUDENTS ENROLLED IN THEIR FIRST SEMESTER AT MEMORIAL UNIVERSITY OF NEWFOUNDLAND

BY

© DAVID LANGDON GILL, B.A., B.A.(Ed.)

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF EDUCATION

Department of Curriculum and Instruction
Memorial University of Newfoundland

October 1982

St. John's Newfoundland

ABSTRACT

The main purpose of this study was to investigate and attempt to determine the extent to which students, in their first semester at Memorial University, were able to demonstrate comprehension of selected mathematical definitions, terminology, structure and principles, as well as an ability to solve routine and non-routine mathematical problems.

A secondary purpose of the study was to determine whether, or not, relationships existed between students' performance on the above tasks and, the size of the high school attended, the mathematics programs offered by the high school attended, and the mathematics programs (honours or matriculation) studied by the student while attending high school.

In attempting to achieve these objectives, 66 items were constructed and tentatively classified as Comprehension, Application, or Problem-Solving level items. A pilot study was then designed to determine the relative suitability and difficulty of these items, and also as an aid in deciding on the format of the instrument. The final instrument consisted of two 9-item subtests referred to as form A and form B with each form containing three items at each of the levels of Comprehension, Application, and Problem-Solving. The categorization of test items at these cognitive levels was accomplished in consultation with a panel of three judges, and by using an item-classification model designed for the study. The instrument was administered to 510 students and data collected and analyzed on 335, of whom 170 completed form A and 165 completed form B.

The results indicated that neither the size of the school, nor the programs offered by the school, bore any significant relationship to the students' performance on the instrument. The students who had studied honours mathematics while attending high school performed significantly better than those who had studied matriculation mathematics, with the noteable exceptions of the Application and Problem-Solving scores on form A, and the Application score on form B.

The extent to which the sample of students were able to demonstrate their comprehension of mathematical structure, concepts and procedures, their ability to apply mathematics, and their capacity to engage in productive mathematical thought, gave cause for concern. The correct-response rate to the items on both forms combined, was 40 per cent, 39 per cent, 17 per cent and 32 per cent on the Comprehension, Application, Problem-Solving and total test scores respectively.

The study concluded with a discussion of the findings, some implications for curricular reform, and also some recommendations for future research.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to the people whose co-operation and assistance made this study possible. In particular, I would like to thank:

Dr. Keith Winter for arranging the necessary leave, and for his continuous support and assistance.

Dr. David Kirby for his invaluable assistance with the initial computer programs.

Dr. Edgar Williams and Dr. Lionel Mendoza for serving on the examination committee.

Professor Mary Veitch for her co-operation and assistance.

The first-year Psychology and Mathematics professors for their willingness to co-operate in the pilot study and the administration of the instrument.

Ms. Harriett Gorman for typing the manuscript.

Very special thanks are extended to Dr. Dale Drost, the thesis supervisor, for his availability, advice, assistance, encouragement and patience, and to my wife and children, for their understanding, support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO THE PROBLEM

BACKGROUND TO THE PROBLEM

The development of students' intellectual capacities is generally accepted to be an important aim of education. The fourth, sixth and seventh aims, as listed in The Aims of Public Education for Newfoundland and Labrador Schools (1965) are as follows:

> 4. To help pupils to mature mentally.
> 6. To ensure that all pupils master the fundamental skills of learning to the limit of their abilities.
> 7. To provide opportunities for the development of pupils' abilities to think critically.

There are those who would claim that the development of the intellect is, in reality, the essence of what education is all about. In explaining how the senior high school program was developed to satisfy the aims of public education, the steering committee analyzed these aims, and printed the following statement in the Handbook for Senior High Schools of Newfoundland and Labrador (1980):

> Considering the pressures which exist to make the school responsible for all aspects of the child's development, and the limited time and resources with which the school works, it is appropriate to make the point that, while the school may work toward the individual's "best and fullest" development, its essential function is the intellectual development of the student. This function

> is expressed primarily in Objectives 4, 6, and 7. The school is responsible for providing all students with programs designed to develop literacy and numeracy and to develop thinking skills at increasingly complex and abstract levels. (p. 4)

In recommending that schools provide programs designed to develop thinking skills at increasingly complex and abstract levels, it would appear that the steering committee had conceptualized cognitive development as an hierarchically structured process. Consequently, students would progress from the simpliest of thought forms, memorization for example, through various levels of increasing complexity, to the higher levels of creativity and abstraction. From time to time attempts have been made to identify and describe the various cognitive levels. Bloom (1956) identified six major cognitive levels as follows: Knowledge (simple recall), Comprehension (lower-level understanding), Application (use of generalizations in particular situations), Analysis (breakdown of materials into component parts), Synthesis (putting together of data to form a whole), Evaluation (judgmental decision making).

Avital (1968) identified three levels of mathematical thinking: Recall or Recognition of materials in the form in which they were presented; Algorithmic Thinking, the straight forward generalization or transfer from the learned material to similar material; Open Search, the highest level of mathematical thinking. Avital associated 'open search' with Bloom's taxonomic levels of analysis and synthesis. He suggested that successful functioning at this level would manifest itself in students' ability to solve non-rountine problems, and to think in a productive, rather than a reproductive manner.

It was out of a concern for the extent to which school mathematics programs should, but possibly do not contribute substantively to students' intellectual development, particularly at the 'open search' level, that this study was undertaken.

Mathematics is a well ordered and developed body of knowledge which, although having limitations, has widespread application. It is simultaneously a foundational science, and an art form deserving of study for its own sake. The real essence of mathematics, however, is that it is a way of thinking, mathematical thinking. Wilson (1971) suggested:

> Mathematics involves a process, a way of being, which can be learned; in fact, "Doing mathematical thinking" can be given as an answer to "What is Mathematics?"

In general terms, mathematical thinking can be subdivided into two distinct styles. Firstly there is deductive or demonstrative reasoning and, secondly, inductive reasoning. Deductive reasoning, or traditional logic, had its beginning with Aristotle and is concerned with definitions, axioms and postulates which guarantee, through syllogistic inference, exactness, consistency and validity of argument. Deductive reasoning is systematic, rigorous, and is possibly best exemplified through the classic works of Euclid. It is the style of reasoning used by mathematicians and textbook writers in presenting or exhibiting mathematics as a finished product. Such formal presentations, while an integral and necessary part of mathematics, do not necessarily give the learner any insight into the birth and development of the mathematical concept or principle in question.

Inductive reasoning, on the other hand, is not concerned with rational deduction from general propositions but on gathering facts, on experimentation, on conjecturing, on testing conjectures and continuing to conjecture, on moving from the particular to the more general. The characteristic attitude of the inductive process is one of curiosity and open-mindedness. Inductive reasoning, for example, would be more concerned with discovering different and shorter proofs than accepting and trying to understand those that are already well established. Polya (1954) claimed that we support our conjectures by plausible reasoning under which he subsumed the process of induction. He said:

> Mathematics is regarded as a demonstrative science. Yet this is only one of its aspects. Finished mathematics presented in a finished form appears as purely demonstrative, consisting of proofs only. Yet mathematics in the making resembles any other human knowledge in the making. You have to guess a mathematical theorem before you prove it; you have to guess the idea of the proof before you carry through the details. You have to combine observations and follow analogies; you have to try and try again. The result of the mathematician's creative work is demonstrative reasoning, a proof; but the proof is discovered by plausible reasoning, by guessing. If the learning of mathematics reflects to any degree the invention of mathematics, it must have a place for guessing, for plausible inference. (p. vi)

In summary, inductive reasoning, as it applies to the study of mathematics, maybe described as the process characterized by such cognitive behaviours as remembering, reflecting, observing, guessing, generalizing, analogizing, analyzing and synthesizing. If Bloom's description of cognitive development is accepted, mathematics would appear to be a discipline ideally suited for the exercise and hence development of students' intellectual capacities.

The description of the inductive process contained in the previous paragraph, could also be taken as the summation of a students' deliberations in attempting to solve a non-routine or challenging problem. Most mathematicians and mathematics educators would undoubtedly agree that the mathematical process, or mathematical thinking, is indeed embodied in the process of problem solving. Polya has insisted, for many years, that problem solving be a principal ingredient of school mathematics programs. Bloom (1956) offered the following:

> Thus, it is expected that when a student encounters a new problem or situation, he will select an appropriate technique for attacking it and will bring to bear the necessary information, both facts and principles. This has been labelled "critical thinking" by some, "reflective thinking" by Dewey and others and "problem solving" by still others. (p. 38)

In adopting a position on the direction that school mathematics reform should take in the 1980s, the directors of the National Council for Teachers of Mathematics have recommended that the mathematics curriculum be organized around problem solving (NCTM, 1980).

Some would contend that mathematics programs, both traditonal and modern, have not been developed in a manner to maximize their potential effectiveness as a tool in facilitating students' full cognitive development.

Polya (1957) exclaimed:

> Mathematics "in statu nascendi", in the process of being invented, has never been presented in quite this manner to the student, or to the teacher himself, or to the general public. (p. vii)

Robinson in the forward to Avital's text (1968) claimed:

> Such evidence as has come forward in the past few years indicates that while students are obviously learning new mathematical concepts and terminology - undeniably desirable objectives in themselves - the 'modern' mathematics programs have made little impact on what has been the central goal of mathematics instruction: the development of the ability to utilize mathematical thinking in the solution of problems. And no progress will be made until we clearly conceptualize the difference between 'lower' (comprehension) and 'higher' (problem solving) modes of mathematical performance and take positive steps to identify and strengthen these latter capacities in our students.

There are also suggestions that teachers of school mathematics have either failed to adequately address the issue of cognitive development at all levels, or at best, have restricted upper-level activity for the more capable students only. Wilson (1971) claimed:

> 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 level behaviours, such as computation and comprehension. There is some logic to claiming that performance at all cognitive levels should be expected for all students. Pity the student who has never been allowed any interesting and challenging mathematics because he is a "slow learner" and hence is never expected or offered anything more than routine computation. (And pity the student who is given too much challenge and never led to success by thoughtful and well organized instruction.) (p. 650)

This study attempted to gather information, relative to levels of cognitive development in mathematics, on a sample of recent Newfoundland high school graduates. It was envisaged that the study would shed some light on the validity, or otherwise, of the above criticisms in relation to mathematics education in Newfoundland.

STATEMENT OF THE PROBLEM

The primary objective of the study was to investigate and attempt to determine the extent to which students, in their first semester at Memorial University, were able to demonstrate:

1. Comprehension of selected mathematical definitions, terminology, structure, concepts and principles.
2. Ability to respond successfully to test items categorized, for the purpose of this study as:
   a) Application
   b) Problem-solving

Secondary objectives of the study were:

3. To compare the performance, in 1 and 2 above, of students who had completed the high school honours mathematics program with those students who had completed the high school matriculation mathematics program.
4. To determine whether, or not, relationships exist between students' performance in 1 and 2 above and:
   a) the size of the high school attended
   b) the mathematics programs offered in the high school attended.

HYPOTHESES TO BE TESTED

An analysis of the data was carried out to test the following hypotheses:

H_1. There is no significant difference between the performance of students who completed the honours program and students who completed the matriculation program.

H_2. There is no significant difference between the performance of students who attended larger high schools and those who attended smaller high schools.

H_3. There is no significant difference between the performance of students who attended schools which offered both honours and matriculation programs and those which offered only the matriculation program.

DELIMITATIONS

The study sought to compile information and data only on students who completed grade eleven in Newfoundland on June, 1980. Data were actually collected on a sample of those students who entered Memorial University in September, 1980, and registered for one of the mathematics courses 101F, 1010, 1150 or 1011 and psychology 1000. The data and information were compiled during the first semester of the 1980-81 academic year (September - December) and was restricted to: the mathematics program from which the student graduated, namely, honours or matriculation; the grade eleven population of the high school attended; the mathematics course in which the student was registered at Memorial; and the score on the instrument designed for the study.

DEFINITIONS OF TERMS

The cognitive ability levels of comprehension, application and problem-solving are fully described in Chapter III.

School Size

School size was determined by the number of students writing the Newfoundland provincial examinations in grade eleven in June, 1980.

In 1974, the Department of Education of the Government of Newfoundland and Labrador introduced a tri-level high school mathematics program. This tri-level program consists of a basic stream, a matriculation stream, and an honours stream. Students from the basic stream are not eligible to enter Memorial University whereas both matriculation and honours students are.

Matriculation Mathematics Program

A three year (grades 9-11) program for students with average ability. The general objectives for this program are stated in the Mathematics Bulletin: Grades 7-11 (1980) as follows:

1. To offer a mathematics program in which essential mathematical concepts and skills are adequately presented utilizing a practical approach, with emphasis on applications and practice rather than emphasis on involved mathematical structure and terminology.
2. To provide a mathematics program which will enable students to acquire the knowledge and essential concepts and skills needed for further educational pursuits, commercial, economical and social endeavours in the life area of their choice.

Honours Mathematics Program

A three year (grades 9-11) program for students with superior mathematical ability. The general objectives for this program are stated in the Mathematics Bulletin: Grades 7-11 (1980) as follows:

1. To provide a more challenging program for the mathematically-gifted student.
2. To provide a program which emphasizes the developmental and structural components of mathematics.
3. To provide recognition of the historical milestones in the development of mathematical ideas - ideas which have helped man in solving many of his problems.
4. To provide awareness of the direct application of mathematics to behavioural, social and applied sciences.

Mathematics 101F

A non-credit preparatory course for students wishing to proceed to Mathematics 1010, whose grade XI matriculation mark and grade XI overall average are both less than 75 per cent. The course consists of remedial work on high school mathematics and some of the algebra normally covered in Mathematics 1010.

Mathematics 1010

A course in pre-calculus mathematics intended primarily for students whose grade XI matriculation mark or grade XI overall average is 75 per cent or better. Ex-honours students whose grade XI mark is less than 75 per cent may also register for this course.

Mathematics 150B(1011)

An introductory course in calculus intended primarily for students whose grade XI honours mark is 75 per cent or better. Students initially registered in this course are required to write a placement test. This test is also available to ex-honours students whose grade XI mark is less than 75 per cent.

Mathematics 1150

A course designed for students desirous of becoming primary or elementary teachers, whose grade XI matriculation or honours mark is 50 per cent or better. The emphasis is on mathematical systems and finite mathematics.

Psychology 1000

An introductory credit course designed to introduce students to the science of behaviour. The course is a prerequisite to all other Psychology courses.

Ex-honours students

Those students who completed the honours mathematics program of studies.

Ex-matriculation students

Those students who completed the matriculation mathematics program of studies.

ITEM CLASSIFICATION MODEL

The first and most critical stage of the study was the development of test items at each of the cognitive levels to be assessed. To assist in this phase of the study a classification model was formulated which, while fundamentally based on the taxonomy of Bloom (1956), resembled more closely the model due to Wilson (1971). The model was also influenced by the scheme devised by Avital (1968). Although items at the knowledge level were not included in the study, the model, which will be presented fully in chapter III, contains the major cognitive levels of Knowledge, Comprehension, Application and Problem-Solving.

SIGNIFICANCE OF THE STUDY

As intimated earlier, the motivation and impetus for this study arose from the following premises:

1. One of the most important aims of education is the full development of students' abilities to engage in independent and productive thinking.
2. Mathematics, possibly moreso than other school subjects, has considerable potential as a medium by which students' intellectual capacities can be developed.
3. "Modern mathematics" programs may not have adequately addressed the role of mathematics in the cognitive development of students.

It has already been suggested that modern mathematics programs may have failed to contribute substantially to the cognitive development of students by having neglected to include real life problems and related problem solving activities in sufficient quantity. If this is the case, there is an urgent need to either supplement the current curriculum with problem solving activities, or, as suggested by NCTM (1980), to reorganize the curriculum around problem solving. Before anything else is attempted, the existing state of affairs should be assessed. By collecting data on student performance, it was the intent of this study to examine and ascertain the current cognitive abilities of Newfoundland high school graduates who attend Memorial University.

OUTLINE OF THE REPORT

Chapter II contains a review of selected related literature and brief reports on three major mathematics achievement studies. A detailed account of the design including the formulation of a model, preparation of test items, pilot studies, testing procedures, research design, and limitations, is given in chapter III. The results of the study are given in chapter IV. These results include student responses to the test items with discussion in each case, an analysis of the data for each of the cognitive levels: Comprehension, Application and Problem-Solving, and an analysis of the data which resulted from the testing of the three hypotheses of the study. Chapter V contains a summary and discussion of the findings of the study. Also included in this chapter is a discussion of the educational implications of the findings with some recommendations for curriculum reform and further research.

CHAPTER II

REVIEW OF RELATED LITERATURE

As already indicated in chapter I, the model used in this study in classifying test items according to cognitive level, was an adaptation of the taxonomy of Bloom (1956) to the subject matter of mathematics. It was therefore necessary to find empirical evidence that Bloom's taxonomy (hereafter referred to as the Taxonomy) provides a valid theoretical framework in which academic achievement, at various cognitive levels, can be reliably ascertained. Before reporting specifically, it is perhaps useful to consider two claims and one acknowledged limitation which pertain to this study.

The Taxonomy, prepared by a committee of college and university examiners and edited by Bloom, attempted to classify and order students' observable behaviour at different cognitive levels and in terms of stated instructional objectives. In identifying the major cognitive levels of Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation, Bloom et al claimed that the cognitive processes can be placed along a cumulative and hierarchical continuum beginning with the level of knowledge and proceeding through to the level of Evaluation. The Taxonomy is therefore so ordered that successful performance at any level, implies mastery of skills and acquisition of knowledge necessary to perform at the preceeding level. Bloom et al (1956) also claimed that the Taxonomy was generally applicable to all subject matter but cautioned:

> One of the major problems in the classification of test items which this study revealed is that it is necessary in all cases to know or assume the nature of the examinees' prior educational experience. Thus, a test problem could require a very complex type of problem-solving behaviour if it is a new situation, while it may require little more than simple kind of recall if the individual has had previous learning experiences in which this very problem was analyzed and discussed. This suggests that, in general, test material can be satisfactorily classified by means of the taxonomy only when the context in which the test problems were used is known or assumed. (p. 21)

Since 1956 the Taxonomy has received widespread usage: to compare the differences between examination items and stated course objectives; to analyze teacher-made tests; to investigate the degree to which teachers verbally question their students at different cognitive levels; to classify textbook questions, exercises and problems; to formulate objectives for teaching various subject matter; to establish standards in evaluating learning outcomes. Investigations connected with the above activities are numerous, and although related to the present study, do not bare directly upon its main intent. According to Begle (1979):

> A number of studies have been carried out, using a variety of subject matters, that together demonstrate that the six levels of the Bloom Taxonomy are empirically as well as conceptually distinct. Studies of this kind, but restricted to mathematics, seem not to be known. (p. 16)

Although Begle's survey of the literature appears to have thoroughly covered the period 1960-75, it did not disclose a study by Avital (1967). This study will comprise the focal point of the literature review for two reasons: Avital's research procedures relate most closely to the design of this study, and as Begle suggested, there is a dearth of such studies pertaining to mathematics.

Avital sought answers to the following questions:

1. Can the categories of the Taxonomy be used to distinguish theoretically and to a major extent exhaustively, different behaviour patterns usually connected with the learning of mathematics?
2. Is it possible to construct a set of items following the categories of the Taxonomy to measure students' attainment in high school mathematics?
3. Can the hierarchical nature of the various categories be sustained or should the order of some pairs of categories be reversed?

In anticipation of the possibility of obtaining affirmative answers to the first three questions, Avital formulated an additional question.

4. What proportion of variance of scores in teacher-made tests in mathematics can be accounted for by students' achievement on items corresponding to the three lowest levels of the Taxonomy?

In arriving at his descriptions of the psychological mechanisms which underly the categories of the Taxonomy in relation to mathematics learning, Avital drew heavily on the writings, teachings and empirical findings of mathematicians, psychologists and mathematics educators. According to Avital, the Knowledge level covers pure recall, either rote or meaningful, whereas the Comprehension and Application levels deal in encoding-decoding from one signal system to another. Comprehension and Application also deal in low level problem solving which can be interpreted as stimulus generalization, the distinction between the two categories being the amount of novelty and remoteness of the required generalization. Analysis which includes Evaluation, is the process of breaking up a given stimulus complex into the underlying parts, and discovering the relationships among the parts, and with the whole. Synthesis is the rearrangement of the parts into a new entity which has a new, not previously seen, meaning.

Avital considered the category of synthesis to be the highest form of performance in mathematics learning. He pointed out that in the context of the Taxonomy, the category of Evaluation becomes a judgmental process carried out by the learner upon finished products. Evaluation in mathematics learning, according to Avital, is a verifying process which deals with the results obtained by the learner, or problem solver himself, and hence of a lesser order than the actual solution generating processes. Avital therefore satisfied himself that the answer to question one was affirmative. The Taxonomy did indeed provide a theoretical framework in which those behaviours peculiar to mathematics learning could be clearly distinguished. Avital concluded that psychological underpinning may be given to five out of six categories included in the Taxonomy. Avital's concern for empirical validation can best be conveyed through his own words:

> However, since education is a social science a basic point in any educational investigation is its practicability. In the case of this study the problem of its applicability is still open. Such a problem could be attacked in at least two ways, either by producing a set of hierarchically ordered objectives and actually organizing a teaching-learning experience based on these objectives, or producing a test which would be rated by a group of informed people as organized along objectives based on the Taxonomy and then applying it to a group of school children to evaluate their attainment of these objectives. The second form of attack is followed in this study. (p. 54)

Tyler (1949) extensively analyzed the relationship between curriculum construction and achievement testing. Based on this analysis, Avital concluded that the crucial problem in his study was the construction of a set of test items which would measure achievement along the various categories of the Taxonomy. The feasibility of such a construction would, in turn, support the claim that objectives directed toward the higher levels of the Taxonomy could be formulated. Fifty-eight items, together with a summary of the Taxonomy, were sent to 16 professional mathematicians. The ratings showed relatively high agreement on the categories of Knowledge, Comprehension, Application and Analysis with rater variance, in almost all cases, being less than 10 per cent of the maximum possible variance. There was less agreement obtained on the items in the Synthesis and Evaluation categories. The raters apparently could not agree that items of the multiple choice format could measure performance at the synthesis level. Some raters contended that the category of Evaluation had no distinct place in mathematical performance.

Forty-eight of the 58 items which had been rated by the professional mathematicians were put together to form the test. The test consisted of six subtests, each composed of eight items, corresponding to the six categories of the Taxonomy. The test was administered to 313 students in grades 10, 11, 12 and 13, and to 112 teacher assistants. The test showed a reliability estimate of about 0.80 as computed by the K-R-20 formula and by the stepped up odd-even split half. The results of the tests indicated a considerable decrease in mean performance with increase in the level of the Taxonomy, with one exception. The mean performance of all high school grades, with the exception of grade 13, on the Comprehension subtest was lower than their mean performance on the Application test. Avital's experience suggested that basic concepts were not comprehended at a satisfactory level, even in grade 12. Avital suggested that abuse of the idea of the spiral curriculum, at all grade levels, may partially account for this phenomenon. The results also revealed a substantial drop in performance of all grades between the categories of Analysis and Synthesis. This study, then according to Avital, supported the contentions of the authors of the Taxonomy that it has a cumulative and hierarchical structure in terms of increasing complexity. This study also showed that it is possible to get a high rater agreement on the classification of test items at levels of Knowledge through to Analysis.

The remaining studies to be reported on fall into two categories; the earlier studies of the 1960s which tend to support the claims of the Taxonomy and the later studies which tended to be more critical of its validity. A crucial step in the development of tests, in which items supposedly evoke response behaviour at various cognitive levels, is the ability of the users to appropriately classify these items. If investigators are to feel confident in their ability to classify such items, they must seek agreement with other competent people in the field. Besides Avital (1967), several other studies report on the extent to which independent judges agree on test item classification.

Stanley and Bolten (1957) selected 227 items from Gerberich's Specimen Objective Test Items and asked eight graduate students, who had studied the Taxonomy for four weeks, to independently classify the test items according to the subcategories of the Taxonomy. The agreement was approximately 60 per cent which Stanley and Bolten concluded to be reliable since each item had to be classified in the exact subcategory.

Stoker and Kropp (1964) conducted a study in which one of the research questions was: Can judges agree on the cognitive processes which an item is intended to measure? Two reading comprehension tests each consisting of 36 multiple-choice questions were constructed. Each test was based on a 900 word reading passage dealing with science. Nine judges, who were doctoral students in educational measurement and familiar with the Taxonomy, were asked to independently classify the items, five for one test and four for the other. The agreement was 77 per cent and 76 per cent. Stoker and Kropp consequently answered their research question in the affirmative.

Cox (1965) asked three judges, familiar with the Taxonomy, to classify 379 multiple-choice test items derived from an introductory natural science course. The judges initially agreed on 85 per cent of the items, and after consultation with a subject-matter expert, agreed on the remaining 15 per cent. In this study all items fell within the first four Taxonomic levels. None were classified as Synthesis or Evaluation. Cox concluded that the Taxonomy was a reliable instrument in classifying test items according to instructional objectives they were designed to measure.

Seddon (1978) attempted to appraise the findings of various investigations which dealt with validating the properties of the Taxonomy. With respect to the reliability of judges classifying test items to the Taxonomic categories, a summary of the relevant studies is presented in Table 1, taken from Seddon. Seddon went on to say:

> As mentioned previously, an obvious and important difference between Fairbrother's experiment and the other experiments was the number of judges taking part in the classification exercise. Also, although not being drawn at random, Fairbrother's judges were almost certainly more representative of a meaningful population of educators (i.e., schoolteachers) than those used in the other experiments. In fact, from the details provided in the actual reports, it is most unlikely that the judges taking part in the other experiments were representative of any population of educators who might wish to communicate with each other about the nature of the educational objectives or items being considered. It is therefore likely that Fairbrother's results and conclusions are the most generalizable to real-life educational contexts. (p. 306)

TABLE I

PERCENTAGES OF PERFECT AGREEMENT
ON JUDGES' CLASSIFICATION OF TEST ITEMS
IN THE SEDDON (1978) REPORT

	Number of judges	Number of items	Subject Matter	Percentages perfect agreement
Scannell & Stellwagen (1960)	2	?	Chemistry	90*
Tyler (1966)	2	384	Geography	75
Cox (1965)	3	379	Natural Science	85
Stoker & Kropp (1964)	4 4	36 36	Chemistry Size relations	31 31
Herron (1966)	5	83	Chemistry	40
Poole (1971) Poole (1972)	6 7	32 44	Social Studies	16 14
Fairbrother (1975)	22 22	40 40	Physics Physics	7 0

* After making allowances for the fact that some items effectively occurred more than once in the total item pool.

(p. 305)

In looking for perfect agreement, Seddon may well have set an unrealistically high standard. Stoker and Kropp (1964), already reported on in this chapter, used nine judges and not eight, as Seddon claimed. Seddon reported 31 per cent perfect agreement for Stoker and Kropp (1964) on the multiple-choice test on size relations, for example. In other words, on 11 of the 36 items all four judges agreed (31 per cent). However, on another 16 items three out of four judges agreed and thereby yielding 85 per cent agreement on 75 per cent of the items. On nine other items, only one judge differed from the other four, with six of these nine disagreements due to the same judge. On 14 of the remaining 16 items, there was agreement by three out of five judges.

Fairbrother (1975) collected poinions from 22 physics teachers on the ability being tested by each item on two multiple-choice examination papers. The papers were the 1970 and 1971 coded answer (multiple choice) papers of the Nuffield Advanced Physics examinations and consisted of 40 items each. Each judge was given a description and illustrative examples of four abilities based on the Taxonomy, Knowledge, Comprehension, Application and Analysis/Evaluation. The judges were also asked to get the abilities clear in their minds before working through each item and assigning it to an ability category. It was also suggested that it would be necessary to refer continually to the descriptions of each ability before making a decision. As reported by Seddon (1978), all 22 judges agreed that the first three items of the 1970 examination belonged to the Knowledge category. This represented seven per cent perfect agreement. There was, however, 50 per cent or better agreement on a further 24 items. Fairbrother did not report on the nature of disagreement as Stoker and

and Kropp had done. It would be informative to know if disagreement tended to be evenly distributed over all judges or if a sizeable proportion of it was due to the same few judges. Item 35, for example, was categorized by four judges as Comprehension, as Application by another 14 and as Analysis by the remaining four. The extent to which item 35 could reliably be placed in the Application section of a test would be greatly enhanced if the investigator knew, for example, that six of the disagreeing judges were consistently disagreeable, or, among the least experienced of teachers. On the 1971 test, 50 per cent or more of the judges agreed on 20 items. Fairbrother suggested that some proportion of the disagreement may have arisen because teachers interpreted instructions differently and hence may have based their analysis on different criteria. He said:

> It would have been better to gather the teachers together to discuss the interpretation of the instructions before asking them to analyze the questions. (p. 205)

According to Fairbrother, some disagreement was due to individuals solving questions differently. A problem which involves Knowledge level ability for one person involves Comprehension or some other ability for others. Item one on the 1971 paper, for example, was rated by 19 judges as Knowledge level and by the remaining three judges as Analysis. Item 24 on the same paper was rated as Knowledge by five judges, Comprehension by eight judges, Application by seven judges and Analysis by two judges. Fairbrother recommended that his research be repeated with a larger group of teachers who had been better briefed.

Studies relating to the ability of judges to independently assign test items at the various Taxonomic levels suggest that such attempts represent onerous tasks. There is a consensus that judges must be thoroughly familiar with the Taxonomy, its claims and limitations, and ideally, should be currently teaching the content upon which the test is based. It appears useful, if not essential, that after studying the pertinent literature and before proceeding with the independent placement of items, judges meet jointly or individually with the investigator(s) so that interpretations of categories, cognitive behaviours etc., become as conceptually streamlined as possible.

As discussed later in chapter III, a major concern of this study was the quality of students' responses to the Problem-Solving Items. Brief descriptions of, or references to, three major studies are also included in chapter III: The International Study of Educational Achievement (IEA), National Assessment of Educational Progress (NAEP) and the National Longitudinal Study of Mathematical Abilities (NLSMA). In the remainder of this chapter brief reports on the mathematical achievement aspects of these studies are presented. These reports pertain to "problem-solving" or "open search" activities, for the age group 16 to 18 years.

The tests of mathematics used in IEA sought to distinguish between the "higher mental processes" and the "lower mental processes". Some attempt was therefore made to determine whether or not students had learned mathematical thinking as well as the "facts" of mathematics. Population 3a was most relevant to the present study, since it consisted of students who were completing the final year of secondary education with mathematics as an integral part of their pre-university program.

In general, students of population 3a appeared to find the 69-item test difficult. The average score of the 12 participating countries was 38 per cent. The highest scoring country was Israel with 53 per cent and the lowest was the United States with 20 per cent. Of the 69 items, 41 were categorized as "lower process" and the remaining 28 were categorized as "higher process". The median correlation of the "higher" with the "lower" process scores was 0.68. According to Husén (1967 B), had this correlation been corrected for unreliability in the separate scores, it would have been in the range of 0.90 - 0.95. Husén (1967 B) went on to say:

> Thus there appears to be some slight difference in the function being measured by the scores, but the difference is slight, and the scores are hardly different enough to make their separate analysis promising. (p. 36)

The National Assessment of Educational Progress conducted its first assessment in mathematics during the academic year 1972-73. In an attempt to assess the problem-solving skills of students, several word problems were included. The achievement level of students can probably be best seen from specific examples. Of the 17-year-old population, 53 per cent were able to correctly solve the following problem:

> Weathermen estimate that the amount of water in nine inches of snow is the same as one inch of rainfall. A certain Artic island has an annual snowfall of 1602 inches. Its annual snowfall is the same as an annual rainfall of how many inches?

Only 44 per cent of 17-year-olds were able to determine the number of votes a candidate received in an election, given that he received 70 per cent of the 4200 votes cast. Twenty-eight per cent of 17-year-olds were able to calculate the area of a square of perimeter 12 inches. According to Carpenter et al (1975), the typical American student appears to have been drilled, through teaching and testing practice, to attempt to solve all problems with a one-step, answer-getting response. Carpenter et al (1975) claimed:

> They need practice with two-step problems, in which they must decide which dimensions are needed to calculate the area or perimeter and cannot plug numbers into formulas by rote. (p. 464)

For NAEP, all target populations were representative national samples. The 17-year-old sample included high school dropouts as well as early graduates.

The second assessment of the National Assessment of Educational Progress reported that performance on one-step word problems was good. These students, however, as did their counterparts who were sampled in the first assessment, had great difficulty in analyzing non-routine or multistep problems. On a problem similar to the following, only 29 per cent of the 17-year-olds were able to arrive at the correct solution.

> Lemonade costs 95 cents for one 56-ounce bottle. At the school fair Bob sold cups holding eight ounces for 20 cents each. How much money did the school make on each bottle?

A major objective of NAEP was to determine change in students' educational attainment. According to Carpenter et al (1980), it appeared that since the first assessment, the performance of 17-year-olds had declined across almost all categories of problems. Carpenter et al (1980) claimed that if any area needed to be singled out as needing urgent attention it would have to be problem-solving. The article claimed that at all age levels, and in virtually every content area, performance was extremely low on exercises requiring problem-solving or application skills. As was true of the first assessment, it appeared that students were learning mathematical skills at a rote manipulation level without understanding the concepts underlying the computation. Students in general did not appear to have any of the basic problem-solving skills. Carpenter et al (1980) claimed:

> Rather than attempting to think through a problem and figure out what is needed to be done to solve the problem, most respondents simply tried to apply a single arithmetic operation to the numbers in the problem.

Students, according to Carpenter et al (1980), were not familiar with such basic problem-solving strategies as drawing a picture or checking the reasonableness of a result.

During the early sixties, the School Mathematics Study Group (SMSG) had instituted a National Longitudinal Study of Mathematical Abilities (NLSMA). The major purpose of this study was to investigate the effects that various mathematics programs had on students' achievement at different cognitive levels, and to use this information in future revisions of mathematics curricula. Students in the Z-population during the third year of the study (1963-64) were high school seniors and numbered 15,364. Of this number, 10,874 students were enrolled in college preparatory mathematics courses. In the spring of 1963-64 these students took nine tests containing 47 items at the Comprehension level, 11 items at the Application level, and 11 items at the Analysis level. Then NLMSA examining model is described in chapter III. The raw score means for all these students on all nine scales were 58 per cent for the Comprehension items, 48 per cent for the Application items, and 66 per cent for the Analysis items. Compared to the assessments discussed earlier, the results on the Analysis items appear to be quite good. The nine tests administered to the Z-population were used to compare the achievement of five distinct textbook groups: CG1(Functions), CGZ(Trigonometry), CG3(Calculus), CG4(Advanced Algebra), and CG5(No Textbook). The CG5 group performed significantly poorer on all tests except the single 11-item Analysis test. On this test, the CG5 group, who having completed a college preparatory program were not studying any mathematics during the year of testing, performed almost as well as the highest performing CG3 group. This would suggest that the Analysis items were for the most part too easy, with some items probably too difficult for most all students tested. According to Romberg and Wilson (1969), the analysis of the Z-population testing found few surprising results. The tests were perhaps too easy as there were several cases, especially with the calculus group, that a ceiling effect may have prevented significant variation.

CHAPTER III

DESIGN AND PROCEDURES

In chapter III, the item-classification model used in the study is presented, and the procedures followed in constructing and categorizing test items are described. Also included, is a description of the pilot study, the instrumentation process, the administration of the instrument, and the sample used. The chapter concludes with a brief discussion of the methodology to be implemented in attempting to achieve both the primary and secondary objectives of the study.

ITEM CLASSIFICATION MODEL

As its primary aim, this study attempted to determine the extent to which students in their first semester at university were able to demonstrate mathematical achievement at different cognitive levels. Accompanying this objective was the assumption that modern mathematics programs had neglected to provide for the development of students' higher cognitive abilities. If this assumption were true, a descriptive model based on the actual curriculum materials used, teacher-made tests and examinations written, and the types of learning experiences provided by the schools, would not have been useful since such a model would not have included the more complex cognitive levels. What was needed was what Taylor (1965) referred to as a normative model. Such a model, having been derived from some theory, would be used as a frame of reference for confrontation with reality and would have to completely describe, in behavioural terms, the whole range of cognitive abilities as they apply to mathematics.

Many evaluation projects, which have been concerned with the multivariate nature of mathematics learning outcomes, have adopted normative models. The International Study of Achievement in Mathematics (Husén 1967) used a model which was almost identical to the model used by the Item Banking Project in England. Wood (1968) described the latter model as follows:

A. Knowledge and information : recall of definitions, notations, concepts
B. Techniques and skill : computation, manipulation of symbols
C. Comprehension : capacity to understand problems, to translate symbolic forms, to follow and extend reasoning
D. Application : of appropriate concepts in unfamiliar mathematical situations
E. Inventiveness : reasoning creatively in mathematics

Similar models were used by the National Assessment of Educational Progress (NAEP) and the organizers of the preliminary study for the International Study of Educational Achievement (IEA). These models were based on the six major levels of cognitive thought identified by Bloom (1956) and can be described as follows:

1. Knowledge - recall of specific facts, terminology, symbols, data, methods or processes, patterns, structures, or settings.
2. Comprehension - low level understanding of material presented in a course or from a textbook etc. Students should be able to illustrate, translate, explain, interpret or extrapolate.

3. Application - use of generalizations, abstractions, principles, theories etc., in particular situations.

4. Analysis - breakdown of materials into component parts; discovery of interrelationships between these parts; discovery of relationships between the parts and the whole.

5. Synthesis - putting together of data to form a whole; establishing relationships between materials which were apparently unrelated in their originally presented form.

6. Evaluation - judgmental decision with respect to the extent that procedures or materials satisfy a given or self-developed set of criteria.

Wilson (1971) devised a model which was an extension of the model used by the School Mathematics Study Group (SMSG) in its National Longitudinal Study of Mathematical Abilities (NLSMA). Both of these models represent adaptations of Bloom's Taxonomy for the purpose of evaluating achievement in mathematics. The major objective of NLSMA was to investigate, in the long-term, the effects that various mathematics programs had on students' achievement at different cognitive levels. To achieve this objective, the NLSMA model was two-dimensional, Content by Cognitive level. Table II is a simplified version of the NLSMA model. The categories A, B, C, D and 1, 2, 3 had several sub-categories in the version actually used by the study.

TABLE II

A MODEL FOR MATHEMATICS ACHIEVEMENT (NLSMA)

Content	A.0. Computation	B.0. Comprehension	C.0. Application	D.0. Analysis
	COGNITIVE LEVELS			
1. 0 Number Systems				
2. 0 Algebra				
3. 0 Geometry				

Wilson (1971) elaborated on the intent and hence, significance of the NLSMA model.

> The model, even in this simplistic form, emphasizes the complexity of outcomes in mathematics learning. Take any mathematics topic, such as whole numbers; there are outcomes for each of the cognitive levels: there are affective outcomes as well. It is not enough that students learn to manipulate whole numbers. Rather, he must also understand something about the system of whole numbers and algorithms for their manipulation; he should be able to use whole numbers and whole number concepts in solving problems; and he should be able to use whole numbers in unfamiliar situations, to apply them in the solution of problems he has not encountered previously, or to generate new algorithms for using whole numbers in solving classes of problems. Further, it is essential that the student derive some enjoyment and appreciation of whole numbers, their use and their structure. (pp. 649-650)

Whereas the NLSMA Model was concerned only with content and cognitive behaviours, the Wilson Model extended the NLSMA model to include the affective behaviours of interests and attitudes (E.O.) and Appreciation (F.O.) (Wilson, 1971 pp. 646-647). Both the Wilson and the NLSMA models renamed Bloom's Taxonomic level of Knowledge to that of Computation and chose to include only four cognitive levels, the highest being Analysis. Wilson's (1971) interpretation of the level of Analysis was as follows:

> This behaviour level is the highest of the cognitive categories - comprising the most complex behaviours. It includes most behaviours described in the Taxonomy (Bloom 1956) as analysis, synthesis or evaluation. It includes what Avital and Shettleworth (1968) have called "open search". Here we will include nonroutine problem-solving, discovery experiences, and creative behaviour as it relates to mathematics. It differs from application-level or comprehension level behaviours in that it involves a degree of transfer to a context in which there has been no practise. (p. 662)

THE MODEL

The model in this study defined and provided psychological descriptions for the cognitive levels of Knowledge, Comprehension, Application, and Problem-Solving. Although the instrument included test items at all levels except Knowledge, the main concern of the study was with students' responses at the Problem-Solving level.

Cognitive Levels

A. Knowledge

A.1 Knowledge of basic facts

A.2 Knowledge of terminology

A.3 Knowledge of algorithms

A.4 Knowledge of concepts, principles, rules, generalizations

A.5 Knowledge of mathematical structure

B. Comprehension

B.1 Ability to carry out algorithms

B.2 Comprehension of concepts, rules, principles, generalizations

B.3 Comprehension of mathematical structure

B.4 Ability to transform problem elements from one mode to another

B.5 Ability to follow a line of reasoning

B.6 Ability to read and interpret a problem

C. Application

C.1 Ability to solve routine problems

C.2 Ability to make comparisons

C.3 Ability to analyze data

C.4 Ability to recognize patterns, isomorphisms, symmetries, or relationships

D. Problem-Solving

D.1 Ability to solve non-routine problems

D.2 Ability to discover relationships, patterns, symmetries or isomorphisms

D.3 Ability to construct proofs

D.4 Ability to criticize proofs

D.5 Ability to formulate and validate generalizations

Psychological Description of the Cognitive Levels

A. Knowledge Level

Students operating at this level are able to repeat, recall or recognize a fact, terminology, routine manipulation or algorithm, a rule, procedure, in the form in which it was explicitly taught or presented. The knowledge level can be associated with memorization more so than any other level.

B. Comprehension Level

Students operating at this level are able to generalize or carry out the process of simple transfer, give examples to illustrate knowledge level phenomenon, translate from one form (mode) to another, recognize an object as being an instance of a concept or principle, interpret a problem, follow a line of reasoning or the steps in a presented proof. The Comprehension level can be associated with understanding knowledge.

C. Application Level

Students operating at this level are able to perform a sequence of responses and select from among available alternatives; read and interpret information - manipulate that information - draw conclusions; recall relevant information - transform problem elements - manipulate these elements in a sequence - recognize a relationship; solve routine problems. The application level can be associated with the notion of using knowledge. Application level behaviours are closely related to the course of study and they deal with activities that are routine. Items similar to these application level items (not identical to them) would have been studied. The student should have solved problems requiring similar response sequences and the prerequisite knowledge information should have been studied in the course.

D. Problem-solving level

Students operating at this level are able to generalize and transfer to a context in which there has been no practice; to manipulate previously learned material in a nonroutine manner, to discover relationships among previously unrelated concepts or principles, to produce something that is entirely new to him or her, to discover new relationships rather than recognizing familiar relationships, to construct (produce) proofs rather than reproduce proofs. The Problem-solving level of behaviour can be associated with productive or creative thinking, whereas lower cognitive behaviours can be associated with reproductive thinking.

PREPARATION OF TEST ITEMS

Initially 107 test items pertaining to the mathematical content common to both the grade eleven matriculation and honours programs, were compiled and categorized by content. Included in the content categories were; real numbers, functions, solution of equations, structure, notation, trigonometry, and geometry. All items, comprising a single content category, were solved and analyzed with respect to the breadth and depth of coverage in each of the grade eleven programs. In this regard, particular attention was paid to the quantity and quality of textbook exercises and problems that students were required to complete. This analysis revealed that the solution(s) to certain items may well evoke different cognitive behaviours depending on whether the student studied honours or matriculation mathematics. For example, several items relating to the analytic geometry of the circle, would have undoubtedly have been classified as D_1 (or higher) level items for students from the matriculation program, while being C_1 (or lower) level items for ex-honours students. In this instance, the honours program provided more challenging and non-routine exercises and problems than did the matriculation program; thus giving ex-honours students an advantage. There were similar instances within other content categories. Upon completion of this analysis, 16 items were eliminated from the item bank.

As indicated in Chapter I, this study sampled the population of grade eleven students who graduated in June, 1980, and who subsequently registered at Memorial University in September, 1980, for at least one of the mathematics courses 101F, 1010, 1150, or 1011. Since there was uncertainty as to when, during the first semester, the test was to be administered, it was decided to remove all items for which it was deemed likely that students' responses would be enhanced as a result of being exposed to the content of any of these first year courses. As a result, the item bank was reduced further by 17 items.

For each of the remaining 74 items, the original solution was analyzed, alternate solutions were considered, and as a result, each item was classified as B_1, B_2, ..., D_4, D_5 in accordance with the model adopted for the study. Although care had been taken to avoid constructing items at the Knowledge level, eight were considered to be at this level and therefore eliminated from the bank. The investigator tentatively categorized the remaining 66 items as: 21 B(Comprehension), 27 C(Application), and 18 D(Problem-solving).

ITEM CLASSIFICATION BY JUDGES

In order to establish an acceptable degree of confidence in the investigator's ability to classify test items reliably, a panel of three judges was asked to independently assign some items to the appropriate level. A sample of 22 items was chosen from the 66-item bank such that seven were randomly selected from the 21 Comprehension level items, nine from the 27 Application level items and six from the 18 Problem-solving level items. This sample, together with a copy of the classification model and the corresponding psychological descriptions, was given to each of three judges. After the investigator met individually with each judge in an attempt to clarify interpretational matters relating to the model, the judges were asked to independently assign to each of the 22 items, the sub-category B_1 ... D_5 which, in their opinion, best suited the expected solution behaviour of the student.

Two of the judges were professors of mathematics with extensive teaching experience in mathematics 1010 and 1011, while the third judge was a professor of mathematics education who had taught mathematics 1010 and 1150 on several occasions. Two of the judges had been high school teachers and two had served on the provincial high school mathematics curriculum for extended periods. Besides being students of Mathematics, all three judges had pursued graduate studies in education (one at the masters level and two at the doctoral level) and were consequently familiar with the Taxonomy. Unlike many studies described in Chapter II, the judges in this study had strong theoretical and practical backgrounds. They were accustomed to teaching students at the senior high school age, were familiar with high school mathematics programs, and hence, were considered to be suitable as item-classification judges. The results of the item classification, including the initial assignment of the investigator, are presented in Table III.

TABLE III

CLASSIFICATION OF TEST ITEMS

JUDGE	ITEM																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
		Δ	●		●	●	Δ		Δ	Δ	●			Δ	●	●	Δ	Δ		Δ	●	
Y	B_1	B_4	C_1	C_1	D_1	B_4	D_1	C_1	B_5	C_4	D_2	D_1	D_1	B_2	B_2	C_1	D_4	C_2	C_2	D_1	B_4	D_1
Z	B_1	B_4	C_1	D_1	D_1	B_6	D_1	C_1	B_5	C_4	D_3	D_1	D_2	B_2	B_2	C_2	D_4	C_2	B_2	D_1	B_4	C_1
W	C_1	B_4	C_4	D_1	D_1	B_6	D_1	D_1	B_5	C_4	D_1	D_2	C_1	B_2	B_3	C_3	D_4	C_2	B_2	D_1	B_6	D_1
*X	C_1	B_4	C_1	C_3	D_2	B_6	D_1	C_1	B_5	C_4	D_2	C_2	D_1	B_2	B_3	C_3	D_4	C_2	B_2	D_1	B_4	C_2

* Investigator
Δ Exact Agreement
● Categorical Agreement

As indicated by the table, all judges and the investigator agreed exactly on eight items (36 per cent) and categorically on 15 items (68 per cent). For items 8, 13, and 19, judges Y and W concurred that their disagreement with the others was due to interpretational difficulties and would, after consultation, reclassify each item to agree with the other raters. This would then give a categorical agreement of 18 items (82 per cent). In comparing each judge with the investigator, the following emerges: with judge Y an exact agreement on 13 items (59 per cent) and a categorical agreement on 17 items (77 per cent); with judge Z an exact agreement on 13 items (59 per cent) and a categorical agreement on 19 items (86 per cent); with judge W an exact agreement on 13 items (59 per cent) and a categorical agreement on 17 items (77 per cent). It was agreed that the investigator would independently classify all items on the instrument to be designed for the study.

PILOT STUDY

The reason for conducting a pilot study was twofold; firstly, unsuitable items had to be eliminated and secondly, the format of the main instrument had to be determined. The main purpose of the study was to ascertain the cognitive abilities of the Newfoundland high school graduate at various specified levels. Consequently, test items needed to be selected on the basis of complexity rather than difficulty. Seddon (1978) had already pointed out that empirical evidence indicated very little correlation between the complexity and difficulty of test items. It was decided that the more difficult items should, and could, be eliminated by inspection. As a result, 52 of the less difficult items were chosen to comprise two pilot instruments. The literature also contained questions regarding the format in which test items were presented. The judges in the Avital (1967) study questioned the suitability of multiple-choice items in measuring achievement at the Synthesis level. Madaus et al (1973) used free-response items at both the Synthesis and Evaluation levels. It was quite possible that items at the Analysis level may also be more appropriately written in the open-ended format.

The first pilot instrument, consisting of 40 multiple choice items, was divided into five subtests with each subtest containing three items at the Comprehension level, three items at the Application level and two items at the Problem-solving level. The five sub-tests were distributed evenly (every fifth student received a different subtest) to 66 students in two sections of mathematics 1010, 53 students in two sections of mathematics 101F and 38 students in one section of mathematics 150B(1011). On the average, the 101F students answered 29 per cent or 2.3 of the questions correctly with eight items not being answered correctly by anyone; the 1010 students answered 46 per cent or 3.7 of the questions correctly with five items not being answered correctly by anyone; the 150B(1011) students answered 67 per cent or six of the questions correctly with all items answered correctly by at least 25 per cent of these students. The time required by students to complete a subtest ranged from 30 to 50 minutes.

The second pilot instrument, consisting of 12 items written in the open-ended format, was divided into three subtests with each subtest containing one Comprehension level item, one Application level item and two Problem-solving level items. These three subtests were also distributed evenly to 36 students in one section of mathematics 1010. As was the case with the other pilot instrument, student participants were informed that their knowledge of mathematics was not being examined, but rather, they were co-operating in a project designed to investigate how students attempted to solve certain problems. They were asked to work seriously at the four items for 25 minutes, showing all attempts to solve each problem, even in cases where a problem was not solved. At the end of 25 minutes, students were given two options, they could hand in their tests and solutions and leave, or, they could remain behind for a discussion of these problems and their solutions. Thirteen students remained and the results of the ensuing discussion are presented in Chapter V. While the achievement of these students on the Comprehension and Application level items approximated closely that of the other 1010 students who had completed the first pilot instrument, there was a substantive decline in achievement on the Problem-solving items. This result further questioned the feasibility of using multiple-choice items for Problem-solving level questions and suggested the need for a further look.

Five items, already used in one of the pilot instruments, were selected and prepared as multiple-choice (five-option) items. These five items constituted subtest M and the same items presented in open-ended format constituted subtest O. Subtests M and O were distributed to 37 students in one section of mathematics 1010 who completed the tests in approximately 25 minutes. The results of this testing are given in Table IV.

TABLE IV

COMPARISON OF STUDENTS' PERFORMANCES
ON MULTIPLE CHOICE VS OPEN-ENDED ITEMS

Item	Subtest M (mult-choice) n = 19		Subtest O (open-ended) n = 18	
	Number of correct responses	%	Number of correct responses	%
1	14	74	12	67
2	7	37	4	22
3	3	16	2	11
4	10	53	9	50
5	7	37	1	6

Item I was considered to be at the Comprehension level, items 3 and 4 were Application level and items 2 and 5 were Problem-Solving level items. The results of this test and the second pilot study suggested that, perhaps, for Comprehension and Application level items, the multiple-choice format provided a satisfactory means of ascertaining achievement; for Problem-solving level items, however, the open-end format appeared to provide a more realistic approach.

THE INSTRUMENT

At the completion of the pilot testing procedure an instrument consisting of two nine-item subtests, referred to as Form A and Form B, was constructed. Both forms consisted of three items at each of the cognitive levels of Comprehension, Application and Problem-solving. Form A consisted of items which pilot testing had indicated to be of an average, or less than average difficulty. Form B items, on the other hand, were considered to be of average, or for four items, greater than average difficulty. These forms were not meant to be parallel, but rather, were meant to provide for a greater coverage of items, both from the standpoint of kind and degree of difficulty. In both forms, the Comprehension and Application items were written in the multiple-choice format while the Problem-solving items were presented as open-ended items.

THE SAMPLE AND POPULATION

Originally, the instrument was to have been administered to students in several sections of mathematics 101F, 1010, 1150, and 150B (1011). Due to administrative difficulties, however, it was decided to use sections of psychology 1000 instead. This procedural change should not have dramatically altered the population since approximately 90 per cent of all first year students studied psychology. Furthermore, only students registered for one of mathematics 101F, 1010, 1150 or 1011 were included in the sample. The instrument was administered to 510 students in 18 sections of psychology 1000. The sample was deemed to be a representative cross section of first year students.

Since one of the objectives of the study was to compare achievement between ex-honours and ex-matriculation students, it was necessary to ensure that a sufficient number of such students were dispersed throughout the psychology 1000 classes. The high school records of 352 students registered in eight sections of psychology 1000, were inspected; 109 students (31 per cent) had studied honours mathematics with the remaining 243 (69 per cent) having studied matriculation mathematics. This information, made available through the Office of Junior Studies, indicated that sampling in this manner, should provide a sufficient number of ex-honours students.

Since testing had been delayed to mid-semester, it was decided to ask the participating psychology instructors to administer the instrument. It was agreed that the rapport, usually well established between teacher and student at this point in time, should be a positive influence in attempting to obtain a high degree of cooperation. All eight participating instructors consented to be examiners and agreed to the following procedures.

PROCEDURE FOR TEST ADMINISTRATION

At the beginning of each period, the examiner made a short presentation addressing the importance of, and the need for, continual educational research. In preliminary briefing sessions, most instructors pointed out that such discussion occurs as a part of the course, and hence, students should relate well to the request. Students were asked to co-operate fully but, should they prefer not to participate, they were informed that they should leave before the period began, or shortly thereafter. Forms A and B were distributed to alternate students and the instructions and remarks for Parts One and Two were read. Students were requested to complete the instrument if possible, otherwise to spend the entire 50-minute period attempting to do so.

RESEARCH DESIGN

The instrument consisted of two nine-item tests, form A and form B. Each form was divided into three subtests corresponding to the cognitive levels of Comprehension (items 1-3), Application (items 4-6) and Problem-solving (items 7-9). Items 1-6 were presented in the multiple-choice format while items 7-9 were open-ended.

The primary objective of the study was to ascertain the extent to which the students sampled were able to demonstrate proficiency at each level specified above. It was determined that this objective could be partially achieved by using the statistics provided by the subprogram, FREQUENCIES, of the SPSS (Statistics Packages for Social Sciences) program. In addition, it was deemed necessary to investigate the nature of students' solutions, and attempts at solution, for items 7-9 and also, the frequency of incorrect choices for items 1-6.

The secondary objective of the study consisted of analyzing the data to test the following hypotheses.

H_1. There is no significant difference between the performance of students who completed the honours program and students who completed the matriculation program.

H_2. There is no significant difference between the performance of students who attended larger high schools and those who attended smaller high schools.

H_3. There is no significant difference between the performance of students who attended schools which offered both honours and matriculation programs and those which offered only the matriculation program.

For each of the above hypotheses, an independent variable was defined. All three independent variables were subjected to multiple regression analysis by using the subprogram, REGRESSION, of SPSS. This subprogram, as part of its output, automatically supplies F tests of significance for each of the independent variables.

LIMITATIONS

The results of this study are only as valid as the items used to obtain them. Although items were constructed at three distinct levels of complexity, form A contained items considered by the investigator to be easy and straightforward, while form B contained some items, especially at the Comprehension level, considered to be of the more demanding variety. In both cases, it was felt that these items should be within the capabilities of most of the students in the sample. It is still possible, however, that the limited number of items at each cognitive level, could give an invalid impression of students' abilities and mathematical maturity.

Even though precautions were taken to ensure that students cooperated fully, there could have been those who did not treat a mathematics test being administered in a psychology class very seriously.

The findings of this study, with any ensuing conclusions and recommendations, can only be generalized to the population sampled, namely, those students who, after completing either matriculation or honours mathematics in Newfoundland as per the 1980 syllabi, entered Memorial University the following year.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

In this chapter, students' responses to forms A and B of the instrument, and data related to their mathematical background and the school attended, are presented and analyzed. The purpose of this analysis was to achieve the objectives of the study as outlined in chapters I and III. Generally stated, these objectives were: (1) to determine the level of understanding that these students were able to exhibit in relation to some basic mathematics, and also to ascertain the degree to which these same students were able to apply fundamental mathematical knowledge both in a routine and non-routine manner; (2) to test the three hypotheses as stated on page 8.

As stated in chapter III, the instrument was administered to 510 first-year students in 18 classes of psychology 1000 at Memorial University of Newfoundland. Of this number, 33 students had completed grade eleven mathematics prior to June 1980, six had completed high school mathematics elsewhere than in Newfoundland, 14 were accepted as mature students and had not completed high school mathematics at all, and 41 were not registered in any mathematics course(s). These students, numbering 94, were therefore removed from the sample. As also mentioned in chapter III, every effort was made to ensure that subjects cooperated fully. All test administrators reported that students, once having decided to spend the full 50 minute period, appeared to work quite conscientiously. To maximize confidence in the results, however, it was decided to eliminate those scripts which showed no evidence of the student having attempted to solve at least one of the three open-ended problems on the instrument. As a result, 81 scripts were eliminated and the sample was finally reduced to 335 students of whom 170 completed form A and 165 completed form B.

The primary objective of the study was to investigate the extent to which students were able to demonstrate (1) comprehension of selected mathematical definitions, terminology, structure, concepts and principles, (2) ability to respond correctly to test items categorized as Application and Problem-solving. To meet this objective two forms of the instrument were developed. Items on form A (TA) and form B (TB) were denoted TA_1, TA_2, ..., TA_9 and TB_1, TB_2, ..., TB_9 respectively. On each form, items 1-3 were Comprehension level, items 4-6 Application level, and items 7-9 Problem-solving level. Each correct response was scored one and each incorrect response was scored zero. The results for the 12 Comprehension and Application level items are presented in the order in which they appeared on the instrument, form A followed by form B. The presentation includes a statement of the item, a table showing the percentage of responses to each distractor for both the ex-matriculation group (M) and the ex-honours group (H), and finally, some discussion. The results for the six Problem-solving level items are presented and discussed in order of increasing difficulty, as determined by the proportion of students who answered each item successfully.

COMPREHENSION ITEMS

<u>Item TA_1</u>

Given that $(x - k)^2 = k^2 + 2x + x^2$, k = ------- ______________

(a) 2, (b) 1, (c) ½, (d) −1, (e) −2.

The percentages of responses to each distractor of Item TA_1 are given in Table V.

TABLE V

PERCENTAGE OF RESPONSES TO ITEM TA_1

GROUP	DISTRACTORS				
	a	b	c	d*	e
M {n(m) = 121}	4	34	7	48	7
H {n(H) = 49}	6	23	4	61	6
TOTALS	5	31	6	52	6

* correct response

Thirty (61 per cent) of the ex-honours students and 58 (48 per cent) of the ex-matriculation students selected the correct response, $k = -1$. Thirty-one percent of all students, however, determined that $k = 1$ was correct. All but 29 students eliminated the set $\{2, -2, \frac{1}{2}\}$ as possibilities for k. It would therefore appear reasonable to speculate that generally, students had realized that if $+2x$ was the second term of the trinomial perfect square $k^2 + 2x + x^2$, then the corresponding binomial expansion had to be $(1 + x)^2$, What may have been overlooked, or not understood, was that $x + 1$ written in the form $x - k$, implies that $k = -1$.

Item TA_2

If a rectangle of width W, is three times as long as it is wide, then its perimeter is -------------.

(a) 10 W, (b) 8 W, (c) 7 W, (d) 6 W, (e) 4 W

The percentages of responses to each distractor of Item TA_2 are given in table VI.

TABLE VI

PERCENTAGE OF RESPONSES TO ITEM TA_2

GROUP	DISTRACTORS

Item TA_3

The equation $\frac{x+2}{2} - \frac{x-1}{3} = 2$ *is equivalent to which one of the following* ----------

(a) $\frac{3(x+2) - 2(x-1)}{6} = 12$ (b) $3x + 6 - 2x - 2 = 12$

(c) $3x + 6 - 2x + 2 = 12$ (d) $\frac{3x + 6 - 2x - 2}{6} = 2$

(e) $3x + 6 - 2x + 2 = 2$

The percentages of responses to each distractor of Item TA_3 are given in table VII.

TABLE VII

PERCENTAGES OF RESPONSES TO ITEM TA_3

GROUP	DISTRACTORS				
	a	b	c*	d	e
M (n(M) = 121)	25	21	31	18	5
H (n(H) = 49)	10	8	76	6	0
TOTALS	21	18	44	15	2

On this item, 37 (76 per cent) of the ex-honours students and 38 (31 per cent) of the ex-matriculation students correctly determined the expression equivalent to $\frac{x+2}{2} - \frac{x-1}{3} = 2$. Incorrect distractors appeared to be chosen for two reasons: either students thought that $-\left(\frac{x-1}{3}\right)$ was equivalent to $\frac{-x-1}{3}$, or they confused finding the least common denominator with the multiplication axiom and hence equated $\frac{x+2}{2} - \frac{x-1}{3} = 2$ with $\frac{3(x+2) - 2(x-1)}{6} = 6(2)$.

Item TB_1

Given that a and b are any two real numbers ($b \neq 0$), a student was informed that $\frac{a}{b} > 0$. The student then deduced that $ab > 0$. Which one of the following statements is correct?----- ________

(a) the deduction is correct for all a and b
(b) the deduction is correct only if $a > b$
(c) the deduction is correct only if $a < b$
(d) the deduction is correct only if $a > 0$
(e) there is some other restriction on the value of a and b.

The percentages of responses to each distractor of Item TB_1 are given in table VIII.

TABLE VIII

PERCENTAGES OF RESPONSES TO ITEM TB_1

GROUP	DISTRACTORS				
	a*	b	c	d	e
M {n(M) = 128}	21	11	11	37	20
H {n(H) = 37}	41	8	0	32	19
TOTALS	26	11	8	36	19

Fifteen (41 per cent) of the ex-honours students and 26 (21 per cent) of the ex-matriculation students selected (a), the correct response. The greatest proportion of students, however, chose the distractor (d) as the correct response. It was conjectured that many of the students in the sample did not understand that $\frac{a}{b} > 0$ implies $a > 0$ and $b > 0$ OR $a < 0$ and $b < 0$ and in both cases, $ab > 0$. Negative values of the variables appeared to not have been considered. This is at a basic level of structural and symbolic comprehension. It is esentially equivalent to: (+)(+) = + and (-)(-) = +. Some students gave evidence of having used this latter rudimentary type of reasoning.

<u>Item TB_2</u>

If an operation $*$ is defined as follows: $a * b = \frac{ab}{a + b}$,

then determine $4 * (4 * 4)$ ----------- ________________

(a) $\frac{3}{4}$, (b) 1, (c) $\frac{4}{3}$, (d) 2, (e) $\frac{16}{3}$

The percentages of responses to each distractor of item TB_2 are given in table IX.

TABLE IX

PERCENTAGES OF RESPONSES TO ITEM TB_2

GROUP	DISTRACTORS				
	a	b	c*	d	e
M {n(M) = 128}	6	14	20	30	30
H {n(H) = 37}	0	0	57	11	32
TOTALS	5	11	29	23	32

Twenty-one (57 per cent) of the ex-honours students and 26 (20 per cent) of theex-matriculation students correctly determined that $\frac{4}{3}$ was the proper response. Thirty-two per cent of all students chose (e) and another 23 per cent chose (d). Some of the students who selected (d) determined $(4 * 4)$ to be 2 and never proceeded further. Twenty-two of the 50 who selected distractor (e) indicated, by their rough calculations, that they had considered $4 * (4 * 4) \equiv 4 * 4 * 4$ and extended the definition of $a * b = \frac{ab}{a + b}$ to $a * b * c = \frac{abc}{a + b + c}$.

Item TB_3

In solving the equation $(2x - 1)^2 = (2x - 3)^2$, a student chose the following procedure:

STEP 1: *Taking the square root of both sides, we have $2x - 1 = 2x - 3$.*

STEP 2: *Subtracting $2x$ from both sides, we have $-1 = -3$.*

CONCLUSION: *There is no solution for the above equation.*

Which one of the following statements is correct? ----

(a) the conclusion is correct ... or the conclusion is wrong because

(b) both sides of the equation are not equivalent

(c) we cannot extract the square root

(d) the only way to solve the equation is to open the brackets first

(e) there is some other error in the procedure

The percentages of responses to each distractor of item TB_3 are given in table X.

TABLE X

PERCENTAGES OF RESPONSES TO ITEM TB_3

GROUP	DISTRACTORS				
	a	b	c	d	e*
M {n(M) = 128}	43	16	15	10	16
H {n(H) = 37}	30	19	13	22	16
TOTALS	40	17	15	12	16

Six (16 per cent) of the ex-honours students and 20 (16 per cent) of the ex-matriculation students selected (e) as the correct response. It was anticipated that many of the students who did not understand that $(2x - 1)^2 = (2x - 3)^2$ implied $\pm (2x - 1) = \pm (2x - 3)$, would have used distractor (d) as a hint for actually checking the validity of their conclusion. It would appear that this did not happen, at least not to any significant extent. It is noted that equal proportions of ex-honours and ex-matriculation students responded correctly to this item. As pointed out in chapter I, it is claimed that the honours course provides a more challenging program for the mathematically gifted student with emphasis on the developmental and structural components of mathematics. The matriculation program on the other hand was designed for students of average ability, and emphasizes practise rather than involved mathematical structure and terminology. It was noted that 40 per cent of all students determined the presented solution to be correct.

In table XI the percentage of correct responses for each of the comprehension items, on forms A and B, for both the matriculation group (M) and the ex-honours group (H) are summarized.

TABLE XI

PERCENTAGE OF CORRECT RESPONSES ON

COMPREHENSION ITEMS

GROUP	ITEM (Per cent Correct)					
	n(TA)=170; n(H)=49; n(M)=121			n(TB)=165; n(H)=37; n(M)=128		
	TA_1	TA_2	TA_3	TB_1	TB_2	TB_3
M	48	79	31	21	20	16
H	61	79	76	41	57	16
TOTALS	52	79	44	26	29	16

If item-difficulty can be determined by using the percentage of correct responses, then the order of increasing difficulty was, TA_2, TA_1, TA_3, TB_2, TB_1, TB_3. It was noted that the performance of the ex-honours and ex-matriculation students did not differ at all on the most difficult and least difficult items TB_3 and TA_2. These results are shown in Figure I.

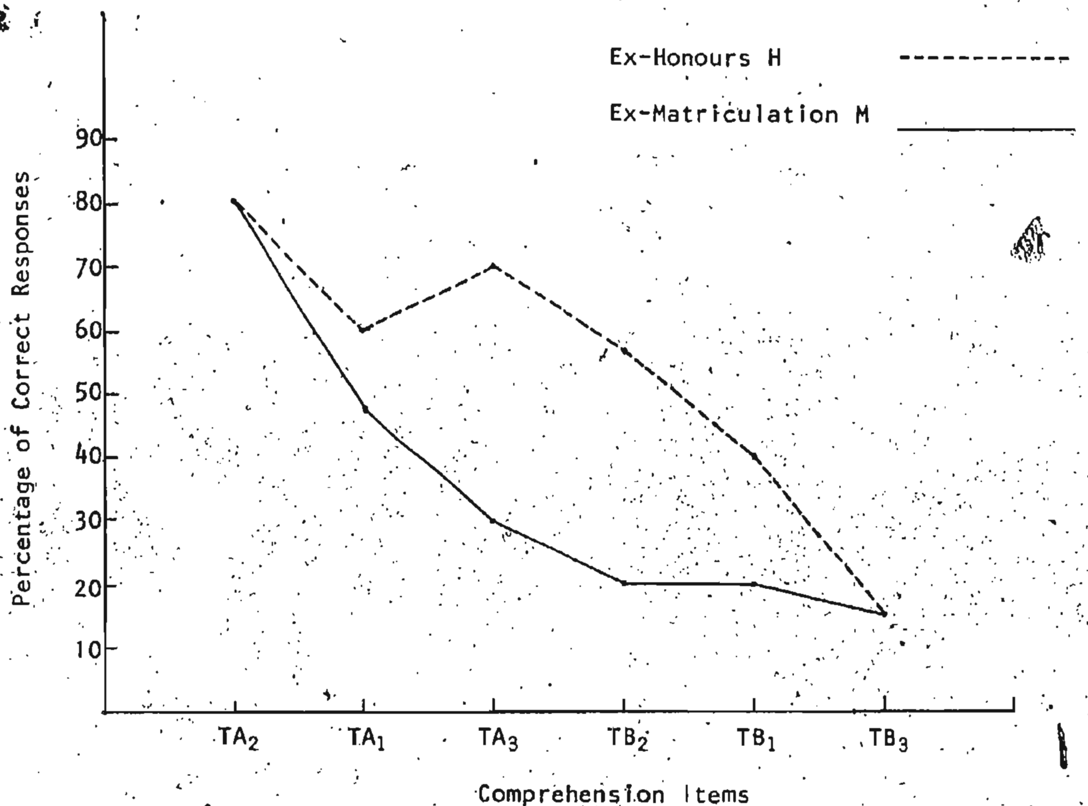

FIGURE 1

A COMPARISON OF THE PERCENTAGE OF CORRECT RESPONSES ON COMPREHENSION ITEMS BETWEEN THE EX-HONOURS AND EX-MATRICULATION STUDENTS

APPLICATION ITEMS

Item TA_4

If $f(x) = x^2 + 2mx + 3$ *and* $f(3) = 0$, *then* m = ---- __________

(a) 2, *(b)* 0, *(c)* -2, *(d)* $-\frac{1}{2}$, *(e)* -3

The percentages of responses to each distractor of item TA_4 are given in table XII.

TABLE XII

PERCENTAGES OF RESPONSES TO ITEM TA_4

GROUP	DISTRACTORS				
	a	b	c*	d	e
M {n(M) = 121}	21	7	70	1	1
H {n(H) = 49}	20	4	76	0	0
TOTALS	20	6	72	1	1

Thirty-seven (76 per cent) of the ex-honours students and 85 (70 per cent) of the ex-matriculation students responded correctly to this item. Thirty-six of the remaining students chose (a) $m = 2$ rather than $m = -2$. This could suggest difficulty in determining whether or not $9 + 6m + 3 = 0$ implies that $6m = -12$ or that $6m = 12$. The item was based on the concept of function and the solution relied on two applications of substitution. This item was the least difficult of all Application items.

Item TA_5

If the diagonal of a square is 4 centimeters, its area in square centimeters is ---------------

(a) 8, (b) 10, (c) 12, (d) 14, (e) 16

The percentages of responses to each distractor of Item TA_5 are given in table XIII.

TABLE XIII

PERCENTAGES OF RESPONSES TO ITEM TA_5

GROUP	DISTRACTORS				
	a*	b	c	d	e
M {n(M) = 121}	45	2	3	3	47
H {n(H) = 49}	51	0	4	0	45
TOTALS	46	2	4	2	46

Twenty-five (51 per cent) of the ex-honours students and 54 (45 per cent of the ex-matriculation students correctly determined the area to be eight square centimeters. It was anticipated that students would apply the pythagorean theorem which practically yields the result at once. Students were encouraged to use the reverse side of part 1 (Comprehension and Application items) for rough work and calculations. There was evidence that one student had constructed a second square through the vertices of the original square with sides perpendicular to its diagonal, but failed to realize that the area of the original square was only one-half that of the new square. It so happened that 46 per cent of all students selected the distractor (e), the response which was twice as large as the correct response.

Item TA_6

Carefully consider the following numerical statements:

(i) $\frac{2}{3} - \frac{1}{2} = \frac{1}{6}$ *(ii)* $\frac{3}{4} - \frac{2}{3} = \frac{1}{12}$ *(iii)* $\frac{4}{5} - \frac{3}{4} = \frac{1}{20}$

Which one of the following is an algebraic generalization of (i), (ii) and (iii) above? ---------

(a) $\frac{n+1}{n-1} - \frac{n}{n+1} = \frac{1}{n^2-1}$ *(b)* $\frac{2n-1}{n} - \frac{n-1}{2n-1} = \frac{1}{2n^2-n}$

(c) $\frac{n-1}{n} - \frac{n-2}{n-1} = \frac{1}{n(n+1)}$ *(d)* $\frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)}$

(e) none of these

The percentages of responses to each distractor of item TA_6 are given in table XIV.

TABLE XIV

PERCENTAGES OF RESPONSES TO ITEM TA_6

GROUP	DISTRACTORS				
	a	b	c	d*	e
M {n(M) = 121}	8	7	17	25	43
H {n(H) = 49}	6	6	16	43	29
TOTALS	7	6	17	30	40

Twenty-one (43 per cent) of the ex-honours students and 30 (25 per cent) of the ex-matriculation students selected the correct response. Sixty-six (40 per cent) of all students selected distractor (e) <u>none of these</u>, as the correct response and hence made known their inability to recognize that $\frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)}$ was a generalization of (i), (ii) and (iii).

Item TB_4

If m men can do a job in d days, then $m + r$ men can do the same job in how many days? ---------

(a) $d + r$, (b) $d - r$, (c) $\frac{d}{m + r}$, (d) $\frac{md}{m + r}$, (e) $d(m + r)$

The percentages of responses to each distractor of item TB_4 are given in table XV.

TABLE XV

PERCENTAGES OF RESPONSES TO ITEM TB_4

GROUP	DISTRACTORS				
	a	b	c	d*	e
M {n(M) = 128}	5	2	48	13	32
H {n(H) = 37}	3	0	35	27	35
TOTALS	4	2	45	16	33

Ten (27 per cent) of the ex-honours students and 17 (13 per cent) of the ex-matriculation students determined correctly that (d) was the proper response. Thirteen (35 per cent) of the ex-honours students and 61 (48 per cent) of the ex-matriculation students determined that (c) was the correct response, These responses suggested that many students did not seem to comprehend that men doing work and the time taken to do it, have an inverse relationship. The rough work revealed the following: m men can do a job in d days therefore 1 man can do the job in $\frac{d}{m}$ days and hence m + r men will do the job in $\frac{d}{m + r}$ days. Some students used more concrete examples before generalizing: 5 men can do a job in 10 days therefore 1 man can do the job in $\frac{5}{10}$ or $\frac{1}{2}$ a day and hence 16 men can do the job in $16 \times \frac{1}{2}$ or 8 days.

Item TB_5

In a group of cows and chicken, the number of legs was 14 more than twice the number of heads. The number of cows in the group is ------------------------

(a) 5, (b) 7, (c) 10, (d) 12, (e) 14

The percentages of responses to each distractor of item TB_5 are given in table XVI.

TABLE XVI

PERCENTAGES OF RESPONSES TO ITEM TB_5

GROUP	DISTRACTORS				
	a	b*	c	d	e
M {n(M) = 128}	13	55	9	10	13
H {n(H) = 37}	5	54	10	11	20
TOTALS	11	55	9	9	16

Twenty (54 per cent) of the ex-honours students and 70 (55 per cent) of the ex-matriculation students determined correctly that (b) was the proper response. The solution relied on students' comprehension-level capacity to transfer from the verbal mode to the symbolic.

Item TB_6

If when x is added to both the numerator and the denominator of the rational number a/b ($a \neq 0$, $b \neq 0$), the value of the rational number is changed to c/d, then x = ------------

(a) $\frac{bc - ad}{bd}$ (b) $\frac{bc - ad}{d - c}$ (c) $\frac{bc - ad}{c + d}$

(d) $\frac{ad - bc}{d - c}$ (e) $\frac{ad - bc}{c + d}$

The percentages of responses to each distractor of item TB_6 are given in table XVII.

TABLE XVII

PERCENTAGES OF RESPONSES TO ITEM TB_6

GROUP	DISTRACTORS				
	a	b*	c	d	e
M {n(M) = 128}	22	13	16	22	27
H {n(H) = 37}	0	27	16	48	9
TOTALS	17	16	16	28	19

Ten (27 per cent) of the ex-honours students and 17 (13 per cent) of the ex-matriculation students determined correctly that $x = \frac{bc - ad}{d - c}$. Some students set up the equation as $\frac{a}{b} + \frac{x}{x} = \frac{c}{d}$ yielding either $\frac{a}{b} + 1 = \frac{c}{d}$ or $\frac{ax + bx}{bx} = \frac{c}{d}$ with neither leading to one of the suggested options. Most of the students who showed their rough calculations, however, set up the correct equation, $\frac{a + x}{b + x} = \frac{c}{d}$. About 30 per cent of these students were able to solve the problem.

In table XVIII the percentage of correct responses for each of the Application Items, on forms A and B, for both the ex-matriculation group (M) and the ex-honours group (H) are summarized.

TABLE XVIII

PERCENTAGE OF CORRECT RESPONSES ON APPLICATION ITEMS

GROUP	ITEM (Per cent Correct)					
	n(TA)=170; n(H)=49; n(M)=121			n(TB)=165; n(H)=37; n(M)=128		
	TA_4	TA_5	TA_6	TB_4	TB_5	TB_6
M	70	45	25	13	55	13
H	76	51	43	27	54	27
TOTALS	72	46	30	16	55	16

Generally, there was less difference between the performance of the ex-matriculation and ex-honours groups for Application Items, than was the case with Comprehension Items. This is illustrated by comparing the graph in figure 2 with the one relating to Comprehension items on page 61. It is noted that on TB_5, the ex-matriculation group showed slightly superior performance to that of the ex-honours group.

In summary, the results of the Application subtest were similar to those on the Comprehension subtest. On all Application Items, the percentage of correct responses was 39 per cent whereas, the Comprehension Items yielded a success rate of 40 per cent. Students had the best success rate on two Items which required a simple well-established procedure and were less successful in cases where the procedure was slightly more involved, or left any room for interpretation.

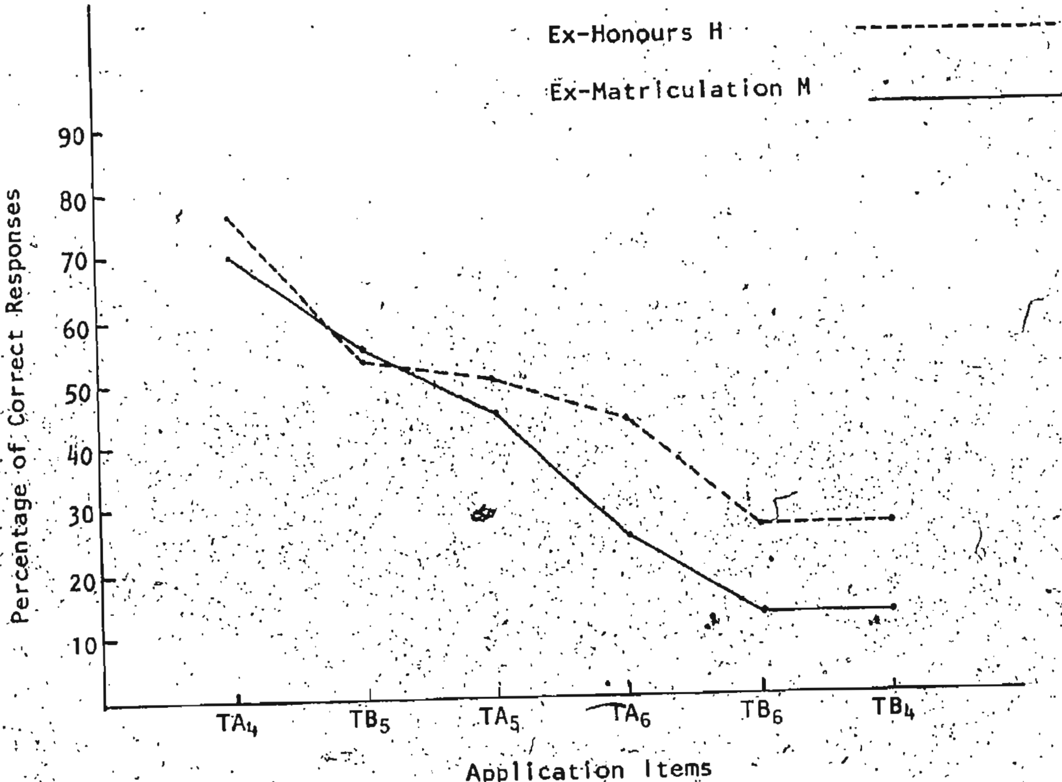

FIGURE 2

A COMPARISON OF THE PERCENTAGE OF CORRECT RESPONSES ON APPLICATION ITEMS BETWEEN THE EX-HONOURS AND EX-MATRICULATION STUDENTS

PROBLEM-SOLVING

Item TA_9

Each girl, in a group of 50, is either blonde or brunette and, either blue-eyed or brown-eyed. If 14 are blue-eyed blondes, 31 are brunettes and 18 are brown-eyed, determine the number of brown-eyed brunettes.

Thirty-four (69 per cent) of the ex-honours students and 42 (35 per cent) of the ex-matriculation students were able to solve this problem. Seventy-four of the 76 successful solvers of this problem used a logical reasoning approach, for example, if 31 are brunettes then 19 are blondes, then 19 - 14 = 5 are brown-eyed blondes, and finally 18 - 5 = 13 are brown-eyed brunettes. Although the method appears concise and efficient, some students who were unable to solve the problem tended to write out complete sentences, for example, "Since there are 31 brunettes, there are 19 blondes", "If 14 of these blondes are blue-eyed, then 5 must be brown-eyed". It is possible that many such sentences tended to confuse rather than clarify the givens. Two students used a rectangular array similar to the 3 x 3 grid shown in Figure 3.

	Blue	Brown	Totals
Blonde	14*	5°	19Δ
Brunette	18°	{13≡}	31*
TOTALS	32Δ	18*	50*

* Step 1 (givens)
Δ Step 2 (two possibilities)
o Step 3 (two possibilities)
≡ Step 4 (solution)

FIGURE 3

3 x 3 GRID SOLUTION TO PROBLEM TA_9

For this problem, it is probably the latter method that would have been chosen by the experienced problem solver. One of the students who chose this method was the only student to score $\frac{8}{9}$, the other scored $\frac{7}{9}$. The first student had $\frac{3}{3}$ on the Problem-Solving composite score, the second $\frac{2}{3}$. The unsuccesful students generally copied and re-copied the given statements and apparently did not perceive the significance of the number 50. Those who came to realize that there were (50 - 31) blondes or (50 - 18) blue-eyed girls, usually solved the problem. There were 14, however, who did not get past 19 blondes, and another five, who stopped at 32 blue-eyed girls. This problem was the least difficult of those on the instrument.

<u>Item TB_7</u>

The last digit in 3^{40} is ?

Twenty-three (48 per cent) of the ex-honours students and twenty-eight (23 per cent) of the ex-matriculation students were able to solve the problem. Most of these students were able to do so with a minimum of calculation. Several students calculated up to 3^4 and then stated that $3^{40} = (3^4)^{10} = (81)^{10} = \ldots\ldots 1$. Other students quickly discovered the pattern of repetition for the last digits and stated the last digit in a verbal fashion. The 118 students who did not solve this problem appeared not to have considered the mathematical strategy of looking for patterns. Many students expanded 3^{40} well past the 3^8 level only to give up. If 'pattern searching' is a useful problem-solving strategy, its development should be enhanced through practise. The nature of these responses provides some evidence that, in all probability, these students had received a minimum of such practise.

Item TA_7

The points $(1, Y_1)$ and $(-1, Y_2)$ are on the graph of the equation $y = ax^2 + bx + c$ and $Y_1 - Y_2 = -8$. Determine the value of b.

Nine (18 per cent) of the ex-honours students and fifteen (12 per cent) of the ex-matriculation students were able to solve this problem. Of the 142 students who did not solve it, approximately 90 students attempted to do so. Some students appeared to have perceived $Y_1 - Y_2 = -8$ to be related to the slope of a line and spent some time attempting to find an expression for $X_1 - X_2$. Other students attempted to gain some insight by graphing $Y = ax^2 + bx + c$ and the points. Most other attempts appeared to be nothing more than statements on various components of the problem but void of any sense of direction. The ingredients of the solution were Knowledge and Comprehension of the principle that if points lie on the graph represented by some equation, then the co-ordinates of the points must satisfy the equation. This should prompt $Y_1 = a + b + c$ (1) and $Y_2 = a - b + c$ (2); since $Y_1 - Y_2 = -8$ (3) is already known, this should suggest (1) - (2) yielding $Y_1 - Y_2 = 2b$ (4); by comparing (3) and (4) $b = -4$. Since students generally did not produce equation (1), it is difficult to know if they were aware, or could recall, the principle of points on a graph.

Item TB_8

A set of n numbers has a sum of S. If each number n of the set is increased by 16 and then multiplied by 7, the sum of the numbers in the new set thus obtained is ?

Seven (19 per cent) of the ex-honours students and five (four per cent) of the ex-matriculation students were able to solve this problem. Many students presented the following: $7(n + 16) = 7n + 112$ therefore the new sum $= 7S + 112$ rather than $7S + 112n$. They did not appear to realize that 112 occurred n times. Other students did not recognize, or, if so, they did not appear to understand the application of the distributive principle in this problem: $7n_1 + 112 + 7n_2 + 112 ---- + 7n_n + 112 = 7(n_1 + n_2 + ---- + n_n) + n(112)$. Instead of arriving at $7S + 112n$, they suggested $S + 7n + 112n$, as if 7 was a term being added each time as was 112.

Item TB_9

A certain number of students can be accommodated in a hostel. If 2 students share a room, then two students will be left without a room. If 3 students share a room, then two rooms will be left over. Determine the number of rooms.

Four (11 per cent) of the ex-honours students and six (five per cent) of the ex-matriculation students were able to solve this problem. Virtually all students recognized the method of solution as being that of two simultaneous equations in two variables. Generally, all students let x = the number of rooms and y = the number of students. However, that was as far as they were able to proceed. There was much evidence of many frustrated efforts, some of which, came very close. Although students had demonstrated some ability to solve routine word problems, TB_5 for example, the extent that this ability had been developed was obviously insufficient to encompass this situation. Many students expressed frustration regarding their ineffectiveness in solving this problem to the test administrators. They related well to the situation of students travelling and living in a hostel; they understood the problem of accommodation; they knew the solution strategy; they just could not set up the proper equations. Many students, however, were able to write the first equation: $2x + 2 = Y$ or $(\frac{Y}{2} = x + 1)$, with some others getting $2x - 2 = Y$. For the second equation, there seemed to be a widespread difficulty in translating: "If 3 students share a room, then 2 rooms will be left over" to "If the student body is divided by 3 $(\frac{Y}{3})$, then $\frac{Y}{3}$ is the number of required rooms which is two less $(x - 2)$ than the number of available rooms, hence $\frac{Y}{3} = x - 2$ or $3x - 6 = y$. This problem required a degree of careful analytic thinking which appeared to be lacking in the sample studies.

<u>Item TA_8</u>

The rational expression $\frac{x + 8}{x^2 + x - 2} = \frac{a}{x + c} + \frac{b}{x + d}$

Determine a, b, c, and d.

Two (four per cent) of the ex-honours students and five (four per cent) of the ex-matriculation students were able to solve this problem. Most students, who factorized $x^2 + x - 2$ as follows:
$\frac{x + 8}{(x + 2)(x - 1)} = \frac{a}{(x + c)} + \frac{b}{(x + d)}$ (1), did not see that they had already determined c and d. Generally, such students proceeded as follows:
$\frac{x + 8}{x^2 + x - 2} = \frac{ax + ad + bx + bc}{(x + c)(x + d)}$ and carried on to greater complications by cross multiplication. At (1) above, some students did realize

that $c = 2$ and $d = -1$ (some had $c = -2$, $d = 1$). In some cases, these students would arrive at: $\frac{x + 8}{x^2 + x - 2} = \frac{ax - a + bx + 2b}{(x + 2)(x-1)}$ but did not procced to $\frac{(a + b)x-a + 2b}{(x - 2)(x + 1)}$ and hence deduce that $a + b = 1$ and $-a + 2b = 8$ and hence, $b = 3$, $a = -2$. One of the common solutions was: $\frac{x}{x + 2} + \frac{8}{x - 1} = \frac{a}{x + c} + \frac{b}{x + d}$ hence $a = x$, $b = 8$, $c = 2$, $d = -1$.

As pointed out earlier, students' capacity to perform at the Problem-Solving level was the main concern of this study. In table XIX the percentages of correct responses to the six Problem-Solving items are presented.

TABLE XIX

PERCENTAGES OF CORRECT RESPONSES ON

PROBLEM-SOLVING ITEMS

GROUP	ITEM (PER CENT CORRECT)					
	n(TA)=170;	n(H)=49,	n(M)=121	n(TB)=165;	n(H)=37;	n(M)=128
	TA_7	TA_8	TA_9	TB_7	TB_8	TB_9
M	12	4	35	23	4	5
H	18	4	69	48	19	11
TOTALS	14	4	45	28	7	6

Over the entire sample and combining all six problems, 173 (17 per cent) of the solutions were correct. Only on items TA_9 and TB_7 did achievement exceed 25 per cent. When the Problem-Solving items were ordered by increasing difficulty, the differential between the levels of performance of the ex-honours and ex-matriculation groups diminished dramatically as

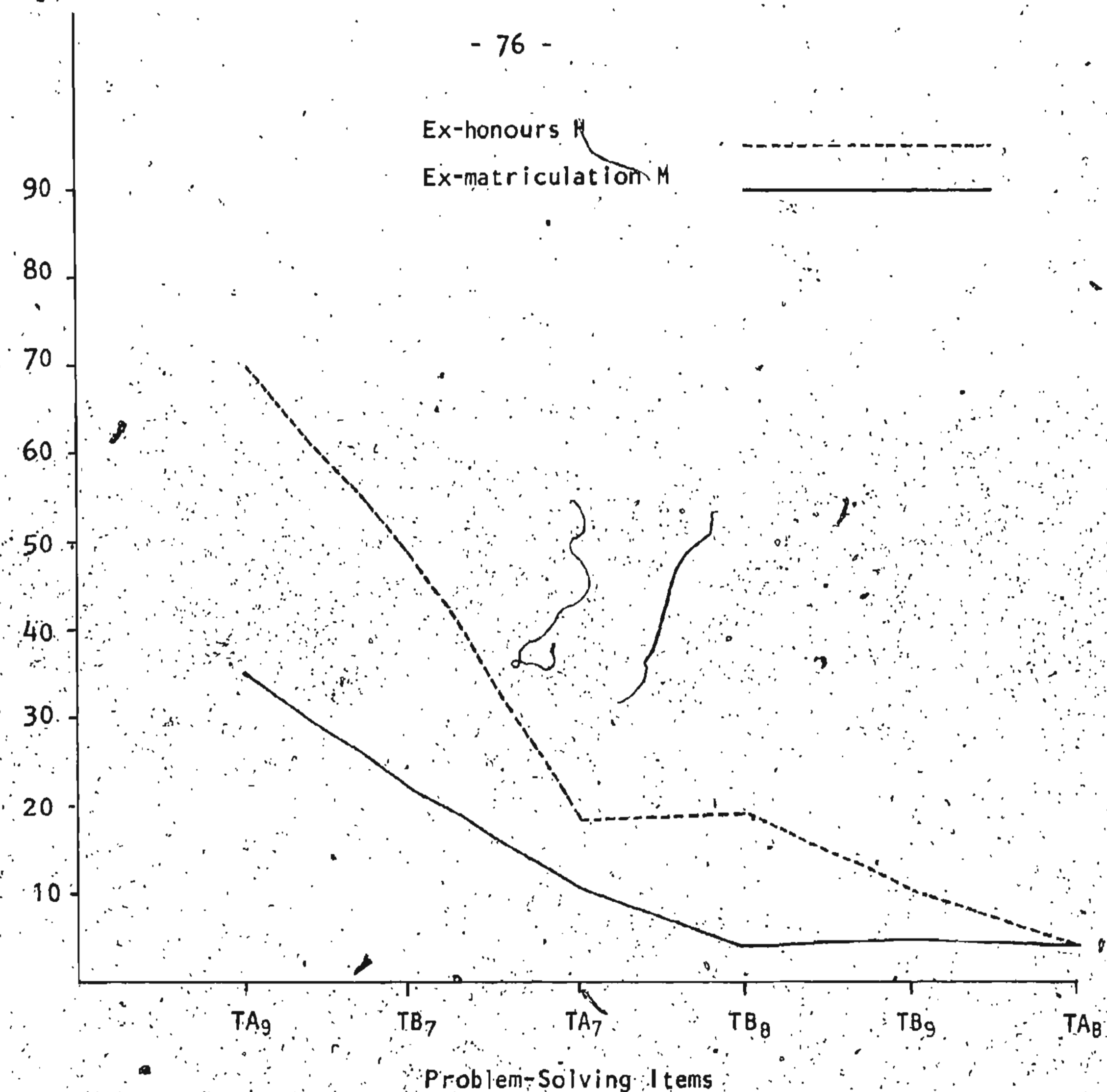

FIGURE 4

A COMPARISON OF THE PERCENTAGE OF CORRECT RESPONSES ON PROBLEM-SOLVING ITEMS BETWEEN THE EX-HONOURS AND EX-MATRICULATION STUDENTS the problems increased in difficulty. This convergent tendency is illustrated in Figure 4. This same phenomenon was also true for the more difficult Comprehension items.

As stated previously, the secondary objective of the study consisted of testing three hypotheses. The statistic used, in each case, was the F-ratio provided by the output of the subprogram REGRESSION of the SPSS program. In preparation for this regression analysis, the following dependent variables, as presented in table XX, were defined. In direct relationship to the hypotheses to be tested, the following three independent variables were defined:

PROG (student program)	...	Ex-matriculation students
		Ex-honours students
SPROG (school program)	...	Schools offering honours and matriculation
		Schools offering matriculation only
SIZE (school size)	...	grade eleven enrollment, range 013-375

TABLE XX

DEPENDENT VARIABLES

DEPENDENT VARIABLE	SYMBOL	DEFINITION
Comprehension Composite Score (Form A)	Comp (A)	$TA_1 + \ldots + TA_3$
Comprehension Composite Score (Form B)	Comp (B)	$TB_1 + \ldots + TB_3$
Application Composite Score (Form A)	Appl (A)	$TA_4 + \ldots + TA_6$
Application Composite Score (Form B)	Appl (B)	$TB_4 + \ldots + TB_6$
Problem-Solving Composite Score (Form A)	Prob (A)	$TA_7 + \ldots + TA_9$
Problem-Solving Composite Score (Form B)	Prob (B)	$TB_7 + \ldots + TB_9$
Comprehension-Application Composite Score (Form A)	Comp (A) + Appl (A)	$TA_1 + \ldots + TA_6$
Comprehension-Application Composite Score (Form B)	Comp (B) + Appl (B)	$TB_1 + \ldots + TB_6$
Total Score (Form A)	TATOT	$TA_1 + \ldots + TA_9$
Total Score (Form B)	TBTOT	$TB_1 + \ldots + TB_9$

Although PROG, SPROG, and SIZE were the only independent variables of direct concern to the study, it was decided to enter the following additional independent variables into the regression analysis.

GRD : Grade eleven mathematics grade, range 51-99

AVE : Grade eleven overall average, range 60-97

UPROG: Mathematics courses registered for at Memorial University; Math 101F, Math 1010, Math 1150, Math 150B(1011).

SEX : Male, Female

The inclusion of GRD, AVE, UPROG, SEX in the analysis can be justified in two ways. Firstly, this information was obtained quite easily and naturally in the data collecting process. Secondly, should PROG, SPROG, and SIZE turn out to be non-significant contributors to the variance of any, or all, of the dependent variables, it may be instructive to know if any of the other independent variables were significant.

Due to low enrollment in Mathematics 1150 and 150B(1011), UPROG (University Mathematics Program) was removed from the set of independent variables. Early in the multiple regression analysis procedure, it was observed that the grade eleven mathematics score (GRD) and the grade eleven overall average (AVE) had a correlation coefficient of 0.79 and consequently, AVE was removed from the analysis. Nie et al (1975) states that only one of a set of highly correlated variables should be used to represent the common underlying dimension (p. 341). As a result of this preliminary analysis, ten standard regression analyses were run; one for each of the dependent variables Comp (A), Appl (A), Prob (A), Comp (A) + Appl (A), TATOT, Comp (B), Appl (B), Prob (B), Comp (B) + Appl (B), TBTOT, with the five independent variables PROG, SPROG, SIZE, GRD and SEX. For each of the dependent variables, the three hypotheses to be tested were:

H_1. There is no significant difference between the performance of students who completed the honours program and students who completed the matriculation program. (PROG)

H_2. There is no significant difference between the performance of students who attended larger high schools and those who attended smaller high schools. (SIZE)

H_3. There is no significant difference between the performance of students who attended schools which offered both honours and matriculation programs and those which offered only the matriculation program. (SPROG)

As can be seen from Tables XXI and XXII, H_1 had to be rejected in all but three cases while both H_2 and H_3 were accepted in every case.

TABLE XXI

F-RATIOS FOR THE INDEPENDENT VARIABLES

(Form A : n = 170)

Dependent Variables	F-Ratios for Independent Variables				
	PROG	SPROG	SIZE	GRD	SEX
TATOT (TA_1 + ... + TA_9)	5.51*	0.16	3.86	34.16**	2.51
Comp + Appl (TA_1 + ... + TA_6)	7.65*	0.34	2.59	16.56**	1.55
Comp (TA_1 + TA_2 + TA_3)	5.53*	0.79	1.95	4.59*	0.01
Appl (TA_4 + TA_5 + TA_6)	3.39	0.00	1.08	15.85**	3.18
Prob (TA_7 + TA_8 + TA_9)	0.44	0.03	0.51	31.79**	1.79

* significant at the .05 level

** significant at the .01 level

TABLE XXII

F-RATIOS FOR THE INDEPENDENT VARIABLES

(Form B : n = 165)

Dependent Variables	F-Ratios for Independent Variables				
	PROG	SPROG	SIZE	GRD	SEX
TBTOT (TB_1 + ... + TB_9)	20.96**	0.12	0.25	5.63*	8.01**
Comp + Appl (TB_1 + ... + TB_6)	10.37**	0.89	0.16	1.30	3.00
Comp (TB_1 + TB_2 + TB_3)	9.62**	0.53	0.67	0.04	0.05
Appl (TB_4 + TB_5 + TB_6)	3.01	3.23	1.37	1.76	4.32*
Prob (TB_7 + TB_8 + TB_9)	9.06**	1.79	0.08	12.44**	1.68

* significant at the .05 level

** significant at the .01 level

For both forms A and B, whether students came from schools of high or low grade eleven enrollment (SIZE), or whether they came from schools offering both honours and matriculation programs or just the matriculation program alone (SPROG), made no significant contribution to the explained variance of any of the dependent variables. For Form A, ex-honours students did significantly better at the .05 level, on the total test score (TATOT), the Comprehension-Composite score (Comp), and on the combined Comprehension-Application score .Comp + Appl. These same students, however, did not achieve significantly better on the Application or the Problem-Solving sections of form A. On form B, the ex-honours students achieved significantly better, at the .01 level, on the total test (TBTOT) and all subtests with the exception of the Application composite score (Appl).

As pointed out earlier in this chapter, although the independent variables GRD (Grade eleven mathematics grade) and SEX (sex) were not of direct concern to the study, they were included in the regression analysis and hence, warrant some comment. On Form A, GRD explained a significant proportion of the variance for all dependent variables. It is also noted that, except for the Comprehension composite score, this significance was at the .01 level. On form B, GRD was significant at the .01 level for the Problem-Solving composite score and at the .05 level for the total score. The independent variable, SEX was significant at the .01 level for the total score, and at the .05 level for Application composite score on form B only. In both cases males performed at a higher level than did females.

CHAPTER V

SUMMARY, FINDINGS, DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

SUMMARY

One of the general aims of education has always been to provide opportunities for the development of students' intellectual capacities. By its nature, mathematics appears to be especially suited to promote and facilitate such development. Modern programs, however, appear to have been more concerned with structural components, rather than with utilizing problem-solving activities as a means of aiding students in the realization of their reflective and productive-thought potential. Teaching and testing practice appear to have emphasized the lower cognitive levels, such as Comprehension, while perhaps placing minimal emphasis on such higher cognitive behaviours as Analysis and Synthesis.

The primary objective of this study was to investigate and attempt to determine the extent to which recent Newfoundland secondary school graduates were able to demonstrate:

1. Comprehension of selected structural and symbolic components, concepts, principles and procedures.
2. The ability to use their Knowledge and Comprehension of mathematics in the solution of routine problems and exercises.
3. The capacity to think productively or generatively, and hence solve problems of a non-routine nature.

The secondary objective of the study was to test the following hypotheses:

H_1. There is no significant difference between the performance of students who completed the honours mathematics program and those students who completed the matriculation mathematics program.

H_2. There is no significant difference between the performance of students who attended larger high schools and those who attended smaller high schools.

H_3. There is no significant difference between the performance of students who attended schools offering matriculation mathematics only, and those who attended schools which offered both matriculation and honours mathematics.

In order to achieve these objectives, 66 items were compiled and tentatively classified at the Comprehension, Application, or Problem-Solving level. By way of establishing an acceptable degree of confidence in the investigator's ability to reliably classify test items, a panel of three judges were asked to independently assign a sample of 22 items to appropriate levels. A categorical agreement between all judges and the investigator approximated 80 per cent and was considered to be satisfactory.

A pilot study was designed to determine the relative suitability and difficulty of items, and also as an aid in deciding on the format of the instrument.

The final instrument consisted of two 9-item subtests referred to as form A and form B, with each form containing 3 items at each of the cognitive levels identified above. In both forms, all Comprehension and

Application items were written in the multiple-choice format, while the Problem-Solving items were presented open-endedly.

The instrument was administered to 510 students and data was collected and analyzed on 335, of whom 170 completed form A, and 165 completed form B.

FINDINGS

The findings of the study supported the acceptance of the hypotheses related to school size and school program. Whether or not students came from schools of high grade eleven enrollment, or whether or not they came from schools which offered both the matriculation and honours mathematics programs or just the matriculation program, made no significant difference to their total test score or any of the cognitive-level composite scores.

The hypotheses relating to the program which students studied during the high school years was, in most instances, rejected. On Form A, the ex-honours students performed significantly better than the ex-matriculation students on the total score, the Comprehension composite score, and the combination Comprehension-Application score. There was no significant difference on either the Application or Problem-Solving composite scores. On form B, ex-honours students performed significantly better on the total test score and on all composite test scores with the exception of the Application composite score, where there was no difference.

The extent to which the sample of students were able to demonstrate their comprehension of mathematical structure, concepts and procedures, their ability to apply mathematics, and their capacity to engage in productive mathematical thought, gave cause for concern. The correct response rate to the items on both forms combined, was approximately 40 per cent, 40 per cent, 17 per cent and 32 per cent on the Comprehension, Application, Problem-Solving and total test scores respectively.

DISCUSSION

As in all studies of this type, the reliability of the results are dependent on the validity of the items used in obtaining these results. The results with any ensuing discussion, conclusions and recommendations, can only be generalized to the population sampled, namely, those students who, after completing either matriculation or honours mathematics in Newfoundland as per the 1980 course outlines, entered Memorial University.

That school size (grade eleven enrollment) was not a significant contributor to the explained variance of students' performance on the instrument, was not suprising. To have expected otherwise, would have suggested an assumption that school size and the quality of mathematical experience are either directly or inversely related in a significant manner. Generally, students who study matriculation mathematics in a school with a small grade eleven enrollment have essentially the same treatment, as students doing matriculation mathematics in a school where the grade eleven enrollment may be as high as 250 students. They use the same texts, complete the same exercises and attempt the same problems. Although the qualifications of teachers vary from school to school, the incidence of enriching and supplementing the standard programs prescribed by the Provincial Department of Education, is normally low. In cases where teachers see a need to add to a program, they usually find that there is just not enough time. There are undoubtedly exceptions to such realities, however, for the students sampled in this study, it would appear that little of this type of activity had taken place. It is noted; however, that the more capable ex-matriculation students who came from the smaller-size schools did not generally have an opportunity to study the honours program. The net effect of this may well have been to inflate the general level of performance of the ex-matriculation students from smaller-size schools. In the case of ex-honours students, those who came from small and medium-size schools performed as well as those who had graduated from larger-size schools.

In the planning stages of this study, it was felt that students who had come from schools which offered the matriculation program but not the honours program, might have been at a disadvantage. In other words, ex-matriculation students who attended schools where the honours program had been offered, would have benefited from student contacts and from instruction, especially in those cases where teachers taught in both programs. One would hope that such spin-offs occur and no doubt they do, however, from the results of this study it was considered that this did not happen to any significant extent. It is possible that the content, exercises and problems of the honours program are, by their nature, not suitable or designed to improve performance on such items as were contained in the instrument.

From the results of this study, it appeared that neither the honours nor the matriculation programs are effective in improving the students' ability to perform on test items which require productive thinking, open search, or creative techniques. As already mentioned in chapter IV, the performance levels of the ex-honours and ex-matriculation students tended to approach each other as the items became more difficult. On the most difficult problem-solving item, both groups achieved at the four per cent success level. On the least difficult problem-solving level item, however, the ex-honours students achieved at the 69 per cent success level whereas the ex-matriculation students achieved at the 35 per cent success level. Considering the fact that the more capable students are generally placed in the honours stream, it would follow that had the honours program placed more emphasis on the development of creative thinking skills than the matriculation program had, the results of this study should have presumably shown the ex-honours students performing at a higher level than the ex-matriculation students for the more difficult items as well. It is quite possible, of course, that the most difficult items were too difficult for all students in the sample. As pointed out in chapter I, both programs were designed to enhance students' comprehension of structure, concepts, principles and procedures. It is not surprising then, that as in the case of school size, the programs offered by a school made no significant contribution to explaining the variance in student performance on either form of the instrument.

It is true, nonetheless, that ex-honours students performed significantly better on both forms of the instrument, and on most subtests within these forms, than did the ex-matriculation students. Generally, the more mathematically capable students are placed in the honours stream and hence, it is to be expected that their performance, on any test of mathematics achievement would be superior to that of their ex-matriculation counterparts. Although this is true, it is noted that on form A, the ex-honours students did not perform significantly better than the ex-matriculation students on the Application and Problem-Solving items. On form B the ex-honours students performed significantly better than the ex-matriculation students at all cognitive levels, with the exception of Application.

The study did not include a measure of intelligence as one of the independent variables. Ebel (1966) claimed that achievement tests constructed along the lines of the Taxonomy tend to measure general ability rather than command of Knowledge. Kropp and Stoker (1966) suggested that as the cognitive level of the test item increases, correlations between reasoning ability scores and taxonomic-level scores also increase. Assuming that general ability would have added substantially, or even moderately, to the variance accounted for on all of the total and individual cognitive-level scores of both forms A and B, the resulting relative contribution of the mathematics program studied would have been consequently reduced. It is quite conceivable that the significance of the high school program could have been rendered nonexistent, in the case of form A, and greatly reduced in the case of form B. It was indicated in chapter III that generally, the items on form B were more difficult that those of form A. Madaus et al (1973) suggested that as items become more difficult, performance tends to correlate more positively with general ability, particularly at the higher cognitive levels. If this is true, it is possible that the significance of the high school program on form B, especially the Problem-Solving composite score, was due largely, or even entirely, to the absence of a measure of general ability from the study.

In general, the level of achievement on both forms of the instrument was considered to be low. The average score on form A was 42 per cent, and on form B, the average score was substantially lower at 22 per cent. If the results are considered by cognitive level, the average for the Comprehension, Application and Problem-Solving subtests were 56 per cent, 49 per cent, and 21 per cent respectively, on form A, and 22 per cent, 28 per cent, and 14 per cent on form B.

It was stated in chapter IV, that 81 scripts were removed because the students had shown no evidence of attempting any of the Problem-Solving items. Had these scripts been included (46 for form A, 35 for form B), the average for the Problem-Solving subtest would have been reduced from 21 per cent to 16 per cent on form A, and from 14 per cent to 12 per cent on form B. It should also be noted that in developing an item bank for the study, problem books, contest books, olympiad reports, reports of national and international studies, and many textbooks were searched in order to find suitable items and determine a reasonable standard. It is felt that only the most difficult items of both forms A and B could be considered comparable with the least difficult of items normally included on the usual mathematics competitions for high school students, even at the grades nine and ten levels. If this is true, and it is assumed that the students sampled in this study are representative of the Newfoundland high school graduates who enter University, then the results are disappointing and disquieting. Noticeable in the results was the incidence of low achievement on items which contained a degree of novelty or non-routiness. These students appeared not to be adept at generalizing, and recognizing or discovering patterns. The development of these skills, which are clearly among the the rudiments of mathematical enquiry, appear not to be successfully addressed by either the matriculation or the honours program. Low achievement on the first and third items of form B indicated that these students were unable to follow a line of reasoning and analyze data, or meant that their comprehension of inequality, especially in relation to zero and square root, was at best, superficial. Low achievement on the second item of form B suggested that the students sampled, although able to transform problem elements from one mode to another, as evidenced in the fifth item of the same form, had not fully developed this capacity. An examination of the high school texts used, and the course outlines prescribed for both the matriculation and honours programs, clearly show that these skills are neither explicit nor implicit objectives. In the case of grade eleven, the provincial examination papers for the past several years, make it amply evident that the major concern is the performance of certain routine and standard tasks for which the courses of study give much practice.

Most of the deficiencies listed in the above paragraph are cognitive behaviours, which the model used by this study places at the Comprehension level. Although the ex-honours students were able to achieve at a higher level on Comprehension items than the ex-matriculation students, the depth of understanding exhibited by both groups left much to be desired. Avital's (1967) study produced similar results, even in superior classes with high academic standing. On eight Comprehension items, the percentages of correct responses for the grade 12 sample (n = 103) ranged from 35 to 87. Avital (1967) claimed:

> It can be reasonably argued that insufficient comprehension has the greatest deterimental effect on mathematics learning. A student's grasp of a concept may be decisive in his understanding of a full lecture ... Our findings provide some evidence for Professor Steels' statement that "our present crop of students can

> write better looking mathematics on paper, but are less aware of the underlying significance of what they are writing". (pp. 111-112)

Most of the Application level items were deemed to be routine, and for that reason, the result was more disappointing than the Comprehension subtest. It should be born in mind, however, that if the Taxonomy possesses the cumulative-hierarchical property as proported, superficial comprehension will also be reflected in scores on Application and higher-cognitive level items. For example, on the forth item of form B, it appeared that students were unable to solve the problem because they did not fully understand that the number of people working and the time taken to complete a job, vary inversely.

During the administration of the second pilot instrument, students were given an opportunity to remain after the testing had been completed and discuss the items and their solutions. Most of the 13 students who remained felt that their biggest drawback in successfully completing the Problem-Solving items was an inability to get started. They appeared to possess few generalized techniques for dealing with novel situations. Others lacked simple confidence in their ability and, unfortunately started with a defeatest attitude. Many students, including those who did not remain, claimed to have had no, or at least limited, experience with non-routine problems in their high school programs. Two students who remained behind after the other 11 had left, expressed disappointment and frustration at not being able to solve such problems. One commented on the practical value of studying mathematics, as they had known it, while the other vaguely recalled a problem that she had encountered. The problem was concerned with matching names to the positions on a baseball team once several relevant statements had been given. The student remembered her frustration at not being able to wade through the conglomoration of seemingly independent and confusing statements. She reflected that the rectangular-array method of tabulating data, employed in one of the problems on the instrument, would clearly be an efficient method in this case. This statement suggested a practical dimension to what could be viewed as the theoretical claims of others. Descartes, the great philosopher and mathematician once said:

> Each problem that I solved became a rule which served afterwards to solve other problems (Polya, 1954)

Polya (1964) also saw the generative value of solving any problem, no matter how semmingly trivial.

> A great discovery solves a great problem but there is a grain of discovery in the solution of any problem. Your problem may be modest; but if it challenges your curiosity and brings into play your inventive faculties, ... Such experiences at a susceptible age may create a taste for mental work and leave their imprint on mind and character for a lifetime. (p. v)

IMPLICATIONS AND RECOMMENDATIONS

From the results of this study, it was concluded that, despite the claims of modern mathematics programs, recent high school graduates seem to have only a superficial comprehension of some of the fundamental mathematical principles, concepts and procedures. The 3[illegible] students sampled in the study were comprised of 86 ex-honours students, and 249 ex-matriculation students whose average grade XI mathematics mark was approximately 80 per cent. This could well mean that teachers should consider developing more efficient methods for ensuring that students will thoroughly understand the mathematics found in their program of studies. If mathematics educators and teachers are to deepen students' comprehension of mathematics, more challenging and thought-provoking kinds of exercises will have to be included in the school mathematics programs. Avital and Shettleworth (1968) suggested:

> . . . teachers can foster good comprehension of new mathematical concepts and operations by presenting them initially, where possible, in more than one way, by relating them to a variety of situations, and by requiring students to identify their applicability to diverse material. (p. 37)

"Good" comprehension implies that, in fact, there well may be an hierarchy related to degrees of understanding. As suggested earlier, it was concluded that the students sampled possessed, at best, a superficial understanding of the mathematics examined. It is quite possible that the level of understanding attained by students was, to some extent, determined by the type of questions, exercises, and problems, that textbooks, teachers and examination papers had required of them.

It was evident that students, regardless of their high school program, experienced great difficulty in coping with non-routine problems. If mathematics teachers and program designers believe that the development of students' critical and independent thinking capacities is an important educational goal, and that higher-level mathematical thinking as manifested in the solution of non-routine mathematical problems is, in fact, critical and independent thinking, then some attempt must be made to introduce such a problem component in the curriculum. On a less philosophical note, if we as mathematics educators are concerned that students be provided with what many, including Polya, feel to be the most efficient and natural way of ensuring that students get to understand, appreciate and fully utilize mathematics, we must seriously consider implementing a problem-solving approach to the teaching of mathematics. Avital and Shettleworth (1968) made the following claim.

> A most important part of the learning process is contact with the methods by which mathematics has developed, the arrangement of familiar but diverse ideas in a new way to generate new conclusions. The best way of establishing this contact and of bringing students to understand the nature of mathematics is for students themselves to participate in mathematical problem-solving. (p. 34)

As pointed out in chapter I, the Board of Directors of NCTM have not only recommended that problem-solving be the focus of school mathematics in the 80s, but they have also suggested how this can be accomplished. Based on the findings of this study, it is suggested that those educators responsible for setting policy and determining secondary school mathematics curriculum consider adopting such a forward looking recommendation. For any such reformation to be effective, however, teachers have to be convinced that the outcomes of mathematics learning are indeed multivariate and that it is their responsibility to make sure that no one level of cognitive activity is overemphasized at the expense of the others. The history of mathematics education bears witness to the fact that school mathematics programs have tended to be polarized with respect to main emphasis. During the pre-sixties it was computation; the sixties and seventies witnessed a major emphasis on comprehension through structure and symbolism; the eighties may well see program planners looking to problem-solving for the answer.

The advent of the revised high school program in Newfoundland should afford an excellent opportunity of introducing a problem-solving component into the curriculum. As the directors of NCTM recommended, however, the problem-solving approach should permeate the whole school program. It therefore will be necessary to design summer institutes, on a provincial level, to deal solely and thoroughly with problem-solving. The writings of such mathematicians and mathematics educators as Avital, Polya, and Wilson should provide appropriate resource literature. Such institutes will clearly have to followed up by district workshops and a comprehensive in-servicing program.

The findings of the study agree very closely with those of Avital. Mathematicians and educators such as Kline and Fremont have been echoing such concerns for the past two decades. The Board of Directors of NCTM, in recommending that problem-solving be the focus of school mathematics in the 1980s, made the following claim: NCTM (1980)

> These recommendations represent both realism and responsibility. They are realistic in their attention to hard data. We are fortunate to have more information about mathematics classroom practice than we have ever had. This useful information comes principally from a series of studies funded by the National Science Foundation and two mathematics assessments of the National Assessment of Educational Progress. (p. i)

It is therefore felt that similar research, if sampled from the same population, will not add anything substantive to the findings of this study. This is not to suggest, however, that this study could not have been improved upon. Studies of a similar purpose and design might include a measure of general ability as one of the independent variables. Also more items from other areas of mathematics could be included, particularly at the Comprehension and Application levels.

The brilliant mathematician and mathematics educator, Polya, has devoted his life to the cause of problem-solving. He has been such a source of inspiration and enlightenment to so many students and teachers, that it was deemed appropriate to conclude this report with yet another quotation from his writings. Polya (1957) stated:

> Thus, a teacher of mathematics has a great opportunity. If he fills his allotted time with drilling his students in routine operations he kills their interest, hampers their intellectual development, and misuses his opportunity. But if he challenges the curiosity of his students by setting them problems proportionate to their knowledge, and helps them to solve their problems with stimulating questions, he may give them a taste for, and some means of, independent thinking. (p. v)

BIBLIOGRAPHY

Aims of Public Education for Newfoundland and Labrador. St. John's, Nfld: Department of Education, Division of Curriculum Bulletin No. 2-A, 1965.

An Agenda for Action: Recommendations for School Mathematics of the 1980s. Reston, Virginia: The National Council of Teachers of Mathematics, Inc., 1980.

Avital, S.M., Higher Level Thinking in Secondary School Students' Attainment in Mathematics. Unpublished Ph.D. thesis, The University of Toronto, 1967.

Avital, S.M., Shettleworth, S.J., Objectives for Mathematics Learning. Toronto: The Ontario Institute for Studies in Education, Bulletin No. 3, 1968.

Begle, E.G., Critical Variables in Mathematics Education. The Mathematical Association of America and the National Council of Teachers of Mathematics, 1979.

Bloom, B.S., Englehart, M.D., Furst, E.J. and others, A Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: McKay, 1956.

Bruner, J.S., Learning and Thinking. Harvard Educational Review. 1959, 29, 184-192.

Bryant, S.J., Graham, G.E. and Wiley, K.G., Nonroutine Problems in Algebra, Geometry and Trigonometry. New York: McGraw-Hill, 1965.

Carpenter, T.P., Coburn, T.G., Reys, R.E. and others, Results and Implications of the NAEP Mathematics Assessment: Secondary School. The Mathematics Teacher. 1975, 68, 453-470.

Carpenter, T.P., Corbitt, M.K., Kepner Jr, H.S. and others, Results of the Second NAEP Mathematics Assessment: Secondary School. The Mathematics Teacher. 1980, 73, 329-338.

Cox, R.C., Item Selection Techniques and Evaluation of Instructional Objectives. Journal of Educational Measurement. 1965, 2, 181-185.

Ebel, R.L., Measuring Educational Achievement. Englewood Cliffs, N.J.: Prentice Hall, 1965.

Fairbrother, R.W., The Reliability of Teachers' Judgement of the Abilities Being Tested by Multiple Choice Items. Educational Research. 1975, 17, 202-210.

Fremont, H., New Mathematics and Old Dilemnas. The Mathematics Teacher. 1967, 60, 715-719.

Greitzer, S.L., International Mathematical Olympiads 1959-1977. The Mathematical Association of America, 1978.

Guttman, L., Image Theory for the Structure of Quantitative Variates. Psychometrica. 1953, 18, 277-296.

Handbook for Senior High Schools of Newfoundland and Labrador. St. John's, Nfld: Department of Education, 1980.

Hiatt, A.A., Assessing Mathematical Thinking Abilities of Sixth, Ninth and Twelfth Grade Students. Unpublished Ph.D. thesis, The University of California, Berkeley, 1970.

Husén, T. International Study of Achievement in Mathematics; A Comparison of Twelve Countries. New York: John Wiley & Sons Inc., 1967B.

Jansson, L.C., The Development of an Instrument to Assess Critical Thinking Ability in Mathematics. Unpublished Ph.D. thesis, Temple University, 1970.

Johnson, R.E., Lendsey, L.L., Slesnick, W.E., and Bates, G.E., Algebra and Trigonometry. Don Mills: Addison-Wesley, 1967.

Kaiser, H.F., Scaling a Simplex. Psychometrica. 1962, 27, 155-162.

Kerlinger, F.N. and Pedhazur, E.J., Multiple Regression in Behavioural Research. New York: Holt, Rinehart and Winston, Inc., 1973.

Kline, M., The Ancients Versus the Moderns. The Mathematics Teacher. 1958, 51, 418-427.

Kropp, R.P. and Stoker, H.W., The construction and validation of tests of the cognitive processes described in the Taxonomy of educational objectives. (Cooperative research project No. 2117). U.S. Office of Education, 1966.

Lord, F.M. and Novick, M.R., Statistical Theories of Mental Test Scores. Reading, Mass: Addison-Wesley, 1968.

Madaus, G.F., Woods, E.M. and Nuttall, R.L., A casual model analysis of Bloom's Taxonomy. American Educational Research Journal. 1973, 10, 253-262.

Mathematics Bulletin (Grades 7-11). St. John's, Nfld: Division of Instruction, Department of Education, 1980.

Nie, H.N., Hull, C.H., Jenkins, J.G. and Others, Statistical Package For the Social Sciences, 2nd Edition. New York: McGraw-Hill, 1975.

Polya, G., How To Solve It. Garden City, New York: Doubleday Anchor Books, 1957.

Polya, G., Mathematical Discovery. New York: John Wiley & Sons Inc., Vol I, 1962; Vol II, 1965.

Polya, G., Mathematics and Plausible Reasoning. New York: John Wiley & Sons Inc., Vol I, Vol II, 1954.

Poole, R.L., Characteristics of the Taxonomy of Educational Objectives: Cognitive Domain. Psychology in the Schools, 1971, 8, 379-385.

Poole, R.L., Characteristics of the Taxonomy of Educational Objectives: Cognitive Domain - A Replication.

Rapaport, E. (trans.), Hungarian Problem Book Based on the EOTVOS COMPETITIONS, 1906-1928. New York: Random House, 1963.

Romberg, T.A., Wilson, J.W., Patterns of Mathematical Achievement in Grade 12: Z-Population. Stanford University: School Mathematics Study Group, 1969.

Salkind, C.T., Annual High School Mathematics Examinations Volume 1 (1950-1960). The Mathematical Association of America, 1961.

Salkind, C.T., The Contest Problem Book II Annual High School Contests 1961-1965). The Mathematical Association of America, 1967.

Seddon, G.M., The Properties of Bloom's Taxonomy of Educational Objectives for the Cognitive Domain. Review of Educational Research. 1978, 48, 303-23.

Stanley, J.C. and Bolton, D.T., A review of Bloom's "Taxonomy of educational objectives" and J.R. Gerberich's "Speciment objective test items, A guide to achievement test construction." Educational and Psychological Measurement. 1957, 17, 631-634.

Stedman, C.H., An Analysis of the Assumptions underlying the "Taxonomy of Educational objectives: Cognitive domain." Journal of Research in Science Teaching. 1973, 10 235-241.

Stoker, H.W. and Kropp, R.P., Measurement of Cognitive Processes. Journal of Educational Measurement. 1964, 1, 39-43.

Taylor, W.D., Decision Making and Problem-Solving. Handbook of Organization. Chicago: Rand McNally, 1965.

Tyler, R.W., Achievement testing and curriculum construction. Trends in Student Personal Work. Minneapolis: University of Minnesota Press, 1949.

Warren, R.J., Public Attitudes Towards Education in Newfoundland and Labrador. Faculty of Education, Memorial University of Newfoundland, 1978.

Wickelgren, W.A., How To Solve Problems. San Francisco: W.H. Freeman and Company, 1974.

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

Wood, R., Objectives in the Teaching of Mathematics. Educational Research. 1968, 10, 83-98.

APPENDIX A

Form A

MEMORIAL UNIVERSITY

Form A

Please complete the following mathematics survey. This survey is attempting to determine the extent to which students, entering Memorial University, are mathematically ready to begin the study of university-level mathematics courses. Your participation, today, could well lead to improvements in both the high school and first-year university programmes of the future, and especially so, in the light of the proposed new grade twelve programme. Your co-operation is therefore greatly appreciated. Thank you.

NAME ____________________ MUN NUMBER __________

PART ONE

Time: 20 minutes or less

In each case, choose the letter corresponding to the correct response.

EXAMPLE:

If $2x = 6$, then $x =$ C

(a) 1, (b) 2, (c) 3, (d) 4, (e) 5

1. If $(x - k)^2 = K^2 + 2x + x^2$, then $K =$ . . . . ____

(a) 2, (b) 1, (c) $\frac{1}{2}$, (d) -1, (e) -2

2. If a rectangle of width w, is three times as long as it is wide, then its perimeter is . . . . ____

(a) 10w, (b) 8w, (c) 7w, (d) 6w, (e) 4w

3. The equation $\frac{x + 2}{2} - \frac{x - 1}{3} = 2$ is equivalent to which one of the following . . . . ____

(a) $\frac{3(x + 2) - 2(x - 1)}{6} = 12$ (b) $3x + 6 - 2x - 2 = 12$

(c) $3x + 6 - 2x + 2 = 12$ (d) $\frac{3x + 6 - 2x - 2}{6} = 2$

(e) $3x + 6 - 2x + 2 = 2$

4. If $f(x) = x^2 + 2mx + 3$ and $f(3) = 0$, then m = ... ______

(a) 2, (b) 0, (c) -2, (d) $-\frac{1}{2}$, (e) -3

5. If the diagonal of a square is 4 centimetres, its area in square centimetres is . . . . ______

(a) 8, (b) 10, (c) 12, (d) 14, (e) 16

6. Carefully consider the following numerical statements:

(i) $\frac{2}{3} - \frac{1}{2} = \frac{1}{6}$ (ii) $\frac{3}{4} - \frac{2}{3} = \frac{1}{12}$ (iii) $\frac{4}{5} - \frac{3}{4} = \frac{1}{20}$

Which one of the following is an algebraic generalization of (i), (ii) and (iii) above? . . . . ______

(a) $\frac{n+1}{n-1} - \frac{n}{n+1} = \frac{1}{n^2-1}$ (b) $\frac{2n-1}{n} - \frac{n-1}{2n-1} = \frac{1}{2n^2-n}$

(c) $\frac{n-1}{n} - \frac{n-2}{n-1} = \frac{1}{n(n+1)}$ (d) $\frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)}$

(e) none of these

PART TWO —(You should spend approximately 10 minutes on each problem.)

Please attempt each problem and include, in the space provided, your attempt(s) to solve each problem. In case(s) where you are unable to complete the solution, it will be extremely useful to see just how far you were able to proceed.

1. The points $(1, y_1)$ and $(-1, y_2)$ are on the graph of the equation $y = ax^2 + bx + c$ and $y_1 - y_2 = -8$. Determine the value of b.

2. The rational expression $\frac{x+8}{x^2+x-2} = \frac{a}{x+c} + \frac{b}{x+d}$. Determine a, b, c and d.

3. Each girl, in a group of 50, is either blonde or brunette and, either blue-eyed or brown-eyed. If 14 are blue-eyed blondes, 31 are brunettes and 18 are brown-eyed, determine the number of brown-eyed brunettes.

APPENDIX B

Form B

MEMORIAL UNIVERSITY

Form B

Please complete the following mathematics survey. This survey is attempting to determine the extent to which students, entering Memorial University, are mathematically ready to begin the study of university-level mathematics courses. Your participation today, could well lead to improvements in both the high school and first-year university programmes of the future, and especially so, in the light of the proposed new grade twelve programme. Your co-operation is therefore greatly appreciated. Thank you.

NAME ______________________ MUN NUMBER ______________

PART ONE

Time: 20 minutes or less

In each case, choose the letter corresponding to the correct response.

EXAMPLE:

If $2x = 6$, then $x =$ . . . . ______

(a) 1, (b) 2, (c) 3, (d) 4, (e) 5

1. Given that a and b are any two real numbers ($b \neq 0$), a student was informed that $a/b > 0$. The student then deducted that $ab > 0$. Which one of the following statements is correct? . . . . ______

   (a) the deduction is correct in all of a and b
   (b) the deduction is correct only if $a > b$
   (c) the deduction is correct only if $a < b$
   (d) the deduction is correct only if $a > 0$
   (e) there is some other restriction on the value of a or b

2. If an operation $*$ is defined as follows:

   $a * b = \frac{ab}{a + b}$ then $4 * (4 * 4) =$ . . . . ______

   (a) $\frac{3}{4}$, (b) 1, (c) $\frac{4}{3}$, (d) 2, (e) $\frac{16}{3}$

3\. In solving the equation $(2x - 1)^2 = (3x - 3)^2$, a student chose the following procedure:

Step 1. Taking the square root of both sides, we have $2x - 1 = 2x - 3$.

Step 2. Subtracting $2x$ from both sides, we have $-1 = -3$.

Conclusion. There is no solution for the above equation.

Which one of the following statements is correct? . . . . . ______

(a) the conclusion is correct . . . or the conclusion is wrong because
(b) both sides of the equation are not equivalent
(c) we cannot extract the square root
(d) the only way to solve the equation is to open the brackets first
(e) there is some other error in the procedure

4\. If m men can do a job in d days, then $m + r$ men can do the same job in how many days? . . . . ______

(a) $d + r$, (b) $d - r$, (c) $\frac{d}{m + r}$, (d) $\frac{md}{m + r}$

(e) $d(m + r)$

5\. In a group of cows and chicken, the number of legs was 14 more than twice the number of heads. The number of cows in the group is . . . . ______

(a) 5, (b) 7, (c) 10, (d) 12, (e) 14

6\. If when x is added to both the numerator and the denominator of the rational number a/b $(a \neq 0, b \neq 0)$, the value of the rational number is changed to c/d, then $x =$ . . . . ______

(a) $\frac{bc - ad}{bd}$ (b) $\frac{bc - ad}{d - c}$ (c) $\frac{bc - ad}{c + d}$

(d) $\frac{ad - bc}{d - c}$ (e) $\frac{ad - bc}{c + d}$

PART TWO (You should spend approximately 10 minutes on each problem.)

Please attempt each problem and include, in the space provided, your attempt(s) to solve each problem. In case(s) where you are unable to complete the solution, it will be extremely useful to see just how far your were able to proceed.

1. The last digit in 3^{40} is

2. A set of n numbers has a sum of s. If each number n of the set is increased by 16 and then multiplied by 7, the sum of the numbers in the new set thus obtained is?

3. A certain number of students can be accommodated in a hostel. If 2 students share a room, then 2 students will be left without a room. If three students share a room, then two rooms will be left over. Determine the number of rooms.