DOCUMENT RESUME

ED 093 717 SE 018 094

AUTHOR	Jurdak, Murad Eid
TITLE	The Effects of Emphasizing Mathematical Structural Properties in Teaching and of Reflective Intelligence on Four Selected Criteria. Technical Report 275.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
REPORT NO	TR-275
PUB DATE	Sep 73
CONTRACT	NE-C-00-3-0065; OEC-5-10-154
NOTE	190p.
EDRS PRICE	MF-$0.75 HC-$9.00 PLUS POSTAGE
DESCRIPTORS	Doctoral Theses; Grade 8; *Instruction; Learning; *Learning Characteristics; Mathematics Education; *Research; *Secondary School Mathematics
IDENTIFIERS	Cognitive Style

ABSTRACT

The purposes of this study were: (1) to compare the effectiveness of two teaching methods having two distinct levels of emphasis on mathematical structure in organizing and presenting the same mathematical content, and (2) to identify the effect of the cognitive ability of reflective intelligence on four cognitive levels of learning a second-order mathematical structure whose learning depends on an already learned first-order system. Integral powers of 2 and 3 were chosen as the mathematical content. Treatment 2 (T2) emphasized explicitly the structural properties in developing operations and algorithms and in manipulating isomorphisms, whereas T1 attempted a direct approach with no explicit emphasis on structural properties. Each of five teachers taught two sections of intact eighth grade classes, using the T1 approach with one section and the T2 method with the other. An immediate posttest and a retention test two weeks later were given to students. Multivariate analysis of variance and discriminant analysis were used on the data. Results showed that T2 was relatively superior to T1 in producing better performance on solving mathematical sentences of the form ab = x, ax = b, and xa = b, and on solving the same type of mathematical sentences in an isomorphic model. Better performance was associated with higher level of reflective intelligence in learning a second-order mathematical system. (DT)

ED 093717

TECHNICAL REPORT

THE WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

The University of Wisconsin Madison, Wisconsin

U.S. Office of Education Center No. C-03 Contract OE 5-10-154

SE 018 094

Technical Report No. 275

THE EFFECTS OF EMPHASIZING MATHEMATICAL STRUCTURAL PROPERTIES IN TEACHING AND OF REFLECTIVE INTELLIGENCE ON FOUR SELECTED CRITERIA

Report from the Project on Conditions of School Learning and Instructional Strategies

by Murad Eid Jurdak

John G. Harvey and Thomas A. Romberg
Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

September 1973

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in the University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

STATEMENT OF FOCUS

Individually Guided Education (IGE) is a new comprehensive system of elementary education. The following components of the IGE system are in varying stages of development and implementation: a new organization for instruction and related administrative arrangements; a model of instructional programing for the individual student; and curriculum components in prereading, reading, mathematics, motivation, and environmental education. The development of other curriculum components, of a system for managing instruction by computer, and of instructional strategies is needed to complete the system. Continuing programmatic research is required to provide a sound knowledge base for the components under development and for improved second generation components. Finally, systematic implementation is essential so that the products will function properly in the IGE schools.

The Center plans and carries out the research, development, and implementation components of its IGE program in this sequence: (1) identify the needs and delimit the component problem area; (2) assess the possible constraints--financial resources and availability of staff; (3) formulate general plans and specific procedures for solving the problems; (4) secure and allocate human and material resources to carry out the plans; (5) provide for effective communication among personnel and efficient management of activities and resources; and (6) evaluate the effectiveness of each activity and its contribution to the total program and correct any difficulties through feedback mechanisms and appropriate management techniques.

A self-renewing system of elementary education is projected in each participating elementary school, i.e., one which is less dependent on external sources for direction and is more responsive to the needs of the children attending each particular school. In the IGE schools, Center-developed and other curriculum products compatible with the Center's instructional programing model will lead to higher morale and job satisfaction among educational personnel. Each developmental product makes its unique contribution to IGE as it is implemented in the schools. The various research components add to the knowledge of Center practitioners, developers, and theorists.

ACKNOWLEDGEMENTS

I wish to express my gratitude to those who have helped make this study possible. Special thanks are due to:

Dr. J. F. Weaver, who as my major professor, has offered guidance, assistance and encouragement throughout the course of this study; Dr. T. A. Romberg for his helpful suggestions and assistance throughout the course of this study; Dr. M. O. Pella for his many helpful suggestions and continuous encouragement; Mr. Ed Haertel for helping in data handling; Mrs. Ethel Koshalek for typing the dissertation.

Special appreciation is expressed to:

Ford Foundation for its financial support of the Arab Science Teachers Program at the University of Wisconsin in which I was a participant; Science and Mathematics Education Center (SMEC) of American University of Beirut for offering administrative help; Wisconsin Research and Development Center for Cognitive Learning for partial financial support of this study.

My gratitude also goes to the students, teachers and administrators of International College (Beirut) and the UNRWA/UNESCO schools in Lebanon for their cooperation.

My wife Muna deserves my greatest thanks for her encouragement and many sacrifices.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

Abstract

THE PROBLEM

This study had two general aims:

(1) To compare, on predetermined criteria, the effectiveness of two teaching methods having two distinct levels of emphasis on mathematical structure in organizing and presenting the same mathematical content.

(2) To identify the effect of a cognitive ability known as reflective intelligence on four cognitive levels of learning a second-order mathematical structure whose learning depends on an already learned first-order system.

Procedure

Integral powers of 2 and 3, as models of an infinite cyclic group, were chosen as a suitable mathematical content for 8th graders in Lebanon where the investigation was carried out. The learning of these two models depends upon the learning of integers with the operation of addition. Two treatments T_1 and T_2 were constructed in such a way that T_2 tended to emphasize explicitly the structural properties of the models in developing operations and algorithms and in manipulating isomorphisms whereas T_1 attempted a direct approach with no explicit emphasis on structural properties.

T_1 and T_2 were piloted and then administered to 5 intact 8th grade classes each (a total of 114 students for each). Each of the five teachers, following the specifications of T_1 and T_2, taught two sections, one according to each. T_1 lasted for six 40-minute lessons and T_2 for seven 40-minute lessons.

The sample was divided, according to the sum score of two parts of Skemp test of reflective intelligence, into three categories; (1) low level (L); (2) medium level (M); and (3) high level (H).

Outcomes were evaluated against four predetermined criteria:

C1: Ability to solve mathematical sentences of the form $ab = x$, $ax = b$ and $xa = b$ in the taught models.

C2: Ability to solve the same type of mathematical sentences in an isomorphic model.

C3 : Ability to select and solve mathematical sentences which "model" decisions in a physical model on which an isomorphic structure is imposed.

C4: Ability to select and solve mathematical sentences which "model" decisions in a generalized model of the taught model, i.e., contains an isomorphic copy of it.

Measurements of the four criteria were taken at two occasions:

(1) immediately following the conclusions of the treatments (achievement) and

(2) two weeks later (retention). X1, X2, {X3, X4}, {X5, X6} were achievement measures of C1, C2, C3 and C4 respectively. X7, X8, {X9, X10}, {X11, X12} were the corresponding retention

measures of the four criteria. As noticed, a pair of measures was associated with each of C3 and C4: the first of the pair was a measure of selecting the correct mathematical sentences which model a given decision and the second a measure of giving the correct solution set.

Five questions were generated from the general aims of the study:

1. Are there treatment differences? For which criteria? for which measures?
2. Are there reflective intelligence differences? For which criteria? For which measures?
3. Are there treatment differences on differences variables? (a difference variable was defined as the difference between the achievement and retention scores on the same scale measuring a criterion).
4. Are there reflective intelligence differences on difference variables? For which variables?
5. Within reflective intelligence, are there treatment differences? For which criteria? For which measures?

Multivariate analysis of variance and discriminant analysis were used to answer these questions.

Results

1. Across reflective intelligence, differences in estimated means between treatments favored (a) T_2 significantly ($\alpha = 0.01$) in the achievement phase on X1 and X2 (measures of C1 and C2) and T_1 on

X3 (one measure of C3); and (b) T_2 marginally ($0.01 < p < 0.05$) in the retention phase on X7 and X8 (measures of C1 & C2).

2. Across treatments and in both achievement and retention phases and for measures of C1, C2, C3 and C4, the difference in estimated means among reflective intelligence levels favored (significantly in most cases and marginally in the rest) the higher level.
3. There were no significant differences between treatments or reflective intelligence levels on difference variables.
4. Within reflective intelligences levels, significant differences between treatments were limited to the achievement phase and to criteria C1, C2 and C3.

Conclusions

1. T_2 is relatively superior to T_1 in producing better performance on measures of C1 and C2.
2. There is no evidence that the emphasis on structural properties in teaching (as in T_2) is conducive to better performance on measures of C3 and C4.
3. Better performance on each of C1, C2, C3 and C4 is associated with higher level of reflective intelligence in learning a second order mathematical system.

Chapter I

PROBLEM AND BACKGROUND

Introduction

The problem of this study belongs to curriculum research in mathematics education. Romberg (1970) identifies a curriculum model with six components: Content, learner, teaching agent, instruction-learning process, operational plan and intended learnings. This study deals with questions concerning the effect and interaction of two curriculum components; namely, learners and instruction (and hence operational plan), while controlling as much as possible the other components: Content, teacher and intended learnings. One aspect of the population of learners to be considered is a cognitive ability related to mathematics learning. As to instruction component, two teaching methods which differ in the organization and presentation of the same mathematical content will be considered.

Within the context of cognitive ability - instruction - interaction, the problem of this study is an instance of a more general problem. The general problem is to identify the different manners in which instructional methods, with different levels of emphasis on the structure of the discipline, (as far as organization and presentation of content is concerned) influence the learning outcomes of a population of learners with different levels of a cognitive ability related to the learning of mathematical structure. Moreover, the constraints of the problem are such that the evaluation of the outcomes will be in terms of predetermined criteria revealing the multivariate nature of mathematics achievement.

In the light of the above mentioned problem, the components of the specific problem of this study can be identified. The cognitive ability to be considered is called reflective intelligence as identified originally by Piaget (1950) and then extended and measured by Skemp (1961). The two teaching methods differ in the level of their emphasis on structure of mathematics in organizing and presenting the same mathematical content. In general, the difference in emphasis takes the form of using structural properties in developing operations and algorithms and in manipulating isomorphisms among different models of the same mathematical theory. The first teaching method (labeled T_1 henceforth) attempts no explicit emphasis on the structure of the models while the second teaching method (labeled T_2 henceforth) deliberately builds in the organization and presentation of the content an explicit emphasis on the structure of the models. The mathematical content consists of integral powers of 2 and 3 with the operation of multiplication (and division). The reasons for selecting this piece of mathematical content will be explained later. The teacher will be controlled by requiring him to follow a specially prepared instructional material which consists of a set of tasks in a certain sequence according to the specification of each teaching method. Four predetermined criteria will be used to evaluate the learning outcomes:

1) Ability to solve equations in the taught models.
2) Ability to solve equations in isomorphic models of the same theory.
3) Ability to select and solve open sentences which "model" decisions in an isomorphic model.
4) Ability to select and solve open sentences which "model" decisions in a generalized model (a generalized model of a model M is a model of the same theory which contains M in the sense of direct summand).

The justification for selecting these criteria and the way they reveal the multivariate nature of mathematical abilities will be discussed in due course. In the pages which follow, a discussion of the rationale, mathematical and psychological backgrounds of the problem will be discussed.

Rationale for the Study

One distinguishing characteristic of what came to be called "modern" mathematics is its emphasis on the structure of the discipline. Scott (1966), after examining contemporary trends in elementary school mathematics, states:

> "In the way of summary, the first and most noticeable feature of the modern (contra traditional) programs is their attention to the structure of mathematics. The ultimate objective of most of the programs appears to be the development within children of an awareness of

mathematics as an entity. The entity is held intact by pervasive ideas or patterns which occur and recur regardless of whether one is studying arithmetic, algebra or geometry." (pp. 23-24)

The reasons given for emphasizing the structure of the discipline are many. Some of the reasons are formulated in terms of relationships of expected social needs and the nature and development of mathematics itself, while some other are formulated in psychological terms. In the first category two reasons are recurrent. The need of modern society for a deeper and more extensive mathematical education is often mentioned as one reason. Moreover, with the expanding growth of mathematical knowledge and its applications, a certain degree of uncertainty is associated with the nature of the needed mathematical skills of tomorrow. Hence, to cope with the situation, curriculum developers in mathematics envision the modern mathematics curriculum as a super structure built on unifying ideas or patterns. The description provided by School Mathematics Study Group (1958) might illustrate the point:

"The world of today demands more mathematical knowledge on the part of more people than the world of yesterday, and the world of tomorrow will make still greater demands. Our society leans more and more heavily on science and technology. The numbers of our citizens skilled in mathematics must be greatly increased; an understanding of the role of mathematics in our society is a prerequisite for intelligent citizenship. Since no one can predict with certainty his future profession, much less foretell which mathematical skills will be required in the future by a given profession, it is important that mathematics be so taught that students will be able in later life to learn the new mathematical skills which the future will surely demand of man of them.

> To achieve this objective in the teaching of school mathematics three things are required. First, we need an improved curriculum which will offer students not only the basic mathematical skills, but also a deeper understanding of the basic concepts and structure of mathematics. Second, mathematics program must attract and train more of those students, who are capable of studying mathematics with profit. Finally, all help possible must be provided for teachers who are preparing themselves to teach these challenging and interesting courses." (n.p.)

Reasons formulated in psychological terms are often given. Principal among these are those based on Piaget's developmental theory which assumes a correspondence between mathematical structures and cognitive structures. The interpretation of Piaget theory in education reinforced the trend of emphasizing structure of the discipline particularly in mathematics. Perhaps Bruner's book Process of Education (1963) is the most illustrative in this respect. Bruner gives four reasons for teaching the fundamental structure of a subject: The first is that understanding fundamentals makes a subject more comprehensible. Second, unless detail is placed into a structured pattern, it is rapidly forgotten. Third, an understanding of structure is the apparent means to achieve adequate transfer of training. At last, this emphasis on structure leads to narrowing the gap between advanced knowledge and elementary knowledge.

Modernizing mathematics curriculum, along the above mentioned lines, is not an exclusive phenomenon in the developed countries. Developing countries are following a similar pattern either through national organizations and more often through the cooperation of international bodies such as Unesco and UNICEF. A somewhat relevant example might be mentioned: Unesco Mathematics Project for the Arab Countries. This project is a regional curriculum improvement project with the cooperation of Unesco. The pattern followed in starting and implementing this project is similar to that of SMSG. The project engaged a large segment of the mathematical and educational community in drawing a modernized secondary mathematics curriculum and sample textbooks. In the suggested syllabus, free use is made of algebraic structures as applied to the characterization and construction of number systems as well as to geometry. Such structures as group, ring, field, vector space and their applications are central in the curriculum (Unesco, 1969). Although this project started as a secondary school project, it is already exerting pressure on junior high school.

Prominent among the unifying ideas which are often used are those which belong to the structural properties associated with operations and algorithms in subsystems of the real number system. It is often found that the structural properties of an ordered semigroup (whole numbers with addition) and those of an ordered group (integers with addition, non-zero integers with multiplication) are central to the development of other structures. Based on these basic structures, structural properties of more involved structures (such as ring, field, vector space) are used to develop other subsystems of real numbers.

Three of the important contexts in which these structural properties are used are: Developing operations, developing algorithms and manipulating isomorphisms, analogies and patterns. It is often seen that the idea of binary operation on a set is central. Properties of this operation are then investigated. The idea of the identity element is used; for example, to organize and structure many examples with this property. Inverses (or opposites) are introduced using the ideas of closure, identity and inverse laws. One prominent aspect in which structural properties of a group are used is that of an inverse operation (Scott, 1966). Subtraction and addition, multiplication and division are conceived as inverse operations. School Mathematics Study Group textbooks (1965a, 1965b) follow essentially the above mentioned pattern. Of course, programs differ in this respect and even the same program differs from one grade to another.

Not only the structural properties are used in developing operations, but also in developing algorithms. What came to be called "traditional mathematics" was often criticized for its failure to put in proper perspective the relation between operation and algorithm. The criticism focuses on the argument that traditional mathematics views algorithms as isolated rules to be practiced up to a mastery in both speed and accuracy. Contemporary programs claim that algorithms should be viewed as shortcuts which derive their meaning from the total system to which they belong. To achieve this goal the sequence of steps (using structural properties at each step) leading to the algorithm is built, emphasized and even practiced at the initial stages of the development. The extensive use of expanded notation and its manipulation using structural properties in developing algorithms is just one illustration.

A third use of structural properties is apparent in the emphasis put on analogies and patterns. The structural properties of whole numbers; for example, are seen to hold also in integers and those of integers to hold in rationale, etc. Inherent in this development is the idea of isomorphism which embeds structures in other structures. Analogies are built between different models and patterns are suggested. Again programs differ in the level on which they explicitly emphasize ismorphisms. They range from an explicit construction of the isomorphism to mere suggestion of analogy and pattern.

The effect of emphasis on structure of mathematics in teaching has been investigated with various motivations and procedures. Some of these studies were short-term comparative status studies, such as Rosenbloom (1960), Ruddel (1961), Cassell-Jerman (1963), Weaver (1963) and Mastain (1964). Some of these were long-term longitudinal multivariate studies, such as National Longitudinal Study of Mathematical Ability (NLSMA). A third category consists of experimental studies in school setting involving one structural property (distributive property, inverse property) as Gray (1963), Coxford (1965) and Osborne (1966). A fourth category consists of experimental clinical studies using a selected number of finite structures and using logic-mathematical criteria. The work of Dienes and Jeaves (1971) fits in this category. Most of these studies will be reviewed in Chapter II.

This research differs from each of the above mentioned studies in at least one of the following: Type, content used, or criteria selected for evaluation. This study is an experimental study in school setting involving controlled variation in the organization and presentation of a well-defined segment of mathematical content-the latter being two models (in toto) of group theory belonging to junior high school. Moreover, the effect and interaction of reflective intelligence with learning mathematical structure has not been studied in experimental setting. The criteria to be used in this study, beside having a logico-mathematical basis, can be interpreted in the multivariate mathematics achievement model which was developed by NLSMA (Romberg and Wilson, 1969).

Significance of the Study

The significance of this research to educational practice in mathematics education might be related to three areas. First, research concerning the effect of emphasizing structure of mathematics in teaching is not consistent and the results are not conclusive, particularly those which involve criteria of higher cognitive levels. Hence, a need for research in this area is still relevant, particularly when and where curriculum development in mathematics is to be undertaken. Second, controlled experimentation which involves subtle variation in the organization and presentation of mathematical content helps to generate necessary information for both practical and theoretical purposes. Moreover, the variation in the presentation and organization of content in this study is planned to approximate, in a specific situation, the more general pattern of the main stream of modern mathematics. It attempts to examine, in a controlled experimental way, one basic assumption of what came to be called modern mathematics. Third, the multivariate criteria to be used will hopefully suggest a pattern of the payoffs and losses of emphasizing structure in mathematics teaching. A discussion of the significance of reflective intelligence for mathematics learning will be included in the discussion of the psychological basis of the problem.

Background

Mathematical Background

The mathematical basis of this study can be discussed in terms of model theory. The meta mathematical treatment of models implies identifying a language and a theory. The language in group theory, for example, consists of:

1. variable symbols $(x,y,z,\ldots.)$

2. constant symbol (e)

3. function symbol (o)

4. logical symbols: $\forall$, V, Λ, $\exists$

Terms are generated as follows:

1. An individual variable symbol or constant symbol is a term.

2. If t_1 and t_2 are terms then $t_1 o t_2$ is a term.

The following non-logical axioms (sentences and formulae using symbols of language only) is the abstract group theory:

G_1 : $\forall(\forall x)(\forall y)(\exists z)(xoy = z)$.

G_2 : $(\forall x\ \forall y\ \forall z)\ [(xoy)oz = xo(yoz)]$

G_3 : $(\forall x)[xoe = x]$

G_4 : $(\forall x)(\exists y)[xoy = e]$

A model of group theory is an interpretation of the above system in a certain set. An interpretation of group theory in a set S is a function from the language to S which maps a variable symbol into a variable over S, a constant symbol into a fixed element of S and the operation symbol into an operation in S. So, the set of integral powers of 2 (or 3) with the operation of multiplication is a model of group theory.

A model of a theory is called a branch of applied mathematics while the language with the theory is called a branch of pure mathematics. As Eves and Newson (1968, pp. 169) suggest: "The difference between applied and pure mathematics is not one of applicability and inapplicability, but rather of concreteness and abstractness."

In the above sense when we are dealing with models we are dealing with applied mathematics. Mathematically, once an operation and an algorithm (which is a finite number of applications of the operation in the model) determined in one model they are determined in all isomorphic models of the same theory. One expects that this "economy" feature of mathematics to hold if learning of mathematics is considered. This logico-mathematical property motivated the inclusion of the second criterion as one criterion for evaluation.

The third and fourth criteria have logico-mathematical components, although they cannot be justified exclusively on logico-mathematical basis. The hypothesis is that selecting and solving mathematical sentences which "model" decision in an isomorphic (to the taught model) model or a generalized model involve also to some extent the "economy" feature of mathematics as discussed above. However, selecting a mathematical sentence involves a problem solving ability which might be related to a cognitive ability such as reflective intelligence.

Psychological Background

Psychologically, the problem of this study is related to cognitive abilities which influence the learning of mathematics. Historically the procedure was to identify "factors" which account for variance of scores. Review of studies of mathematical abilities is given in Wrigley (1958). Peel (1971) comments on findings and theories of mental factors:

> "The findings and theories offered are not at all conclusive on the nature of mathematical ability. We do not on the whole obtain a clear-cut picture, save that general intellectual ability is important (a trite observation to make to any experienced teacher) and there is a group factor we might call mathematical ability. Attempts to analyze this in greater detail reveal number, verbal, spatial, thinking elements intricately mixed up with material content" (pp. 153)

A rather different approach to the study of cognitive abilities related to mathematics learning is provided by the developmental psychology of Piaget. Piaget (1950) conceives a close relation between formal logic (the axiomatic method) and psychology of intelligence and this relation can be described as that of "deductive geometry and positive or physical geometry (pp. 28). Piaget defines intelligence as constituting "the state of equilibrium towards which tend all the successive adaptions of a sensori-motor and cognitive nature, as well as all assimilatory and accommodatory interactions between the organism and the environment." (pp. 11)

For this definition to make sense according to Piaget theory of intelligence some clarification of terms are necessary. Piaget, following an evolutionary biological trend, assumes a continuous interaction between subject and object. "Assimilation may be used to describe the action of the organism on surrounding objects, in so far as this action depends on previous behavior involving the same or similar object" (pp. 7). If this definition is restricted to psychology, then the modifications which the organism imposes are determined by "movement, perception or interplay of real or potential actions (conceptual operations)". Accommodation refers to the action of the environment on the organism and psychologically it refers to the process in the sense that the pressure of circumstances always leads, not to a passive submission to them, but to a simple modification of the action affecting them. Adaptation is "an equilibrium between assimilation and accommodation." In this sense intelligence as defined is an extension and a perfection of all adaptive processes.

Piaget identifies two distinct types of intelligence: Sensori-motor and reflective intelligence. Sensori-motor intelligence is the permanent state of equilibrium towards which all successive adaptations of sensori-motor nature tend (including perceptual activity, the formation of habits and pre-verbal or pre-representative intelligence itself). On the development scale, this kind of intelligence is completed before appearance of language ($1\frac{1}{2}$ - 2 years).

Reflective intelligence is the permanent state of equilibrium towards which all successive adaptations of cognitive nature tend (including representative thought, intuitive thought, concrete operational thought). This reflective thought, which is characteristic of the adolescent, exists from the age of 11-12 years when the person is capable of reasoning in a hypothetic-deductive manner. Concrete operation thought consists of first-degree grouping (group of actions in the mathematical sense), whereas reflective thinking "consists of reflecting (in the true sense of the word) on these operations and, therefore, operating on operations or on their results and consequently effecting a second degree grouping of operation."

The transition from sensori-motor intelligence to the reflective intelligence level requires three essential conditions. These conditions are identified by Piaget (1950) as:

> "Firstly, an increase in speed allowing the knowledge of the successive phases of an action to be moulded into one simultaneous whole: Next, an awareness not simply of the desired results of an action, but its actual mechanism, thus enabling the search for the solution to be continued with a consciousness of its nature. Finally, an increase in distances, enabling actions affecting real entities to be extended by symbolic actions affecting symbolic representations and thus going beyond the limits of near space and time." (pp. 121)

The findings of the Geneva School concerning the growth of pupils' thinking had an impact on the teaching of mathematics.

Z. Dienes (1969) developed an elementary mathematics program whose psychological basis is built on Piaget ideas. Again, Dienes distinguishes between constructive thinking (which the Geneva School calls pre-operational) and analytic thinking leading to operational thinking.

R. R. Skemp started an attempt to begin in the field of mathematics what Piaget did in the field of number. Skemp (1961) assumes that the transition from arithmetic to mathematics requires the exercise of reflective intelligence. Arithmetic requires perception of numbers and their relationships and making correct responses without being aware of the relationships or methods. Mathematics on the other hand, requires awareness of concepts, operations and their interrelations. Skemp (1961) modifies Piaget definition somewhat as follows:

> "Reflective intelligence is the functioning of a second order system which:
>
> 1. can perceive and act on the concepts and operations of sensori-motor system,
>
> 2. can act on them in ways which take account of these relationships and of other information from memory and from the external environment,
>
> 3. can perceive relationships between these concepts and operations." (p. 49)

Based on this definition, he prepared a four-part test: the first is on concepts formation, the second on reflective activity with concepts, the third on operation formation and the fourth on reflective activity with operations. Using this instrument he studied the relationship of reflective intelligence to mathematics achievement. His study will be reviewed in Chapter II.

From the above presentation there seems to be a relationship between reflective intelligence and learning of mathematics. The correlation between reflective intelligence and mathematics learning was studied by Skemp (1961) and Harrison (1967). This cognitive ability becomes more significant to the learning and teaching of mathematics if its effect is confirmed and known and if one could identify procedures to develop it (still an open possibility). For these reasons reflective intelligence was included as one factor in this study.

Beside the psychological basis of the Skemp test, some practical considerations motivated its selection for the present study. Because standardized intelligence tests for Lebanon are not available yet and because the Skemp test is non-verbal and almost culture-free, the latter was convenient and practical to use in the present study.

Purpose of the Study

The specific question of this study can now be formulated. Two factors are considered in this study: First, teaching method with two levels of emphasis on structure T_1 and T_2 and second reflective intelligence. Generally T_1 and T_2 differ (a characterization will be given in Chapter III) in that T_2 attempts an explicit emphasis on the structural properties of the structure of models (in this case infinite group structure) in developing operations, algorithms and in manipulating ismorphisms. T_1 does not attempt such an explicit emphasis in these areas.

The two general aims of this study are:

(1) To compare, on predetermined criteria, the effectiveness of two teaching methods having two distinct levels of emphasis on mathematical structure in organizing and presenting the same mathematical content.

(2) To identify the effect of a cognitive ability known as reflective intelligence on four cognitive levels of learning a second-order mathematical structure whose learning depends on an already learned first-order system.

The questions of the study become:

(1) Is there differential effect of T_2 or T_1? If yes for what criteria and for what measures of these criteria?

(2) Is there reflective intelligence differences? If yes for what criteria and for what measures?

(3) Is there treatment differences on difference variables? For which criteria? (A difference variable is defined as the difference score between the achievement and retention scores on a scale measuring a given criterion).

(4) Is there reflective intelligence differences on difference variables? For which criteria?

(5) Within reflective intelligence levels, are there treatment differences? For which criteria? For which measures?

Chapter II

RELATED RESEARCH

For the sake of convenience, studies directly related to mathematical structure in teaching will be divided into three categories: (1) Studies which involve one structural mathematical property in school settings; (2) experimental clinical studies and (3) National Longitudinal Study of Mathematical Abilities (NLSMA). Studies related to reflective intelligence will be reviewed separately.

In this chapter, it is not the intention to be exhaustive in reviewing related studies. Many of these relate only peripherally to the problems investigated in this study, as is the case in evaluation studies (such as Rosenbloom, 1961; Weaver, 1963; Cassel-Jerman, 1963) which attempt to compare contemporary mathematics programs with traditional programs. These studies, for one thing, evaluate indirectly, and some tangentially, the effect of emphasis on the structure of mathematics in teaching--the latter being one noticeable feature of contemporary school mathematics. Due to the great number of such studies and the difficulty in drawing a demarcation line between what is related and what is not, it was decided not to review such studies in this chapter. (A review of such studies is given in Begle and Wilson, 1970).

Studies Related to the Teaching/Learning of Mathematical Structure

Studies Which Involve One Structural Mathematical Property

These studies relate to the present study in that they involve controlled variation in mathematical content using one mathematical structural property or in that they identify factors which influence the learning of such a property. Among studies which belong to this category are those of Gray (1963), Coxford (1965), Osborne (1966) and Weaver (1973).

Gray (1963) investigated the effectiveness of a program of instruction at third-grade level in introducing multiplication which was based on the development of an understanding of the distributive property of multiplication over addition. Two sets of eighteen lessons each were developed. One set used the distributive property in introducing multiplication and the other did not. Posttests and retention tests of multiplication and transfer ability were given. In addition, an interview test was given to ascertain whether a subject was able to give a rational explanation of the multiplication procedures he used. The results indicated that the treatment which used the distributive property produced significantly higher scores on posttest of transfer ability and also on retention tests of multiplication and transfer. Also, significant differences, favoring the same treatment, were found in the number of subjects who used the distributive property in finding answers for all items. Gray concluded (among other things) that an understanding of distributive property does not appear to develop unless specifically taught and that, in as much as the distributive property is a structural property, teaching for an understanding of structure can produce superior results in terms of pupils' growth.

Coxford (1965) conducted a study, one of whose purposes was to measure the effect of an instructional program (labeled PPW) for teaching subtraction using extensively the relationship between addition and subtraction as compared with the traditional take-away approach (labeled TA). The structural property under consideration in PPW was the set-analogue of the inverse property. IQ and achievement were used as covariates. Criterion measures included subtraction transfer and application, subtraction computation, addition computation, problem solving and total arithmetic achievement. The experiment lasted for one year and involved six first-grade classes. The results indicate that for higher ability group there were reliable (significant?) differences favoring the TA approach in subtraction computation. Also, although the differences were not reliable, there was a definite tendency for higher ability in PPW approach to have higher scores for subtraction applications. Coxford concluded:

1. The TA approach led to a greater immediate proficiency in solving subtraction sentences than did the PPW approach in the higher ability group.

2. The PPW approach tended to facilitate solutions of applications of subtraction to a greater extent than did the TA approach in the higher ability group.

Osborne (1966) conducted a follow-up study (in the second grade) of the pupils in Coxford study. Tests to measure knowledge (recall), application, transfer and "structure-deduction" were used. Again Osborne concludes that significantly greater mastery of knowledge, application and "structure-deduction" was exhibited at the end of grade two by the children in the PPW approach than by those in the TA approach.

Weaver (1973) reported a status study designed to ascertain whether there existed differential achievement effects of distributive property associated with such factors as context (regrouping sets, multiplication-addition), distributve form (left, right), item-stem format (horizontal, vertical). . . . Twelve 9-item tests were constructed to incorporate all identified factors. The tests were given to intact 4th, 5th, 6th, and 7th grade classes from two midwestern city school districts. A coding scheme was used to categorize students' responses. Weaver gives some tentative conclusions among which the following seem relevant:

1. . . ., Pupils' sensitivity to the use of the distributive property was relatively low.
2. Across grade levels it appears that "regrouping sets" examples were less difficult than "multiplication-addition"

examples.

3. Based upon the response categorization, pupils had only a limited tendency to respond in the same way across the set of examples within a test.

Weaver's conclusions seem to suggest that simply teaching the distributive property, inasmuch as it is a structural mathematical property, does not seem sufficient to promote sensitivity towards its use, nor does it seem to produce a consistent regularity in responses towards relevant items if such factors as context and symbolic representations are varied. Weaver advances the following conjecture:

> "Work with properties which give 'structure' to some particular aspect of mathematics is no guarantee that pupils will be exempt from rote learning and 'symbol pushing'." (p. 5)

Clinical Experimental Studies

These studies relate to the present study inasmuch as they involve the effect of learning one or more mathematical structures on transfer to analogous structures. Among studies which belong to this category are those of Dienes and Jeeves (1970), Branca and Kilpatrick (1972) and Scandura (1967).

Dienes and Jeeves (1965, 1970) started a series of investigations in a theory-building framework. The results of their investigations are reported in two books, Thinking in Structures (1965) and the Effect of Structural Relations on Transfer (1970). The first book, which is related only tangentially to the present study, deals with the problem of identifying strategies and evaluations (a verbal retrospective account of the moves the subject used) and the relations among them. Since the second book is more relevant, it will be reviewed in some detail.

The basic question to which Dienes and Jeeves (1970) addressed themselves was concerned with the manners in which relationships between structures affect the successive learning of other structures. Three general relationships between structures were identified:

1. Recursion which is of two kinds:
   - generalization which is the process of passing from a structure to a wider structure which has the same generating rules, as in passing from cyclic three group to cyclic five group.
   - particularization which is the reverse process of generalization.
2. Embeddedness: A structure A is embedded in B if A has one isomorphic image (simple) or several isomorphic images (multiple).
3. Overlapping which means that there is one (simple) or several (multiple) structures common to two structures.

The following diagram (Dienes and Jeeves, 1970), p. 63) gives the various relationships that were used in the order in which the tasks were presented.

A special electrical machine, which was capable of generating the structures mentioned, was devised. Two kinds of dependent variables were used: (1) measures of the degree of success or failure of a subject (3 measures); and (2) measures of the way in which the subjects tackled the tasks (5 measures). Two groups of subjects participated in the experiment: university adult students and children of average chronological age of 11 years. Based on their results, Dienes and Jeeves concluded (among other things) that:

(i) children consistently find it more difficult to generalize than adults,

(ii) children consistently find it easier to particularize than to generalize,

(iii) children find the difficult-easy (5)3 treatment easier to handle than the (3)5 treatment.

(iv) on the (3)5 treatment adults do better than children, but on the (5)3 treatment the performance of the adults and children approximate. On an S-R-O model there should be no difference between the 5-task under the two conditions, 5 preceded by 3, or 5 given first - but there is,

Task I	Task II		Task III		Task IV
Klein 4	3*	→ Recursion (generalisation)	5	→ Recursion (generalisation)	7
Klein 4	5	→ Recursion (particularisation)	3	→ Recursion (generalisation)	7
Klein 4	3	→ Recursion (generalisation) plus Embeddedness (simple)	6	→ Recursion (generalisation) plus Overlap (simple)	9
Klein 4	3	→ Recursion (generalisation) plus Embeddedness (simple)	6	→ Overlap (multiple)	9A**
Klein 4	6	→ Recursion (particularisation) plus Embeddedness (simple in reverse)	3	→ Recursion (generalisation) by a factor) plus Embeddedness (simple)	9
Klein 4	6	→ Recursion (particularisation) plus Embeddedness (simple in reverse)	3	→ Embeddedness (multiple)	9A

* numbers refer to the order of the group
** 9A is $3 \oplus 3$ (direct sum of two 3-cyclic groups).

(v) on the (3)6 order, the adults again perform better than the children, and once again the performances of adults and children approximate on the (6)3 order,

(vi) adults and children both find embeddedness much harder than generalisation,

(vii) adults and children both find overlap more difficult than generalisation,

(viii) where multiple embeddedness, e.g., (3)9A, is replaced by simple embeddedness and recursion; e.g., (3)9, the narrow margin between adults and children becomes much bigger,

(ix) recursion by a factor is much harder than multiple embeddedness for children; e.g., (3)9 is much harder than (3)9A,

(x) adults find recursion by a factor, (3)9, easier than multiple embeddedness, (3)9A.

Branca and Kilpatrick (1972) raised the question as to how consistent subjects' strategies and evaluations were across different embodiments of the same group structure and also embodiments of another structure which was (in this case) a network structure. One hundred subjects from a private residential school participated in the experiment. The Klein group structure was embodied in a four-color game and a switch-light apparatus. The network structure was embodied in a map game. Measures of the strategies used by each subject and his evaluations were taken. Among the conclusions the authors give is that subjects showed substantial consistency across tasks in the evaluations they gave. However, consistency in the strategies used across tasks involving the two embodiments of group structure was not evident. In addition, there was no indication that subjects who used a particular strategy on the group structure tended to use a similar strategy on the network structure task.

Scandura started an attempt to build a theory of learning mathematics based on the assumption that mathematical behavior is rule-governed. A rule is conceived as the basic behavior unit and is defined as a function: $[(S_i, R_i)\ i = 1,\ldots,n,\ldots]$ in which each stimulus S_i is paired uniquely with a single response R_i. The scope of the rule is the domain of the function. In one study, Scandura (1967) reported an experiment which dealt with rule generality and consistency in mathematics learning. The study aimed to ascertain whether:

1. Successful responding is only within the scope of the learnt rule and no systematic within-scope differences exist.
2. Within-scope use of a rule imply beyond scope use when no information is given as to when a rule is or is not appropriate.

A number game which can be characterized by an ordered pair of positive integers (n,m) was used. Three treatments were constructed: (1) (S) taught a rule which can be applied only to two fixed values of n and m; (2) (SG) taught a rule which can be applied for a specific n and all values of m; (3) (G) taught a rule which can be applied for all n and m. The sample consisted of 85 undergraduate university subjects. The criterion test consisted of 3 problems, one in the domain of each rule. Based on the results the conclusions were:

1. Performance on within-scope problems did not differ appreciably, ..., and successful problem solving was limited almost exclusively to within-scope problems.
2. The rules taught tended to be used consistently on all problems whether they were appropriate or not.

National Longitudinal Study of Mathematical Abilities (NLSMA)

NLSMA was a long-term (5 years) large scale (112,000 students) longitudinal study which attempted among a host of other things to obtain some quantitative information on the cumulative and comparative effectiveness of mathematics curricula as embodied in textbooks. Textbooks were classified as "modern" if they explicitly used structural properties of the real number system and "conventional" if they were relatively or completely untouched by the ideas associated with recent curriculum reform. Scales to measure mathematics achievement were constructed according to a mathematics achievement model developed by NLSMA (Romberg and Wilson, 1969). Three populations of students were identified: X-population, Y-population and Z-population. X-population and Y-population were tested at grades 4 and 7 respectively in 1962 and for five consecutive years. Z-population was tested at grade 10 in 1962 and for 3 consecutive years.

NLSMA's data which relate to the present study include comparative results on scales which measured explicitly structural properties for X-population and Y-population. Four scales measured structural properties explicitly: (1) Whole number structure 1 which was given to grade 4 of X-population; (2) Whole Number Structure 2 which was given to grades 6 and 8 of X-population; (3) Structure of Rationals which was given to grades 7 and 8 of X-population and grade 7 of Y-population and (4) Structure which was given to grade 7 of X-population. These scales were classified as comprehension scales.

Although no comprehensive conclusion can be given as to the comparative effectiveness of "modern" textbook groups and "conventional" textbook group, data in NLSMA reports by Carry and Weaver (1969), Carry (1970) and McLeod and Kilpatrick (1969) give evidence to a restricted conclusion; for each of the four scales the mean of at least one "modern" textbook group was significantly greater than the means of all "conventional" textbook groups.

A Framework for the Present Study

It seems that the results of the reviewed studies are not consistent. This might be due partly to the grade level at which the studies were conducted and/or the nature of the study and/or nature of criteria used for evaluating outcomes:

1. The grade level varied from one study to another.
2. The studies reviewed here are clinical studies involving one or more structures (in toto) not necessarily belonging to ordinary school curricula or studies conducted in a school setting involving one structural property or multivariate longitudinal studies. In the first category any educational implication is an extrapolation. Although studies in the second category pertain to school curriculum, they do not apply to situations, particularly at preparatory and secondary levels, where most or all of the structural properties of one given mathematical structure are to be used. The results of the studies in the third category which are status studies, do not lend themselves to interpretation as causal consequences of the manipulation of independent variables.

3. The criteria used for evaluation vary; however, it is possible to divide them into three categories: Gray, Coxford, Osborne and Weaver used achievement criteria which were not necessarily meant to fit in any theoretical achievement model; Dienes and Jeeves, and Branca and Kilpatrick and Scandura used criteria which have necessarily some logic-mathematical basis. NLSMA used a theoretical model (Romberg and Wilson, 1969).

The present study proposes to investigate the effect of controlled variation in content resulting from using structural properties of one structure in a school setting. Moreover, the criteria for evaluation, beside having a logico-mathematical basis, are expected to fit in a theoretical model - the mathematics achievement model developed by NLSMA (Romberg and Wilson, 1969).

Studies Related to Reflective Intelligence

Skemp (1961) constructed an instrument to measure the reflective intelligence as defined by Piaget and extended by Skemp himself. (A fuller discussion was included in Chapter I). The instrument consists of four parts: Concept formation, reflective activity on concepts, operation formation, and reflective activity on operations (a redrawn version of the last two parts is given in Appendix E). The first part, which is a preliminary one, consists of items each of which has three examples of a certain concept, three non-examples of the same concepts and three instances to be judged by the subject whether they belong to the concept or not. A similar pattern holds for operation items except that the subject has to apply the operation (after discovering it) on three given instances. The reflective activity with concepts is the logical multiplication of two attributes of known concepts. The reflective activity with operation consists of the operation of combining two known operations, reversing a known operation or combining and reversing two known operations. Using this instrument, Skemp investigated the hypothesis whether the presence of reflective intelligence is a necessary, but not sufficient, condition for mathematics achievement. Skemp correlated reflective intelligence measures with mathematics achievement of IV form and V form students of a grammar school (10th and 11th grades) in England. The correlations for V form between a mathematical criterion and each of reflective activity on concept, use of operations and reflective activity on operations were 0.58, 0.42 and 0.72. The corresponding correlations for IV form were 0.56, 0.48 and 0.73. Skemp concluded that the data support the hypothesis. The two highest correlations being those between the two tests of reflective activities and mathematics, although an unexpected relatively high correlation between the use of operation and mathematics was observed.

Harrison (1967) investigated the power of the Skemp test in predicting mathematics achievement. He tried to answer the following questions:

1. Does the addition of measures of reflective intelligence to measures of general intelligence significantly improve the prediction of one's performance in mathematics?
2. How is the relationship between reflective intelligence and mathematics performance affected by student anxiety toward testing situations?
3. Are there significant differences in levels of reflective intelligence in age categories from ten to sixteen.

The study was carried out in two phases. In the first phase, six eighth-grade classes (125 students) were administered in May and June of 1966 a battery of tests. This battery included general intelligence tests, Sarason's Test Anxiety Scale for Children and the four parts of Skemp test. In the second phase, the sample consisted of two classes of students at each of the grade five, six, seven, eight, nine, ten and eleven levels (340 students). These students wrote the four Parts of Skemp test only and their scores were categorized according to age (ten to sixteen) and sex (boy or girl).

Based on his results Harrison concludes:

1. Measures of <u>Operation Formation</u> and <u>Reflective Activity with Operations</u> of Skemp test significantly improve the

prediction of mathematics performance (in grade 8 at least).

2. Students' anxiety toward testing situation (as measured by Sarason's <u>Test Anxiety Scale for Children</u>) does not significantly affect the relationship between reflective intelligence and mathematics performance.

3. Fifteen to sixteen-year-olds were found to operate significantly higher than ten to twelve-year-olds on Skemp test.

In the present study, it is the intention to investigate the relationship of reflective intelligence not to mathematics performance as a whole, but to four cognitive levels of mathematics achievement. Due to the static nature of two subtests of Skemp test: <u>Concept Formation</u> and <u>Reflective Activity with Concepts</u> and due to their non-significant contribution to the predictive power of Skemp test (as reported by Harrison), it was decided to use only the remaining two parts: <u>Operation Formation</u> and <u>Reflective Activity with Operations</u>.

Chapter III

PROCEDURE

Introduction

Two factors were considred in the present study: aptitude and instruction. The aptitude was a cognitive ability known as reflective intelligence. Instruction focused on two teaching methods T_1 and T_2 which differed in the presentation and organization of the same mathematical content. The general plan of the present study is shown in Figure 1. The major steps were:

1. A population of learners was identified and a sample was drawn from this population.
2. A mathematical content was identified and then was organized and presented from the point of view of each of T_1 and T_2.
3. Three levels of reflective intelligence were determined using the sum score of a two-part test developed by R. R. Skemp. The sample was divided equally into three categories: Low(L), Medium(M) and High(H).
4. Classes were assigned to one of the two treatments T_1 and T_2.
5. Outcomes of the two treatments were measured against predetermined criteria. Measures X1 to X6 were taken immediately after the conclusion of the treatments. Measures X7 to X12 were taken two weeks later.

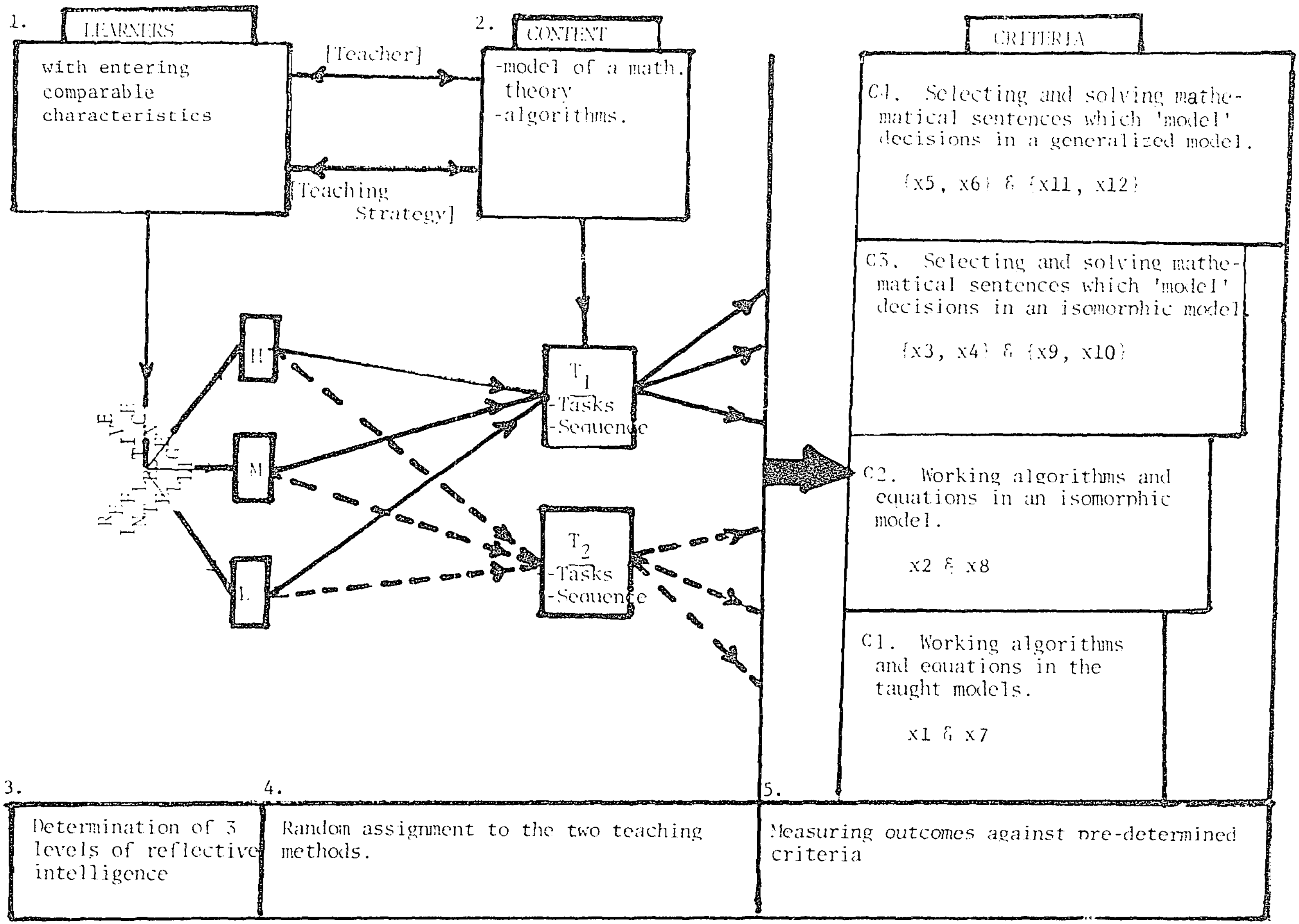


Figure 1: Research plan

Population and Sample

Population

The sample was drawn from two school systems: UNRWA/UNESCO education system and International College (I.C) school system. The United National Relief and Work Agency (UNRWA) is the United Nations body which coordinates and administers different services (including education services) to the Palestinian refugees in the host Arab countries and in the occupied territories. Education in UNRWA is technically supervised by UNESCO, and the school system is referred to as UNRWA/UNESCO school system. The Palestinian refugees live in "camps" provided by UNRWA. The "camp" is usually a collection of one or two-room units which have minimal physical facilities. It is not very unusual to assign a family of 6 to 8 to a two-room units. UNRWA provides for a subsistence allowance and a humble medical care system; however, the majority of the refugees have jobs in or outside the camp. This sub-population seems comparable to the lower socio-economic class in Lebanon.

Education in UNRWA is provided for and administered by UNRWA, but it is technically supervised by UNESCO. The teachers are themselves Palestinians. In terms of qualifications, teachers in UNRWA are comparable to teachers in Lebanese public schools. The majority of them are graduates of either preservice training centers (two years training after secondary education) or the inservice UNRWA/ UNESCO Institute of Education (two years training for elementary teachers plus two more years for subject teachers in junior high school). The schools are usually crowded buildings inside or near the camp. However, the students seem to be fairly motivated to learn probably because education is conceived by many as a means for economic and social improvement and as means to achieve national aspirations.

English is the language of instruction in mathematics and sciences on the junior high level. However, English is taught as a foreign language in the elementary cycle.

The other school system is that of I.C. I.C is one of the oldest, largest and most expensive private schools in Lebanon. It was established some 90 years ago and was for some time connected with American University of Beirut. Because of its prestige, I.C accommodates a select group of students in terms of academic aptitude. In general, the higher categories on the economic scale can afford the expenses of I.C. I.C. has a spacious campus with modern recreation and sport facilities. Classrooms are adequately equipped with seating facilities and teaching aids. The teachers hold at least a bachelor degree in a school subject plus a teaching diploma in education. Because I.C. can provide lucrative salaries and benefits it usually draws teachers with good and established experience. Again English is the language of instruction at the junior high level.

It should be mentioned that no attempt was made to obtain evidence as to what degree the population from which UNRWA and I.C.

students come is representative of the population of students in Lebanon.

Sample

Samples were drawn from a pool of 8th grade classes in UNRWA and I,C. school systems; however, the samples were not strictly random samples. Random sampling would, for one thing, disrupt the on-going school operation, and, for another, would require liberty in sampling from a large population. Both of these requirements were not administratively possible. Teachers were selected according to predetermined criteria (to be described later), and their classes were randomly assigned to treatments. The students of these classes form the sample for this study.

Ten 8th grade classes participated in the study: Six from URNWA/UNESCO schools and four from I.C. A total of 228 students (out of 310) survived the criteria for inclusion in the analysis. The schools which participated in the study were: Haifa School and Jerusalem School from UNRWA/UNESCO School System and the Intermediate Section of I.C.

Treatments

The treatments were the two teaching methods: T_1 and T_2. T_1 and T_2 differ in the organization and presentation of the same mathematical content.

Identification of Mathematical Content

The mathematical content consisted of the integral powers of 2 and 3 with the operations of multiplication and division. This piece of mathematics was chosen for four reasons:

1. This content is an important segment of secondary school mathematics. In fact, it is a part of the official Lebanese syllabus for grade eight (Lebanese Ministry of Education, 1970).
2. This content uses in a necessary way the integers and the operation of addition on them. This is relevant to the cognitive ability considered in this study; i.e., reflective intelligence. Reflective intelligence was defined as a second order system which consequently, can perceive and act on concepts and operations from sensori-motor system. The concepts of integers and of the operation of addition on them supposedly belong, according to Skemp (1962) to arithmetic and are learnt through the exercise of sensori-motor intelligence. Hence it is legitimate to assume that learning of integral powers of 2 and 3 with the operations of addition and subtraction requires the exercise of reflective intelligence.
3. This content has aesthetic value due to the conciseness and simplicity it brings by unifying many ideas which otherwise seem involved and complicated.

4. This content has applications in mathematics such as logarithms, exponential functions, numeration in different bases; and these have applications in other disciplines.

Characterization of T_1 and T_2

Although the mathematical content for both treatments was the same, a variation of emphasis in T_1 and T_2 could result in different sets of tasks and different possible sequences. The emphasis referred to here is the explicit emphasis on structural properties (in this case infinite cyclic group properties) in developing operations and algorithms and in manipulating isomorphisms among different models of the same theory. Before proceeding to characterize T_1 and T_2, two points ought to be clarified.

The emphasis referred to here is the explicit emphasis which is built consciously and deliberately in the presentation of the content as opposed to mere implicit existence of structural properties which are not deliberately or consciously built in the presentation of the content. It is possible, of course, that many of the implicit structural properties might be inferred by a student without explicitly emphasizing them in teaching.

Explicit emphasis on structure in teaching of mathematics is a matter of degree. For a mathematician, for example, a structure is determined by a set and a theory and the validity of any mathematical sentence is established by a chain of valid propositions (proof).

This level of emphasis was ruled out because of its unfeasibility at the level of instruction proposed in this study. Hence, it was opted for intuition whenever and wherever possible instead of rigor and for avoiding as much as possible technical nomenclature and symbolism.

Specifically T_1 and T_2 differ in the kind of tasks they include and in the possible sequences of presenting these tasks. The mathematical content was analyzed from point of view of each of T_1 and T_2 and schematic diagrams of this analysis are given in Appendix A. The following component behaviors of T_1 were identified:

1. Identifies the relation between the n^{th} power of 2 and the product of n factors of 2 (n is a positive integer).
2. Uses the terms power, base and exponent correctly in examples.
3. Performs the operation defined by $2^n \times 2^m = 2^{n+m}$ where n, m are positive integers.
4. Performs the operation defined by $2^n \div 2^m = 2^{n-m}$ where n, m are positive integers.
5. Deduces (if $n < m$ where n, m are positive integers) that
$$2^{n-m} = \frac{1}{2^{m-n}}.$$
6. Deduces (if $n = m$ where n, m are positive integers) that $2^0 = 1$.

The following target behaviors for T_1 were identified:

7. Solves equations of the form $2^n \times 2^m = \square$ where n and m are integers.

8. Solves equations of the form $2^n + 2^m = \square$ where n and m are integers.

9. Solves equations of the form $\square \times 2^n = 2^m$ and $2^n \times \square = 2^m$ where n and m are integers.

In appendix A, a schematic diagram relating component behaviors, prerequisite behaviors and target behaviors for T_1 is included. Behaviors are numbered in the sequence in which they were taught.

The following component behaviors were identified for T_2:

1. Identifies the relation between the n^{th} power of 2 and the product of n factors of 2 (n is a positive integer).

2. Uses the terms power, base and exponent correctly in examples.

3. Cognizes the set $S_1 = \{2^n$: n is a positive integer$\}$.

4. Performs the operation defined by $2^n \times 2^m = 2^{n+m}$ where n and m are positive integers.

5. Cognizes the set $S_2 = \{2^n$: $n > o$, n is an integer$\}$.

6. Verifies that $2^{-n} = \frac{1}{2^n}$ using $2^o = 1$ and $2^n \times 2^{-n} = 2^o$

7. Cognizes the set $S = \{2^n$: n is an integer$\}$

The following target behaviors were identified for T_2:

8. Solves equations of the form $2^n \times 2^m = \square$ where n and m are integers.

9. Solves equations of the form $2^n \times \square = 2^m$, $\square \times 2^n = 2^m$ where n and m are integers.

10. Solves equations of the form $2^n \div 2^m = 2^{n-m}$ where n and m are integers.

In appendix A, a schematic diagram relating component behaviors, prerequisite behaviors and target behaviors for T_2 is included. Behaviors are numbered in the sequence in which they were taught.

In addition to the above target behaviors for T_2, three additional target behaviors were identified for T_2. These belong to the objective mentioned earlier concerning manipulating isomorphisms among different models. These additional three target behaviors are:

11. To relate the elements of (B,+), (S,x) and (A,x) by: $n \leftrightarrow 2^n \leftrightarrow 3^n$ (n is an integer).
    $B = \{n: n \text{ is an integer}\}$, $S = \{2^n: n \text{ is an integer}\}$
    $A = \{3^n: n \text{ is an integer}\}$
12. To relate the operations of (B,+), (S,x) and (A,x) by: $(n+m) \leftrightarrow 2^n \times 2^m \leftrightarrow 3^n \times 3^m$
13. To relate mathematical sentences in (B,+), (S,x) and (A,x):
    $n + m = \square \leftrightarrow 2^n \times 2^m = 2^{\square} \leftrightarrow 3^n \times 3^m = 3^{\square}$
    $n + \square = \leftrightarrow 2^n \times 2^{\square} = 2^m \leftrightarrow 3^n \times 3^{\square} = 3^m$
    $\square + n = m \leftrightarrow 2^{\square} \times 2^n = 2^m \leftrightarrow 3^{\square} \times 3^n = 3^m$

Differences Between T_1 and T_2

Based on the analysis in Appendix A, the following differences and similarities between T_1 and T_2 can be deduced:

Target Behaviors

The target behaviors of T_1 form a subset of the target behaviors of T_2.

Tasks and Sequence

a) In T_1 powers with zero exponents, negative exponents and multiplication of powers were related to rational numbers and their multiplication (see 3, 4, 5 and 6 of schematic diagram for T_1 in Appendix A).

b) In T_2 the following sequence was used:

1. Multiplication of positive powers was introduced by relating it to multiplication of positive integers (see 4 in the diagram for T_2 in Appendix A).
2. 2^0 was added to the set as a new symbol (see 5.1.3 in the diagram).
3. The operation of multiplication was extended to the new set (see 5.1.1 in the diagram).
4. Using identity property, 2^0 was given the meaning of number 1 (see 5.1 in the diagram).
5. Again using closure property, identity property and

inverse property, symbols for negative powers were introduced to the set (see 7.1 in the diagram)

6. Using these properties, negative powers were given the meaning of reciprocals of positive powers (see 6 in the diagram).

7. The operation of multiplication was then extended to the set of integral powers (see 9 in the diagram)

c) In T_1, division was introduced as a separate operation by relating it to division of integers. Multiplication and division were developed separately, but simultaneously. A separate algorithm was developed for division (see 4 and 5 of the diagram for T_1). No explicit attempt was made to relate multiplication and division.

d) In T_2, division was introduced as an inverse operation to multiplication (see 11.1 of the diagram for T_2). Its introduction was delayed almost until the end and no separate algorithm was developed for division.

e) As a result of a, b, c, and d, T_1 and T_2 have different tasks and different possible sequences. In Appendix A, tasks for T_1 and T_2 are numbered in the sequence in which they were presented.

f) In T_2 isomorphisms were constructed and manipulated among the three models: (S,x) (A,x) and B,+). Nothing explicitly was done along these lines in T_1.

Table 1 gives a summary of the differences between T_1 and T_2.

Table 1

Summary of Differences Between T_1 and T_2

	T_1	T_2
1. Multiplication	Related to rational numbers	Introduced as an operation on a set
2. Zero power	Related to rational numbers	Introduced using identity property
3. Negative power	Related to rational numbers	Introduced as inverse of non-negative powers, using the ideas of closure, identity and inverse
4. Division of powers	a. Separate operation b. Division and multiplication were developed separately but simultaneously c. Separate algorithm was developed	a. Inverse operation to multiplication b. Delayed almost till the end c. No separate algorithm was developed
5. I somorphisms	---------	I somorphism among the models (B,+), (S,x) and (A,x) were constructed and manipulated.
6. Sequence	As shown in the schematic diagram in Appendix A.	As shown in the schematic diagram in Appendix A.

Preparation of Instructional Material

Based on the analysis given in Appendix A, the ideas in each treatment were built into a sequence with each idea or a related group of ideas put in a "frame" (a frame means simply that this idea or group of related ideas were conceived as one unit) Teaching, however, involves, besides identifying and sequencing a content, a strategy or a set of strategies to present the content in a given sequence. In the preparation of the instructional material for this study, the investigator did not decide to use exclusively the discovery approach or the expositary approach. As it can be seen from the lessons of each treatment (Appendix B), teachers were supposed to use a variety of strategies which do not follow exclusively one paradigm. Among these strategies; for example, motivating the introduction of a new concept by examples, giving direction to discussion, asking questions and eliciting responses, guiding individual work of students, verbalizing a conclusion sometimes etc. . . . . . Students were supposed to work examples, discover relations and patterns, respond to questions, work individually, etc. . . . . . In a way, this set of strategies approximates normal teaching, i.e. teaching in ordinary classrooms. An effort was made to keep this general pattern of strategies in both treatments.

The preparation of the instructional material passed through successive stages of modification. These were the main stages:

1. Tasks for each behavior in T_1 and T_2 were constructed. These tasks were put in frames which were sequenced accordingly. The product of this stage was a series of frames which describe in some detail the development of the content from the point of view of T_1 and T_2. No attempt was made at this stage to divide the frames into lessons.

2. The instructional material was then divided into lessons based on common sense. The frames within each lesson were divided into:

   <u>"S" frames:</u> These were meant for individual work by the student under the supervision and guidance of the teacher.

   <u>"T" frames:</u> These were meant to be presented by the teacher for the whole class.

   The product of this stage was a set of lessons (6 for T_1 and 7 for T_2). Each lesson was a set of frames labeled either as "S" or "T".

3. The instructional material was then tried in a pilot study which will be described later. Based on experience gained in this pilot study, the following modifications were made:

   a. The content of each lesson, i.e., the number of frames in each lesson was modified in order that each lesson be covered in 40 minutes approximately.

b. The "S" frames for each of T_1 and T_2 were produced in separate booklets called worksheet booklets. Originally it was thought that the "S" frames could be copied by students on their work books while the teacher wrote them on the board. However, in the pilot study it became obvious that this process was time-consuming.

c. Although the teachers in the pilot study were familiarized with the objectives of each treatment, they were not able to conceive satisfactorily of each lesson as a unit with specified terminal behaviors. For this reason, terminal behaviors for each lesson were inserted as a part of the instructional material.

4. The final form in which the instructional material was used appears in Appendix B. "S" frames were reproduced in separate worksheet booklets for individual students.

A word should be mentioned about lesson 6 in T_1 (which is lesson 7 in T_2). The content of this lesson was not mentioned in the Analysis in Appendix A. However, since some of the criteria measures include "word problems" involving powers, it was thought that some experience with this type of problems should be provided. These two lessons were included for that particular purpose.

Implementation

The project was implemented in two stages: The pilot study and the experiment itself.

Pilot Study

The aims of the pilot study were:

1. To obtain information concerning the teachability and learnability of the instructional material in the form in which it was produced.
2. To obtain examples of actual classroom teaching using the instructional material.
3. To obtain information concerning the feasibility of criteria measures and in particular language level, instructions and possible misconceptions.

Two teachers from UNRWA/UNESCO Malkiyyah school participated in the pilot study. Each teacher taught one section of grade eight according either to T_1 or T_2. A total of 66 students participated in the study.

The teachers were familiarized with the purpose of the research. They were trained in the use of the instructional material in three sessions each of two hours duration. In addition, the investigator visited them during instruction and discussed their feedback and their difficulties. The teachers were encouraged to give their impressions about the instructional material and the responses of the students in the classrooms.

The following techniques were used to obtain the above mentioned information:

1. Recording of all sessions for each of T_1 and T_2.
2. Direct observation of classroom activities and testing sessions.
3. Use of progress tests. These tests were short achievement tests of 5 items each. The tests were prepared to measure the attainment of terminal behaviors of each lesson. Each test was given immediately after the lesson whose objectives the test measure.
4. Feedback from the cooperating teachers. Using these procedures the following information was collected:

Teachability of the Instructional Material

Records of the lessons indicated that it was possible for the teacher in T_1 or T_2 to conform to a high degree to the ideas in "frames" in the suggested sequence.

Records of the lessons, direct observation of classes by the investigator gave some possibilities for improving the instructional material. These possibilities and how they were incorporated were discussed earlier.

Learnability of the Instructional Material

Results of progress tests suggested that in general the instructional material for each of T_1 and T_2 could be learned satisfactorily by eighth graders except for three lessons in T_2. Table 2 gives the percentage of students who scored 4 or more in each test within each treatment. Table 2 shows that in T_1, 75% or more scored 80%

Table 2

Percent of Students Scoring 4 or More in Each Test in Each Treatment in the Pilot Study

	Treatment										
	T_1					T_2					
Test	1	2	3	4	5	1	2	3	4	5	6
% of students with a score of 4 (out of 5) or more	75	83	83	78	75	97	100	39	68	65	90
% of students with a score less than 4	25	17	17	22	25	3	0	61	32	35	10

or more in each of the five tests. In T_2 only in three tests (Tests 1, 2 and 6), 75% scored 80% or more. In each of the remaining three tests of T_2, 70% or less scored 80% or more. These three tests indicated a fairly low achievement level with test 3 the clearest indicator of low achievement (Only 39% scored 80% or more with more than 50% scoring 50% or less). Reasons were sought to explain this in two directions: the nature of the instructional material and the background of the students.

A study of the content of these three lessons revealed that heavy use was made of addition operation on integers and solution of simple linear equations in one unknown. In T_1 no such heavy use was made of the operation of addition on integers because most of the time addition and subtraction of positive numbers were used. Also linear equations in one unknown were never explicitly used in T_1.

A study of the comments of the teacher who was following T_2 revealed that his students experienced difficulties in adding integers. Although the students had studied addition of integers, it was clearly deficient particularly in cognizing the structural properties of this operation. Clearly the solution of the linear equations in one unknown was closely related to the operation of addition on integers. It was realized that adding integers and solving linear equations with one unknown were indispensible for T_2. Hence it was decided that all teachers in the experiment should review these topics with their students and make sure that their students can add two integers and can solve a simple equation with one unknown (the solution set being the set of integers) before starting the experiment. Except for these difficulties in lessons 3, 4 and 5, the teachers reported a higher degree of participation from their students than their usual classes. Since in the majority of lessons 75% scored 4 or more, the latter was adapted as the acceptable achievement level.

Recorded Material for Training

Examples of the recorded lessons in each of T_1 and T_2 were selected with a view to using them as background material in the training of cooperating teachers in the experiment.

Feasibility of Experimental Criteria Scales

Direct observation of the testing sessions by the investigator and cooperating teachers yielded information which was used to improve the tests themselves.

Testing time for each test was modified based on the gained experience. A maximum of 40 minutes was allotted to each of part II and part III (tests of criteria 3 and 4, respectively). A maximum of 25 minutes was allotted for part I (test of criteria 1 and 2).

Instructions were originally given in written English. These instructions turned out to be too much involved to be understood by mere reading. It was decided that, in the experiment, instructions should be given orally and should be translated (to Arabic) and/or repeated whenever necessary. However, pictures and illustrative examples were to be included in a written form in the instructions.

Table 3 gives the K-R reliability coefficients of scales X1 to X6 as calculated from pilot study data. The fact that the reliability coefficients were not high might be attributed to both the small variance and small sample size (66 students). The K-R coefficient for X1 was particularly low, so the number of items for scale X1 was increased from 10 to 15.

Table 3

K-R Coefficient for X1 to X6 from Pilot Study Data

Scale	K-R Coefficient
X1	0.16
X2	0.40
X3	0.42
X4	0.51
X5	0.43
X6	0.65

It was decided not to give a pretest to check the comparability of the base-line knowledge of the students. This decision was motivated by the reports of the cooperating teachers that students were not taught and did not have the chance to know the integral powers of 2 and 3 with multiplication and division operations on them.

Implementation of the Experiment

The experiment was implemented along the lines suggested by the pilot study. The stages of implementation will be discussed under the following headings: selection and training of the teachers, description of the sample, description of the classroom activities and description of students' progress during the experiment.

Selection and Training of Teachers

The teachers who participated in the experiment were selected from International College and UNRWA/UNESCO education system through the proper channels and according to the following criteria:

1. Teachers should be certified mathematics teachers. UNRWA requires at least four years post secondary training (in-service or college) in mathematics and its teaching. I.C. requires at least a university degree with a good background in mathematics.

2. They should be judged to be good mathematics teachers by mathematics supervisors.
3. They should be teaching at least two sections of second preparatory (8th grade).
4. They should voluntarily agree to participate and conform to the requirements of the study.

Four teachers from UNRWA/UNESCO were originally selected, but one was later dropped because of administrative difficulties in his school. Two teachers were selected from I.C. Each teacher taught one of his sections according to T_1 and the other according to T_2.

The training of teachers started almost four weeks before the experiment. Training was done in a combination of self-study by teachers and group discussion. The teachers met in two groups with the investigator: the UNRWA/UNESCO three teachers as one group and the two teachers of International College as another. The group session lasted for six to eight hours.

In the first session each group of teachers was generally oriented to the purpose of the experiment, their expected responsibilities in it and the background which their students should have in order to be able to participate in the experiment. The instructional material for both treatments was distributed and the differences between the two approaches in T_1 and T_2 were explained. At the end of the session the teachers were asked to do two things: first, to start preparing their students for the experiment and in particular to review (or reteach) addition on integers and solution of linear equations with one unknown.

Second, to study the first two lessons of each treatment and, in particular, the idea in each frame and the sequence of the frames.

In the next session, the comments of the teachers were discussed and their queries answered. The phrases "to conform to the idea in each frame" was discussed and illustrated by examples until a general agreement was reached. Segments of the recorded material from the pilot study were used and the group evaluated the extent to which the teacher conformed to the treatment he was following. This pattern continued until the lessons of each treatment were covered. In the last session each teacher was given a kit containing:

(a) A copy of the instructional material.

(b) A copy of instructions for teachers (Appendix C).

(c) Enough copies of "worksheet booklets."

(d) Enough copies of "progress tests."

(e) Progress sheet.

(f) A cassette recorder and tapes.

Description of the Sample

It was anticipated that the sample would be reduced by attrition. Originally, the number of students in the ten 8th grade classes was 310 students. This number was reduced to 228 students. For the purposes of the experiment a student was included in the analyses if:

1. attended all lessons of the treatments to which he was originally assigned.

2. took all tests as scheduled.

240 students survived the two criteria (on the average one student was absent from each lesson or a testing session). However, since an orthogonal design was desirable (equal numbers per cell), random elimination was used to balance the design. Table 4 gives a breakdown of the sample in each school according to sex, age and reflective intelligence. Table 5 gives a breakdown of the sample in each treatment according to the same factors.

Table 4
Number of Students in the Sample in Each School According to Sex, Age and Reflective Intelligence

	Sex		Mean Age (in years)	Reflective Intelligence		
	M	F		(L)	(M)	(H)
UNRW/UNESCO Haifa	-	76	14.3	50	19	7
Jerusalem	51	-	13.8	6	29	16
I.C.	101	-	13.3	20	28	53
Total	152	76	13.8	76	76	76

Table 5

Number of Students in the Sample in Each Treatment According to Sex, Age and Reflective Intelligence

	Sex		Mean Age (in years)	Reflective Intelligence		
	M	F		(L)	(M)	(H)
T_1	80	34	13.8	38	38	38
T_2	72	42	13.7	38	38	38
Total	152	76	13.8	76	76	76

Description of Classroom Activities

The cooperating teachers in the experiment were asked to follow in their classroom teaching the pattern suggested by the treatments (Appendix C). In particular, they were asked to conform to the idea in each frame and to the sequence of the frames. "T" frames were to be presented in the "normal way of teaching" which the teacher was used to following. When the sequence called for an "S" frame each student had to work individually in his own "worksheet booklet" with the teacher guiding individual students. In the last five minutes progress tests were administered. At the end of the lesson the teacher collected the worksheet booklets and stored them with him for the next lesson. After the class, the teachers corrected the progress tests and recorded the scores on the progress sheet. Teachers were only to proceed to the next lesson if at least 75% of the class scored 4 or more. No homework was given during the experiment.

Three factors were possible contributors to variance between treatments in an unknown way: effect of the teacher, non-conformation to the treatments and contamination of the two treatments.

Although it was impossible to completely control these factors, some steps were taken to reduce their effect. It was hoped that the teacher contribution to the variance might be reduced by having the same teacher teach two sections according to T_1 and T_2. Moreover, teachers were trained and instructed explicitly and strongly to follow the treatments as prepared and to follow in their teaching the pattern described earlier. It was hoped that this last step would lead to a better adherence to the requirements of each treatment and would reduce the possibility of contamination. The teachers were instructed to record every other lesson in each treatment. The evidence that the teachers did conform to each treatment was based on occasional visits by the investigator to the classrooms and his evaluation of the recorded lessons. However, the tapes were later lost in an unfortunate accident and consequently it was not possible to give more objective evidence in that respect. The fairly smooth progress of the students in each treatment which shall be described next might support the hypothesis of conformity of the teachers to the treatments.

To reduce contamination by students no homework was given during the experiment. All work was done during the class hour and all materials were collected at the end of each lesson.

Progress of Students During Experiment

Progress tests were used to monitor the achievement of students during instruction. The progress tests were tests of 5 items each constructed along the terminal behaviors of each lesson.

Teachers were instructed not to proceed to the next lesson, unless 75% of the students score 4 or more on the test of the previous lesson. In case this level was not achieved, teachers were instructed to find the difficulties of their students and spend sometime on them. Table 6 shows that at least 83% of the whole sample achieved 4 or more in each lesson of each treatment.

Table 6

Percent of Students Scoring 4 or More in Each Test in Each Treatment in the Experiment

Treatment	T_1					T_2					
Test	1	2	3	4	5	1	2	3	4	5	6
% of scores of 4 or more	96	97	93	90	83	97	97	95	87	89	85
% of scores less than 4	4	3	7	10	17	3	3	5	13	11	15

Experimental Criteria and Measures

Experimental Criteria

The following four criteria were selected as experimental criteria:

C1: Solving the three types of mathematical sentences (which are solvable in a group) in the taught models, i.e., $ax = b$, $xa = b$, $ab = x$, where a, b are powers of 2 or powers of 3.

C2: Solving the same types of mathematical sentences in an isomorphic model, i.e., $ab = x$, $ax = b$, $xa = b$, where a and b are powers of 5.

C3: Selecting and solving mathematical sentences which "model" decisions in an isomorphic model.

C4: Selecting and solving mathematical sentences which "model" decisions in a generalized model.

There are at least two reasons for selecting C1-C4 as criteria: The first is logico-mathematical and the second is empirical. C1 is a direct achievement criterion since the solutions of the mathematical sentences mentioned in it were target behaviors for both T_1 and T_2. C2 is accounted for on a mathematical basis. If two models of the same theory are isomorphic then the solvable mathematical sentences in one model (and their solutions) correspond in a natural way. C3 and C4 cannot be explained exclusively on a mathematical basis, since selecting a mathematical sentence which "model" decisions in an isomorphic model or a generalized model implies a problem solving ability. However, it is logical to assume that, should there be an effect of any treatment on problem-solving ability, it ought to show up, if ever, in dealing with decisions (problems) in an isomorphic model.

The second reason for selecting these criteria is that they reveal the multivariate nature of mathematical achievement - a hypothesis which was strongly supported by National Longitudinal Study of Mathematical Abilities (NLSMA). The four criteria fit respectively in the four categories of the achievement model of NLSMA (Romberg-Wilson, 1969): computation, knowledge, application, and analysis.

Wilson (1970) describes and illustrates in some detail the four cognitive levels. One important subcategory of computation level is identified as "ability to carry out algorithms." Restated, "this is the ability to manipulate elements of a stimulus according to some learned rules" (p. 660). C1 involves solving equations, the method of whose solution (algorithm) was taught and practiced. The behaviors implied in the first criterion seem to fit in the subcategory of "ability to carry out algorithm" of the computation level.

A subcategory of comprehension is "knowledge of principles, rules, and generalizations." Items in this subcategory "pertain to relationships among concepts and problem elements which the student can be expected to know as a result of his course of study." In C1, i.e., solving the same types of equations (as those of C1) in an isomorphic model, it is assumed that the students already studied the rules involved in two examplars and in this way solving the same types of equations in an isomorphic model is a new examplar of previously taught rules. In this sense, C2 belongs to a subcategory of comprehension level.

More clearly, C3 belongs to application level because it involves a sequence of responses closely related to the course of study. The sequence in C3 basically involves selecting and carrying out algorithms which were taught and practiced. This sequence of responses is closely related to what the students studied previously.

In describing the subcategory of analysis level identified as "ability to solve non-routine problems," Wilson (1970) states that:

> "...the objective is to develop the ability to solve problems unlike those which have been solved previously. Such problem solving may involve separating problems and exploring what can be learned about each part. It may involve reorganizing the problem elements in a new way in order to determine a solution. In all cases, the student is given a problem situation for which an algorithmic solution is not available to him ..." (p. 662).

In C4, problem solving is involved since C4 calls for selecting and solving a mathematical sentence. Besides, the items which belong to C4 are in many ways unlike those which have been studied (lesson 6 in T_1 and 7 in T_2) and no algorithmic solution was taught. Moreover each problem had to be separated into at least two parts and each explored separately and in relation to the other part. It seems that C4 fits quite well as a subcategory of analysis level.

Selection of Criteria Measures

At least two measures were taken for each criterion; one immediately following the treatments (achievement) and one two weeks later (retention). Table 7 shows the distribution of measures to criteria.

It will be noted that the words "scale," "measure" and "variable" were used almost interchangeably. Usually "scale" was used to refer to a set (or an equivalent) set of items which might or might not be intact. A scale produced a "measure" of achievement or retention. These measures are the achievement and retention variables. The same symbol was used to denote either scale, measure, or variable. In this sense scale X1 is the same as scale X7 (both use the same test items) but measure or variable X1 is different than measure or variable X7 since the first is an achievement variable and the second a retention variable. The reference should be clear from the context.

Table 7

Classification of Criteria Measures According to Achievement and Retention

	Criterion			
	C1	C2	C3	C4
Achievement	X1	X2	X3, X4	X5, X6
Retention	X7	X8	X9, X10	X11, X12

Two measures were taken for each of C3 and C4 in each of the achievement and retention case in order to provide separate measures for selecting the mathematical sentence which "model" a decision in an isomorphic (or generalized) model and for solving such sentences (i.e., giving the correct answer). So

X3, X9 were measures of selecting the correct mathematical sentences (or any equivalent set) which "model" decision in an isomorphic model.

X4, X10 were measures of giving the correct answers for such mathematical sentences.

X5, X11 were measures of selecting the correct mathematical sentences (or any equivalent set) which "model" decision in a generalized model.

X6, X12 were measures of giving the correct answers for such mathematical sentences.

Incorrect responses for X3, X9, X5, and X11 are difficult to interpret. A wrong response in any one of them might be interpreted in many ways. For one thing it might mean that the student does not know how to write a mathematical sentence (or sentences) which describe the problem. Also it might mean that the student did not care to write the mathematical sentences although he knew how (the students were instructed to write their method of solution). A third possibility is that the students wrote incorrect sentences.

Measures X4, X10, X6 and X12 are less difficult to interpret. They refer to a correct answer irrespective of the existence or nonexistence of a correct method. However, a correct response in any of them does not necessarily imply that the student conceived of a correct method.

Selection of Testing Models

For C1 the models are obviously integral powers of 2 and 3 with the operation of multiplication. For C2, the integral powers of 5 with the operation of multiplication was chosen as a model which is group-isomorphic to the previous models.

For the C3, a model which is described in Appendix D was selected. This model consists of a machine with buttons (R) and (L) and a screen (S) on which numbers appear. If (R) is pressed once, any number on (S) will be doubled. If (L) is pressed once any number on (S) will be halved. The number 1 is assumed to be on the screen originally.

The "pressing" function is a function from the integers to integral powers of 2 defined as follows:

$$P(n) = \begin{cases} P_R(n) = 2^n & n > o \quad P_R\text{: Press (R).} \\ P_L(n) = 2^n & n < o \quad P_L\text{: Press (L).} \end{cases}$$

It follows that:

$$P(n+m) = 2^{n+m} = 2^n \times 2^m = P(n) \times P(m).$$

Since P is 1-1 and onto, P is a group isomorphism from the group of integers with addition and the group of integral powers of 2 with the operation of multiplication.

For C4, a model consisting of two machines similar to the one described above was chosen. It can be thought of the "pressing" function as acting on an ordered pair of integers in this manner:

$$P_1(n,m) = (P(n), P(m)) \qquad P_1\text{: } I \times I \longrightarrow S \times S$$

$$\begin{aligned} P_1((n_1,m_1) + (n_2, m_2)) = P_1(n_1 + n_2, m_1 + m_2) &= (P(n_1+n_2), P(m_1+m_2)) \\ &= (P(n_1)xP(n_2), P(m_1)XP(m_1)) \\ &= (P(n_1), P(m_1))x(P(n_2), P(m_2)) \\ &= P_1(n_1,m_1)x\ P_1(n_2, m_2) \end{aligned}$$

So P_1 can be shown to operate as a group isomorphism from I x I onto S x S which is a generalized model of S.

Universe of Items

1. The universe of items for scales X1 and X7 is the union of the following sets of mathematical sentences:

$S_1 = \{2^n \times 2^m = \square : n, m \in I\}$ (I is the set of integers)

$S_2 = \{2^n \times \square = 2^m : n, m \in I\}$

$S_3 = \{\square \times 2^n = 2^m : n, m \in I\}$

$S_4 = \{3^n \times 3^m = \square : n, m \in I\}$

$S_5 = \{3^n \times \square = 3^m : n, m \in I\}$

$S_6 = \{\square \times 3^n = 3^m : n, m \in I\}$

2. The universe of items for scales X2 & X8 is the union of the following sets of mathematical sentences:

$U_1 = \{5^n \times 5^m = \square : n, m \in I\}$

$U_2 = \{5^n \times \square = 5^m : n, m \in I\}$

$U_3 = \{\square \times 5^n = 5^m : n, m \in I\}$

3. The universe of items for scales X3, X4, X9 and X10 is the set of decision rules which are generated by the isomorphism:

$P(n+m) = P(n) \times P(m) \qquad n, m \in I.$

" "

$2^{n+m} = 2^n \times 2^m$

Three types of open sentences can be generated

$2^n \times 2^m = 2^{\square}$

$2^n \times 2^{\square} = 2^m \qquad n, m \in I$

$2^{\square} \times 2^n = 2^m$

The set of decision rules based on these three types of sentences, i.e., $(\bigcup_{i=1}^{3} S_i)$ is the universe of items for scales X3, X4, X9, and X10.

4. The universe of items for scales X5, X6, X11, and X12 is the set of decision rules which are generated by the isomorphism:

$$P_1\ (n_1, m_1) + (n_2, m_2) = P_1\ (n_1, m_1) \times (n_2, m_2)$$

which is equivalent to:

$$\begin{array}{lcl} P(n_1 + n_2) = P(n_1) \times P(n_2) & & 2^{n_1 + n_2} = 2^{n_1} \times 2^{n_2} \\ & \text{or} & \\ P(m_1 + m_2) = P(m_1) \times P(m_2) & & 2^{m_1 + m_2} = 2^{m_1} \times 2^{m_2} \end{array}$$

Nine types of pairs of open sentences (referred to henceforth component sentences) can be generated. If (n,m) denoted that the box (variable) at the n^{th} position of the first sentence and m^{th} position of the second sentence then the possibilites are:

(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)

The set of decision rules based on these nine types of pairs of open sentences, i.e., $\left(\bigcup_{\substack{i=1,2,3 \\ j=1,2,3}} S_i \times S_j\right)$ is the universe of items for scales X5, X6, X11 and X12.

A decision was made to consider a subuniverse of each universe defined above by restricting n & m as follows: $|n| \leqslant 9$, $|m| \leqslant 9$. This decision may be justified for the following reasons:

1. Our interest in this study is in these mathematical sentences as far as they indicate learning of structure and it is felt that large numbers would complicate computations in a way which might mask the efficiency of learning which took place.
2. Considering a subuniverse will simplify sampling procedures.

Sampling Procedures

Three sampling procedures were contemplated:

a) Random sampling: This was ruled out because a sample of reasonable size would result in tests which require unpractical testing-time.

b) Item sampling: This was ruled out because such a procedure in this case (where the population is originally small) will seriously reduce the power of statistical tests to be used.

c) Stratification with randomizing factors: this procedure was adapted. Important factors in each universe were identified and conscious (hopefully rational) decisions were made as to the number of items to be selected from within types of items according to each factor. The exponents themselves were randomly selected, signs randomly assigned to them and then powers were randomly assigned to types of sentences.

Sampling Items for X1

The following factors were identified:

1. Base: 2 or 3
2. Signs of the exponents: (+ and +), (+ and −), (0 and +), (−&−) and (0 and −).
3. Position of the box in the mathematical sentence: First (1), second (2) and third (3).
4. Kind of operation: Multiplication or division.

15 items were selected using the sampling procedure described above according to the distribution in Table 8.

Table 8

Distribution of Items for X1 According to Base, Sign of Exponent, Position of Box, and Kind of Operation

Base	2														
Sign of exponents	(+ & +)			(+ & −)			(− & −)			(0 & +)			(0 & −)		
Position of box	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Number of items	0	1	0	0	1	1*	1	0	1*	0	0	1	1	0	0

Base	3														
Sign of exponents	(+ & +)			(+ & −)			(− & −)			(0 & +)			(0 & −)		
Position of box	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Number of items	1	0	1	1	0	1	0	1	1*	1	0	0	0	1	0

* The sentences in those items were given in the division form, i.e. $a \div b = c$ where a, b, c are powers of either 2 or 3.

Sampling Items for X2

The same factors as for X1, were identified for X2. Eight items were selected using the sampling procedure described before and according to the following distribution in Table 9.

Table 9

Distribution of Items for X2 According to Base,
Postion of Box, and Kind of Operation

Base	5														
Sign of exponents	(+ & +)			(+ & -)			(- & -)			(0 & +)			(0 & -)		
Position of box	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Number of items	0	2	0	1	0	1*	1	0	1	0	0	1*	0	1	0

* Mathematical sentences in those items were given in division form, i.e. $a \div b = c$ where a, b and c are powers of 5.

Sampling Items for X3 and X4:

The procedure followed here is essentially as before. Mathematical sentences were selected then problems were constructed accordingly. Since a mathematical sentence which involves division is equivalent to some mathematical sentence which involves multiplication, and since both sentences will result in the same problem, no division mathematical sentences were selected. Base 2 was chosen and not base 3.

Five items were selected using the sampling procedure described before and according the following distribution in Table 10.

Table 10

Distribution of Items for X3 and X4 According to Base, Sign of Exponent and Position of Box

Base	2														
Sign of Exponents	(+ & +)			(+ & -)			(- & -)			(0 & +)			(0 & -)		
Position of box	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Number of items	0	1	0	0	1*	0	0	0	1	1	0	0	0	0	1

* For this particular item one term in the sentence having base 3 was included to give one type of a problem with no solution.

Sampling Items for X5 and X6

In principle, 4 integers are needed as exponents for every sentence of type (n, m), however, it was decided to have one of the exponents equal in the component sentences in order to generate more problematic situations. Hence a 3-tuple of integers are needed for every sentence of type (n,m). Again no division forms of sentences are included.

The sentences were selected first and problems constructed accordingly. Five items (i.e., 5 pairs of sentences of type (n,m)) were selected using the sampling procedure used for other previous scales and according to the distribution in Table 11:

Table 11

Distribution of Items for X5 and X6 (Base 2) According to
Sign of Exponents and Position of Box

	Position of box								
Sign of Exponents	(1,1)	(1,2)	(1,3)	(2.1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
1. (+ + +) or (+ + 0)	0	0	0	1	0	0	0	0	0
2. (+ + −) or (+ + 0)	0	0	1	0	0	0	0	0	0
3. (+ − −) or (+ − 0)	0	0	0	0	0	0	0	1	1
4. (− − −) or (− − 0)	0	0	0	0	0	1	0	0	0

Construction of Scales. The items were constructed using the procedures described earlier and then were divided into three papers. The first paper contained items which belonged to scales X1, X2. The second paper contained items which belong to Scales X3, X4. The third paper contained items which belong to Scales X5, X6. Two equivalent forms of each paper were prepared. The papers were administered immediately following the conclusion of the treatments (achievement) and two weeks later (retention). In general, those who took one form in the first administration took the other one in the second administration.

Most of the instructions were given verbally and the investigator supervised almost all of them. Appendix D includes the instruction and the tests used.

Two parts of Skemp test-operation formation (SK6:Part 1) and reflective activity with operations (SK 6: Part II) were used. They were reproduced from a microfilm of Harrison's Ph.D. dissertation (Harrison, 1967). Permission to duplicate and use the tests was obtained from Drs. R. R. Skemp and D. B. Harrison. Appendix E gives the form in which these two parts were used and how they were administrated.

Chapter IV

HYPOTHESES AND STATISTICAL PROCEDURES

In this chapter specific hypotheses are identified. Statistical procedures including experimental design, types of analyses and statistical decision rules are discussed. At last summary statistics of the scales are given.

Hypotheses

This study dealt with five questions:

1. Are there significant treatment differences? For which criteria? For which measures?

2. Are there significant reflective intelligence differences? For which criteria? For which measures?

3. Are there significant treatment differences on difference variables between retention and achievement? For which criteria? For which measures?

4. Are there significant reflective intelligence differences on difference variables between retention and achievement? For which measures?

5. Are there significant differences within each reflective intelligence level? For which criteria? For which measures?

Each question, when applied to the particular criteria and variables, generated a family of null hypotheses. In the following, each family was characterized. The families were identified as Family 1, Family 2, . . . according to the question to which they belong.

Treatment Hypotheses [Family 1]

1. The mean vectors of T_1 and T_2 on achievement and retention variables are equal.

   1.1 The mean vectors of T_1 and T_2 on achievement variables are equal.

      1.1.i (i=1,...,6) The corresponding components of the mean vectors of T_1 and T_2 on achievement variables are equal.

   1.2 The mean vectors of T_1 and T_2 on retention variables are equal.

      1.2.i (i=7,...,12) The corresponding components of the mean vectors of T_1 and T_2 on retention variables are equal.

Reflective Intelligence Hypotheses [Family 2]

2. The mean vectors of the three reflective intelligence levels on achievement and retention variables are equal.

    2.1 The mean vectors of the three reflective intelligence levels on achievement variables are equal.

        2.1.1 The mean vectors of medium and low reflective intelligence levels on achievement variables are equal.

            2.1.1.i (i=1,...,6) The corresponding components of the mean vectors of medium and low reflective intelligence levels on achievement variables are equal.

        2.1.2 The mean vector of high reflective intelligence level is equal to the average of the mean vectors of low and high reflective intelligence levels on achievement variables.

2.1.2.i (i=1,...,6) Each component of the mean vector of high reflective intelligence level is equal to the average of corresponding components of mean vectors of medium and low reflective intelligence levels on achievement variables.

2.2. The mean vectors of the three reflective intelligence levels on retention variables are equal.

2.2.1 The mean vectors of medium and low reflective intelligence levels on retention variables are equal.

2.2.1.i (i=7,...,12) The corresponding components of the mean vectors of medium and low reflective intelligence levels on retention variables are equal.

2.2.2 The mean vector of high reflective intelligence level is equal to the average of mean vectors of medium and low reflective intelligence levels on retention variables.

2.2.2.i Each component of the mean vector
(i=7,...,12)
of high reflective intelligence is equal to the average of corresponding components of the mean vectors of medium and low reflective intelligence levels on retention variables.

<u>Treatment Hypotheses on Difference Variables [Family 3]</u>

3. The mean vectors of T_1 and T_2 on difference variables between achievement and retention are equal.

   3.1 The corresponding components of the mean vectors of T_1 and T_2 on difference variables are equal.

<u>Reflective Intelligence Hypotheses on Difference Scores [Family 4]</u>

4. The mean vectors of the three reflective intelligence levels on difference variables are equal.

   4.1 The mean vectors of medium and low reflective intelligence levels on difference variables are equal.

      4.1.i The corresponding components of the mean vectors
      (i=1,...,6)
      of medium and low reflective intelligence levels on difference variables are equal.

   4.2 The mean vector of high reflective intelligence level is equal to the average of mean vectors of medium and low reflective intelligence levels on difference variables.

      4.2.i Each component of the mean vector of high reflective
      (i=1,...,6)
      intelligence level is equal to the average of the

corresponding components of the mean vectors of medium and low reflective intelligence levels on difference variables.

Treatment Hypotheses within Reflective Intelligence [Family 5]

5. The mean vectors of T_1 and T_2 within each reflective intelligence level are equal on all achievement and retention variables.

5.1 The mean vectors of T_1 and T_2 within each reflective intelligence level are equal on achievement variables.

5.1.1 The mean vectors of T_1 and T_2 within low reflective intelligence level are equal on achievement variables.

5.1.1.i (i=1,...,6) The corresponding components of the mean vectors of T_1 and T_2 within low reflective intelligence level are equal on achievement variables.

5.1.2 The mean vectors of T_1 and T_2 within medium reflective intelligence level are equal on achievement variables.

5.1.2.i (i=1,...,6) The corresponding components of the mean vectors of T_1 and T_2 within medium reflective intelligence level are equal on achievement variables.

5.1.3 The mean vectors of T_1 and T_2 within high reflective intelligence level are equal on achievement variables.

5.1.3.i (i=1,...,6) The corresponding components of the mean vectors of T_1 and T_2 within high reflective intelligence levels are equal on achievement measures.

5.2 The mean vectors of T_1 and T_2 within each reflective intelligence level are equal on retention variables.

5.2.1 The mean vectors of T_1 and T_2 within low reflective intelligence level are equal on retention variables.

5.2.1.i (i=7,...,12) The corresponding components of the mean vectors of T_1 and T_2 within low reflective intelligence level are equal on retention variables.

5.2.2 The mean vectors of T_1 and T_2 within medium reflective intelligence level are equal on retention variables.

5.2.2.i (i=7,...12) The corresponding components of the mean vectors of T_1 and T_2 within medium reflective intelligence level are equal on retention variables.

5.2.3 The mean vectors of T_1 and T_2 within high reflective intelligence level are equal on retention variables.

5.2.3.i (i=7,...12) The corresponding components of the mean vectors of T_1 and T_2 within high reflective intelligence level are equal on retention variables.

In Figure 2, the hypotheses are given in a symbolic concise form for an easy and quick reference. Each null hypothesis is expressed in terms of expected means on some parameters. The first parameter represents the source of variation under consideration and the remaining parameters represent the variables under consideration. A code is included to interpret each source and variable parameters.

Statistical Procedures

Data Unit

The data unit was a vector of repeated measurements on the same subject. Implicit is the assumption that vectors associated with two distinct subjects are independent. Although it was true that students were subject to some common influences, the latter were partially controlled as explained in Chapter II. For the purposes of the analyses which were undertaken, measurement vectors were assumed to be independent.

Experimental Design

The variables in this study were repeated measurements which were related in an unknown way. Consequently, multivariate techniques were used. Two basic designs were used: A two-way additive main-effect orthogonal design and a two-way additive nested main-effect orthogonal design.

Additive main-effect design. This design was used to answer the first four questions. The two-way design has *a* classes in the A-way

	Source of Variation	Variables
<u>Question 1</u> Are there significant treatment differences? For which criteria? For which measures? <u>Treatment Hypotheses [Family 1]</u> 1. $\mu(T_1 - T_2; X1, \ldots, X12) = 0$ 1.1 $\mu(T_1 - T_2; X1, \ldots, X6) = 0$ 1.1.i $\mu(T_1 - T_2; Xi) = 0$ $(i = 1, \ldots, 6)$ 1.2 $\mu(T_1 - T_2; X7, \ldots, X12) = 0$ 1.2.i $\mu(T_1 - T_2; Xi) = 0$ $(i = 7, \ldots, 12)$	T_1: Treatment one T_2: Treatment two $T_1 - T_2$: Contrast between T_1 and T_2	X1: Achievement measure of criterion one (C1) X2: Achievement measure of criterion two (C2) X3, X4: Achievement measures of criterion three (C3) X5, X6: Achievement measures of criterion four (C4) X7: Retention measure of criterion one (C1) X8: Retention measure of criterion two (C2)
<u>Question 2</u> Are there significant reflective intelligence differences? For which criteria? For which measures? <u>Reflective Intelligence Hypotheses [Family]</u> 2. $\mu(L; X1, \ldots, X12) = \mu(M; X1, \ldots, X12) = \mu(H; X1, \ldots, X12)$	L, M, H: low, medium, high reflective intelligence levels	X9, X10: Retention measures of criterion three (3) X11, X12: Retention measures of criterion four (C4)

Figure 2. Schematic diagram of hypotheses.

2.1 $\mu(L; X1, . . ., X6) = \mu(M; X1, . . ., X6)$ $= \mu(H; X1, . . ., X6)$		
2.1.1 $\mu(M - L; X1, . . ., X6) = 0$	M - L: A contrast between M and L.	
2.1.1-i $\mu(M - L; Xi) = 0$ (i = 1, . . ., 6)	ML - H: A contrast between the average of (M, L) and H	
2.1.2 $\mu(ML - H; X1, . . ., X6) = 0$		
2.1.2-i $\mu(ML - H; Xi) = 0$ (i = 1, . . ., 6)		
2.2 $\mu(L; X7, . . ., X12) = \mu(M; X7, . . ., X12)$ $\mu(H; X7, . . ., X12)$		
2.2.1 $\mu(M - L; X7, . . ., X12) = 0$		
2.2.1.i $\mu(M - L; Xi) = 0$ (i = 7, . . ., 12)		
2.2.2 $\mu(ML - H; X7, . . ., X12) = 0$		
2.2.1.i $\mu(Ml - H; Xi) = 0$ (i = 7, . . ., 12)		

Figure 2 continued

<u>Question 3</u> Are there significant treatment differences on difference scores between retention and achievement? For which criteria? For which measures? <u>Treatment Hypotheses on Difference Scores [Family 3]</u> 3. $\mu(T_1 - T_2; D1, \ldots, D6) = 0$ 3.i $\mu(T_1 - T_2; Di) = 0$ $(i = 1, \ldots, 6)$ <u>Question 4</u> Are there significant reflective intelligence differences on difference scores between retention and achievement? For which criteria? For which measure? <u>Reflective Intelligence Hypotheses on Difference Scores [Family 4]</u> 4. $\mu(L; D1, \ldots, D6) = \mu(M; D1, \ldots, D6) = \mu(H; D1, \ldots, D6)$ 4.1 $\mu(M - L; D1, \ldots, D6) = 0$ 4.1.1 $\mu(M - L; Di) = 0$ $(i = 1, \ldots, 6)$		Di = Xi - X(i + 6), where Xi is an achievement measure and X(i + 6) the corresponding retention measures

Figure 2 continued

4.2 $\mu(ML - H; D1, \ldots, D6) = 0$

4.2.i $\mu(ML - H; Di) = 0$

$(i = 1, \ldots, 6)$

Question 5

Are there significant treatment differences within each reflective intelligence level? For which criteria? For which measures?

Treatment Hypotheses within Reflective Intelligence [Family 5]

5. $\mu(T_1 - T_2 \mid L; X1, \ldots, X12) = \mu(T_1 - T_2 \mid M; X1, \ldots, X12)$ $= \mu(T_1 - T_2 \mid H; X1, \ldots, X12)$ $= 0$	$T_1 - T_2 \mid L$:	Contrast of T_1 and T_2 within low R. I.
5.1 $\mu(T_1 - T_2 \mid L; X1, \ldots, X6) = \mu(T_1 - T_2 \mid M; X1, \ldots, X6)$ $= \mu(T_1 - T_2 \mid H; X1, \ldots, X6)$ $= 0$	$T_1 - T_2 \mid M$:	Contrast of T_1 and T_2 within medium R.I.
5.1.1 $\mu(T_1 - T_2 \mid L; X1, \ldots, X6) = 0$ 5.1.1.i $\mu(T_1 - T_2 \mid L, Xi) = 0$ $(i = 1, \ldots, 6)$	$T_1 - T_2 \mid H$:	Contrast of T_1 and T_2 within high R. I.

Figure 2 continued

5.1.2 $\mu(T_1 - T_2|M; X1, \ldots, X6) = 0$

5.1.2.i $\mu(T_1 - T_2|M; Xi) = 0$

$(i = 1, \ldots, 6)$

5.1.3 $\mu(T_1 - T_2|H; X1, \ldots, X6) = 0$

5.1.3.i $\mu(T_1 - T_2|H; Xi) = 0$

$(i = 1, \ldots, 6)$

5.2 $\mu(T_1 - T_2|L; X7, \ldots, X12 = \mu(T_1 - T_2|M; X7, \ldots, X12)$

$= \mu(T_1 - T_2|H; X7, \ldots, X12)$

$= 0$

5.2.1 $\mu(T_1 - T_2)|L; X7, \ldots, X12) = 0$

5.2.1.i $\mu(T_1 - T_2|L; Xi) = 0$

$(i = 7, \ldots, 12)$

5.2.2 $\mu(T_1 - T_2|M; X7, \ldots, X12) = 0$

5.2.2.i $\mu(T_1 - T_2|M; Xi) = 0$

$(i = 7, \ldots, 12)$

5.2.3 $\mu(T_1 - T_2|H; X7, \ldots, 12) = 0$

5.2.3.i $\mu(T_1 - T_2|H; Xi) = 0$

$(i = 7, \ldots, 12)$

Figure 2 continued

classification and b classes in the B-way classification. A description of this model is given in Bock (1968). A model assuming additivity of main-effects is

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \varepsilon_{ijk}$$

μ: p x 1 vector representing the general mean of each variable

α_j, β_k: p x 1 vectors representing the effects of the jth class of A and the kth class of B, respectively.

ε_{ijk}: p x 1 vector representing errors.

ε_{ijk} is assumed to have multivariate normal distribution with zero mean vector and covariance matrix Σ. Again, a common error covariance matrix is assumed in all subclasses.

For the first four questions of the present study, one main effect was treatment with two subclasses (T_1 and T_2) and the other was reflective intelligence (L, M, and H). The design is orthogonal since the number of subjects in each cell was the same (38 subjects). Normality of error distribution was assumed and the pooled error covariance matrix was used. Table 12 lists for each of the first four families the source of variation and associated contrasts with a partitioning of the available five degrees of freedom (between classes). For treatment two orthogonal contrast are considered. No interaction is assumed.

Nested additive main-effect design. This design is a variation of the additive main-effect model except that one main effect is nested within reflective intelligence. Table 13 gives the source of variation and contrasts with a partitioning of the degrees of freedom

Table 12

Source of Variation and Contrasts Associated

with Families 1, 2, 3 and 4

Family	1 and 3	2 and 4		
Source of Variation	Treatment	Reflective Intelligence		Interaction
Contrast	$T_1 - T_2$	ML - H	M - L	
df	1	1	1	2

$T_1 - T_2$: a contrast between the means of T_1 and T_2.

ML - H: a contrast between the average of means of M and L and the mean of H.

M - L: a contrast between the means of M and L.

Table 13

Source of Variation and Contrasts Associated

with Family 5

Source of Variation	Treatment			Reflective Intelligence
Contrast	$T_1 - T_2 \mid L$	$T_1 - T_2 \mid M$	$T_1 - T_2 \mid H$	
df	1	1	1	2

$T_1 - T_2 \mid L$: A contrast between treatment means nested within low reflective intelligence level.

$T_1 - T_2 \mid M$: A contrast between treatment means nested within medium reflective intelligence level.

$T_1 - T_2 \mid H$: A contrast between treatment means nested within high reflective intelligence level.

(between classes). No contrast under reflective intelligence was included since contrasts generated under reflective intelligence in this case were exactly the same as those in the previous model.

Analyses

Multivariate analysis of variance. Before explicitly testing the hypotheses in a family, the statistical significance of discrepancies between the fitted model and the data was established. Bock (1968, p. 106) states that "when the simple additive model is assumed, a test of its goodness-of-fit is equivalent to testing for interaction of main-class effects." Consequently, the first step in the analyses was to test for interaction of main effects. If no significant interation was found, a sequence of analyses for each family of hypotheses was carried out. Figure 3 illustrates the sequence of analyses used for a source of variation on two contrasts

a. A MANOVA for the source of variation on the 12 variables, X1, . . ., X12 was run. If the hypothesis of no differences was rejected then this indicated a significant effect on at least one dependent variable.

b & c. Two MANOVAS were rerun for the same source on achievement variables (X1, . . ., X6) and retention variables (X7, . . ., X12). If the hypothesis of no differences was rejected then this indicated significant effect on at least one dependent variable in each case.

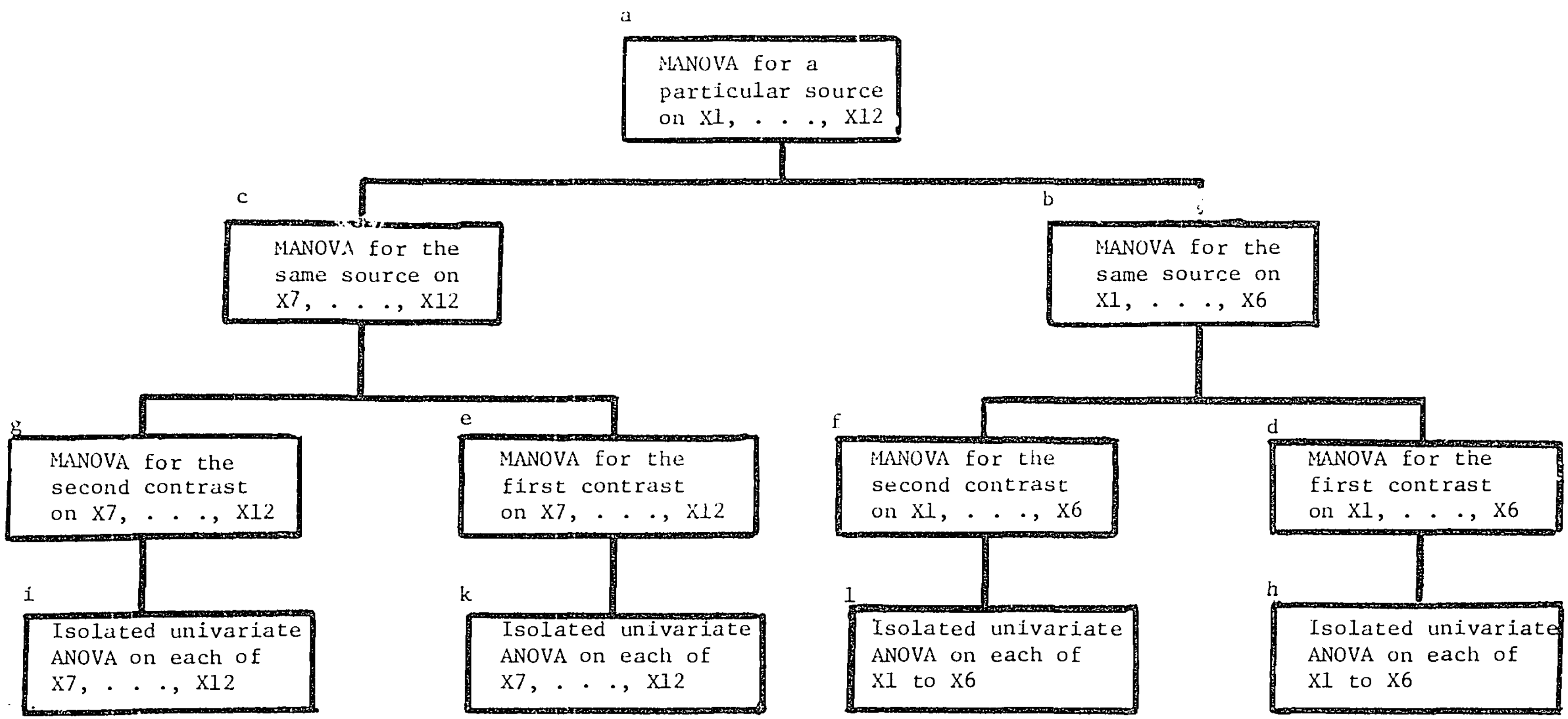


Figure 3. Schema of analyses used to study one particular effect.

GPO 804-930-7

d, e, f & g. a MANOVA was rerun for each contrast belonging to the source under consideration. If the contrast was rejected, then this indicated that the contrast was significant at least on one of the dependent variables.

h, k, l & i. Univariate F-statistics were examined to isolate the significant contrast on the dependent variable.

A similar straregy was used for each source of variation under consideration. In the case of Questions 3 and 4 the difference variables (D1, . . ., D6) were used.

Discriminant analysis. An additional means of characterizing contrasts for a certain effect was discriminant analysis. A linear function of the dependent variables was determined which maximally separated the groups with respect to between group variation. If the number of degrees of freedom was two or more, a second discriminant function was determined, statistically independent of the first. In case the discriminant function was readily interpretable, a description of its interpretation was given. Analyses were done using MULTIVARIANCE FORTRAN IV program as adopted by Madison Academic Computing Center.

Statistical Decision Rules

Interest in the present study focused on the five families of hypotheses as separate entities as well as on the hypotheses within each family. Each family helps give a global answer to one of the five basic questions whereas the hypotheses in each family help particularize the global answer to particular criteria and variables. Although the five families are separate and distinct, hypotheses within the same family are not statistically independent. In the 2 x 3 multivariate design, at most five independent multivariate contrasts can be constructed. Some families of hypotheses contain more than five hypotheses (Families 2 and 5). Moreover each of the five families includes a large number of univariate hypotheses. Given enough statistical tests, the risk is that some of them might come out to be significant by chance.

The area of multiple comparison is already confusing in the univariate case (Games, 1971). One would assume that multiple comparison is more involved in the multivariate case particularly that few techniques exist (at least in applied statistics book). Roy (1957) gave a procedure to construct a simultaneous confidence interval through which an infinite number of contrasts can be tested under the same type I error. However, Roy did not use the F-distribution but the greatest root distribution.

This being the case, a plausible approach was to cast decision rules in a conservative form. One way to do that was to decrease the probability of type I error which was kept at .01 or less for each hypothesis. It was decided also not to test a hypothesis in a family unless the immediately preceding hypothesis had been rejected. This last decision would prevent awkward results such as rejecting a null hypothesis while accepting a hypothesis of larger scope which subsumes it. Moreover it would reduce the number of hypotheses to be tested thus reducing the family-wise type I error.

The univariate F's should be conceived only as "isolated" tests on a particular variable, in the sense that they do not take in consideration the relationships among other variables. Most of the time our interest in a variable is not only in its isolated effect but also in its contribution to discrimination between classes when other variables are considered. The discriminant function provided this latter technique. If the discriminant function was significant at $\alpha = .01$ as tested by Barlett test (Bock, 1965), then its standardized coefficients were examined for a meaningful characterization of the contribution of each variable to the discrimination between classes under consideration.

Each null hypothesis was tested against the alternative of no mean differences. In the case of univariate hypotheses, directional alternative hypotheses were considered. Direction was judged from the observed algebriac value of the contrast under consideration.

Summary Statistics of Scales

Table 14 shows summary statistics of the scales X1 to X6. Scales X7, . . ., X12 are equivalent (one-by-one) to scales X1 to X6.

Included in Table 14 are the mean, standard deviation of forms A and B for each scale and the number of subjects who took each form. The t test was used to establish the statistical

Table 14

Summary Statistics for Scales X1 to X6

Scale	X1		X2		X3		X4		X5		X6	
Form	A	B	A	B	A	B	A	B	A	B	A	B
Mean*	9.33	9.53	5.22	5.04	1.75	2.10	2.65	2.61	1.24	1.14	1.24	1.12
Standard Deviation*	3.26	3.40	1.98	1.73	1.33	1.39	1.53	1.55	1.40	1.49	1.54	1.63
Number of Subjects	114	114	114	114	110	118	110	118	107	121	107	121
t-Statistic	.19 < 1		.26 < 1		.62 < 1		.06 < 1		.17 < 1		.10 < 1	
Hoyt Coefficient	.78	.80	.72	.62	.61	.63	.64	.67	.72	.77	.78	.85

* of correct responses

equivalence of forms A and B for each scale. Each of the t-statistics in Table 14 was quite small (less than 1) indicating that the two means of forms A and B were statistically equal.

Hoyt coefficient are also included for each form of each scale. Computations were done at Madison Academic Computing Center using the Generalized Item Test Analysis Program (GITAP).

Chapter V

RESULTS

Results are presented for each of the five families of hypotheses: (1) treatment hypotheses, (2) reflective intelligence hypotheses, (3) treatment hypotheses on difference variables, (4) reflective intelligence hypotheses on difference variables and (5) treatment hypotheses within reflective intelligence. In each family, analyses of hypotheses are presented in the order given in Chapter IV and analysis was not carried on if the immediately preceding multivariate hypothesis in a family was not rejected. The results of multivariate analyses of variance for one or more hypothesis are reported in tables each of which is identified by (1) the family to which the hypothesis belongs, and (2) a statement of the hypothesis in a symbolic form as it exactly appears in Figure 3 of Chapter IV. All entries in the tables were rounded to two decimal places (computer print-outs give 4 decimal places).

Table 15 gives the vectors of observed means and standard deviations of achievement and retention measures. Table 16 gives the estimates and associated standard errors of treatment and reflective intelligence contrasts on achievement and retention variables. Figure 4 gives the achievement profiles of T_1 and T_2 and Figure 5 gives the retention profiles of T_1 and T_2.

Table 15

Vector Cell Means and Vector Cell Standard Deviation on Achievement and Retention Variables

Cell	Variables											
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
T_1: Low(L)	7.55 (2.64)	3.71 (1.78)	1.33 (1.17)	2.03 (1.60)	0.47 (0.86)	0.45 (0.89)	7.13 (2.58)	3.76 (1.75)	1.10 (1.18)	1.26 (1.54)	0.34 (0.78)	0.45 (1.16)
Medium(M)	7.68 (2.78)	4.37 (1.46)	2.32 (1.14)	2.87 (1.19)	1.13 (1.19)	0.84 (1.08)	8.53 (2.83)	4.55 (1.66)	1.89 (1.13)	2.63 (1.46)	0.55 (0.95)	0.89 (1.09)
High(H)	11.39 (2.42)	6.08 (1.50)	2.84 (1.13)	3.53 (1.33)	2.13 (1.63)	2.18 (1.69)	11.58 (2.42)	6.13 (1.17)	2.24 (1.40)	3.89 (1.13)	1.66 (1.70)	2.32 (1.68)
T_2: Low(L)	7.84 (2.70)	4.63 (1.53)	1.29 (1.06)	1.55 1.13)	0.71 (1.11)	0.34 (0.94)	8.13 (2.88)	4.39 (1.76)	1.16 (1.00)	1.68 (1.32)	0.32 (0.81)	0.29 (0.77)
Medium(M)	9.45 (3.42)	5.24 (1.73)	1.74 (1.33)	2.29 (1.39)	0.89 (0.98)	1.00 (1.56)	9.42 (2.59)	5.08 (1.55)	1.45 (1.18)	2.15 (1.41)	0.50 (1.06)	0.82 (1.35)
High(H)	12.66 (2.03)	6.76 (1.30)	2.00 (1.69)	3.53 (1.47)	1.79 (1.92)	2.26 (1.91)	11.92 (2.41)	6.39 (1.31)	1.87 (1.47)	3.45 (1.41)	1.39 (1.81)	2.47 (1.75)

Note - The maximum mean score of X1 and X7 is 15; for X2 and X8 is 8; and for the rest it is 5.

- Standard deviations are in parentheses.

Table 16

Least Square Estimates and Standard Errors of Treatment and Reflective Intelligence Contrasts on Achievement and Retention Variables

Contrast	Variable											
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
$T_1 - T_2$ [a]	-1.11 (0.36)*	-0.82 (0.21)	0.51 (0.17)	0.35 (0.18)	0.11 (0.18)	-0.04 (0.19)	-0.75 (0.35)	-0.47 (0.20)	0.25 (0.16)	0.17 (0.18)	0.11 (0.17)	0.03 (0.18)
ML-H [b]	-7.79 (0.76)	-3.87 (0.44)	-1.47 (0.36	-2.68 (0.38)	-2.32 (0.36)	-3.13 (0.39	-6.89 (0.74)	-3.63 (0.43)	-1.30 (0.35)	-3.47 (0.39)	-2.20 (0.35)	-3.57 (0.38)
L-M [c]	-0.86 (0.44)	-0.63 (0.25)	-0.68 (0.21)	-0.79 (0.22)	-0.42 (0.22)	-0.53 (0.23)	-1.34 (0.43)	-0.74 (0.25)	-0.54 (0.20)	-0.92 (0.22)	-0.20 (0.20)	-0.49 (0.22)

Note - Standard error in parenthesis.

[a] $T_1 - T_2$: A contrast between treatment one and two.

[b] ML-H : A contrast between the average of medium and low and that of high reflective intelligence levels.

[c] L-M : A contrast between low and medium reflective intelligence levels.

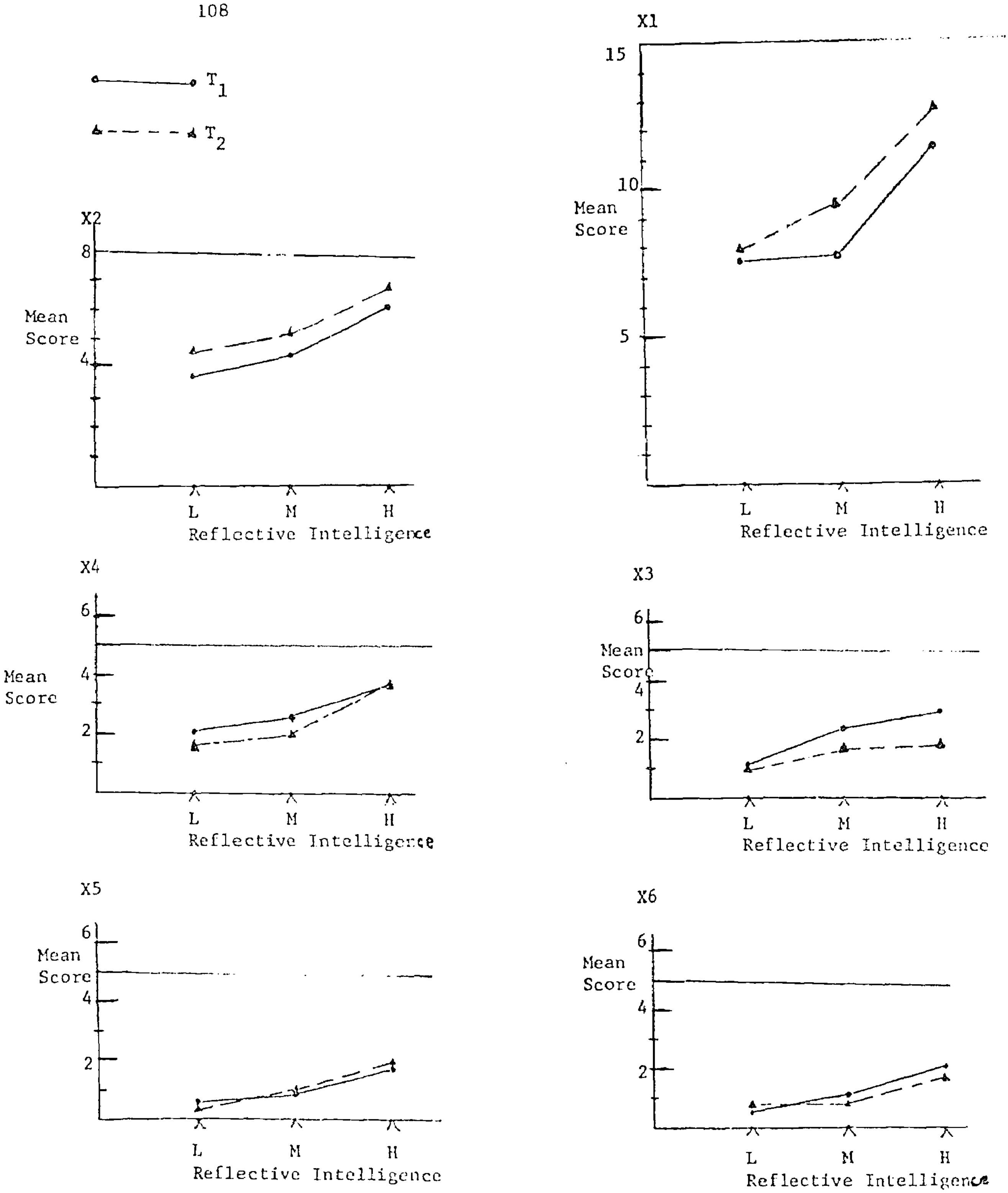


Figure 4

Achievement Profiles for T_1 and T_2

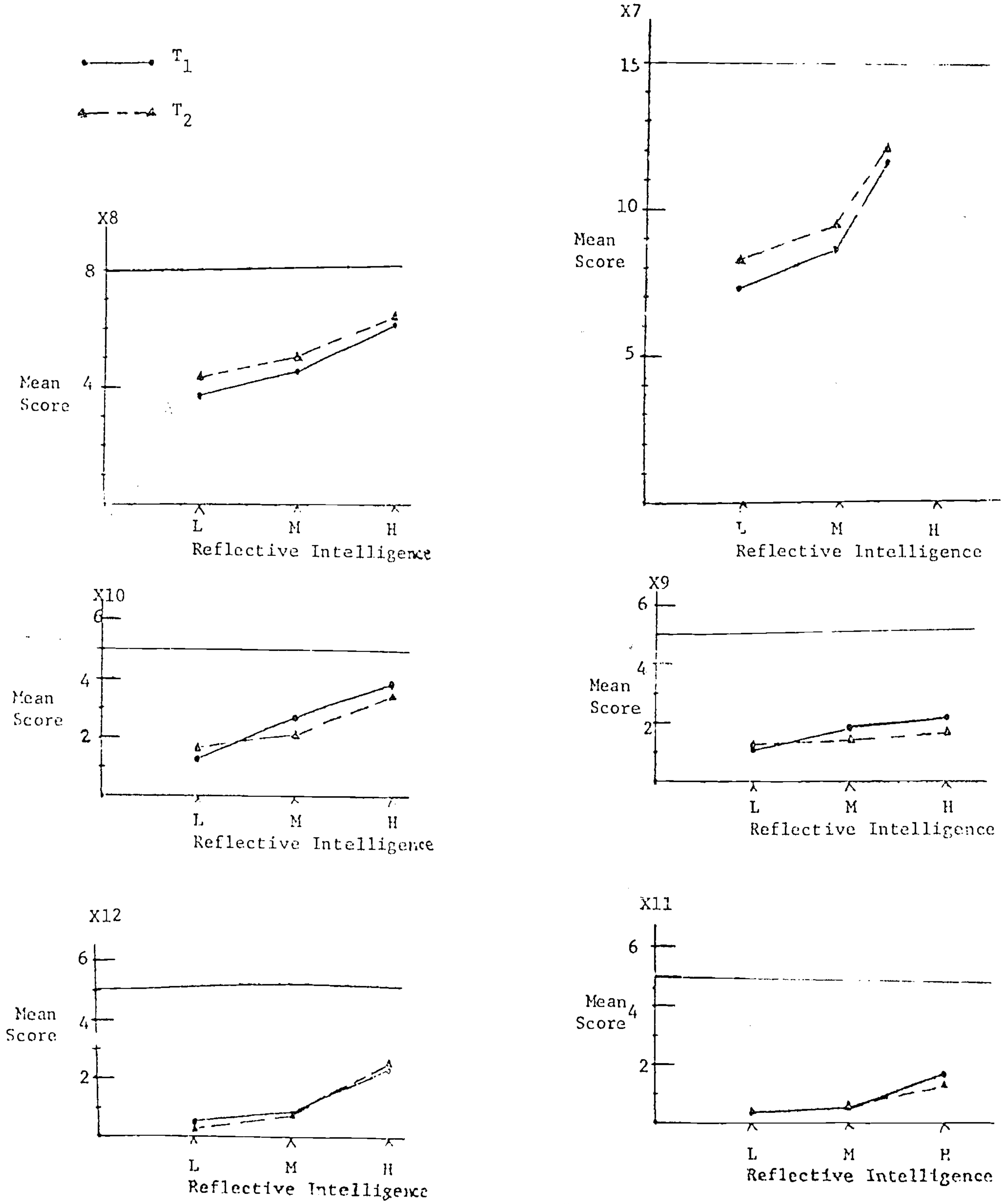


Figure 5
Retention Profiles

Treatment Hypotheses [Family 1]

Table 17 gives the results of ANOVA of treatment and reflective intelligence interaction on achievement and retention variables. The multivariate F was not significant at $\alpha = 0.01$ ($p < 0.05$). Since a test of interaction is equivalent to a test of goodness-of-fit of the data to the simple additive model, the latter was accepted as a tenable model.

Hypothesis 1

This hypothesis focused on the mean vector of T_1 on achievement and retention variables as contrasted with the mean vector of T_2 on the same variables. Results of ANOVA for the hypothesis are reported in Table 18. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypothesis 1.1 and 1.1.i (i=1,...,6). Hypothesis 1.1 focused on the mean vector of T_1 on achievement variables as contrasted with the mean vector of T_2 on the same variables. Hypothesis 1.1.i(i=1,...,6) focused on the corresponding components of the two mean vectors. The results of ANOVA's for the hypotheses are reported in Table 19. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). Each of the isolated univariate F for X1, X2 and X3 was significant at $\alpha = 0.01$ ($p < 0.00$). The coefficients (raw and standardized) of discriminant function associated with hypothesis 1.1 are reported in Table 20. The discriminant function was significant at $\alpha = 0.01$ ($p < 0.00$). The standardized coefficients

Table 17

MANOVA for Interaction of Treatment and Reflective Intelligence on Achievement and Retention Variables

F-ratio for Multivariate Test of Equality of Mean Vectors = 1.55

df = 24 and 422 , p less than 0.05

Table 18

Treatment Hypotheses

Hypothesis 1: $\mu\ (T_1 - T_2;\ X1, \ldots, X12) = 0$

F - ratio for Multivariate Test of Equality of Mean Vectors = 3.40

df = 12 and 211 p less than 0.00

Table 19

Treatment Hypotheses

Hypothesis 1.1: $\mu\ (T_1 - T_2;\ X1, \ldots, X6) = 0$ and Hypotheses 1.1.i: $\mu\ (T_1 - T_2; Xi) = 0$ $(i = 1, \ldots, 6)$

F-ratio for Multivariate Test of Equality of Mean Vectors = 6.68

df = 6 and 217 p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	69.63	9.57	0.00
X2	38.75	15.93	0.00
X3	14.75	9.11	0.00
X4	7.02	3.79	0.05
X5	0.74	0.42	0.52
X6	0.11	0.06	0.81

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 20

Discriminant Function for Hypothesis 1.1: $\mu(T_1 - T_2;\ X1, \ldots, X6) = 0$

Variable	Raw Coefficient	Standardized Coefficient
X1	-.18	-.49
X2	-.38	-.59
X3	.47	.60
X4	.29	.39
X5	-.03	-.04
X6	.04	.05

Accounts for 100% of canonical variance

$\chi^2 = 37.10$ with df = 6, p less than 0.00

indicate that the discriminant function acts as a contrast between X1, X2 and X3, X4 (the remaining coefficients were relatively small). A plausible description of this discriminant function is that it is a contrast of measures of criteria C1 & C2 (computation and comprehension) and a measure of criterion C3 (application).

Hypothesis 1.2 and 1.2.i (i=7,...,12). Hypothesis 1.2 focused on the mean vector of T_1 on retention variables as contrasted with the mean vector of T_2 on the same variables. Hypotheses 1.2.i (i=7,...,12) focused on the corresponding components of the two mean vectors. The results ANOVA's for the hypotheses are reported in Table 21. The multivariate F for the test of equality of mean vectors was not significant at $\alpha = 0.01$ ($p < 0.05$). None of the univariate F's was significant at $\alpha = 0.01$. Table 22 gives the raw and standardized coefficients of the discriminant function associated with hypothesis 1.2. The discriminant function was not significant at $\alpha = 0.01$ ($p < 0.06$).

Reflective Intelligence Hypotheses [Family 2]

Hypothesis 2

This Hypothesis focused on the differences among the mean vectors of the three reflective intelligence levels on achievement and retention

Table 21

Treatment Hypotheses

Hypotheses 1.2: $\mu(T_1-T_2;X7,\ldots,X12=0$ and Hypotheses 1.2.i: $\mu(T_1-T_2;Xi)=0$ $(i=7,\ldots12)$

F-ratio for Multivariate Test of Equality of Mean Vectors=2.05
df = 6 and 217 p less than 0.06

Variable	Hypothesis Mean Square	Univariate F	p less than
X7	31.69	4.61	0.03
X8	12.79	5.34	0.02
X9	3.69	2.40	0.12
X10	1.58	0.83	0.36
X11	0.74	0.47	0.49
X12	0.04	0.02	0.88

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 22

Discriminant Function for Hypothesis 1.2: $\mu(T_1-T_2;\ X7,\ldots,X12)=0$

Variable	Raw Coefficient	Standardized Coefficient
X7	.18	.46
X8	.41	.63
X9	-.40	-.46
X10	-.24	-.33
X11	-.03	-.04
X12	-.05	-.07

Accounts for 100% of canonical variance
χ^2 = 12,07 with df = 6 , p less than 0.06

variables. The results of ANOVA for the hypothesis is reported in Table 23. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypothesis 2.1. This hypothesis focused on the differences among the mean vectors of the three reflective intelligence levels on achievement variables. The results of ANOVA for the hypothesis are reported in Table 24. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). The discriminant coefficients (raw and standardized) for the first two discriminant functions are reported in Table 25. The first discriminant function accounted for 93% of canonical variance and was significant at $\alpha = 0.01$ ($p < 0.00$). The second discriminant function accounted for only 7% of the variance and was not significant at $\alpha = 0.01$ ($p < 0.08$). For the first discriminant function, almost all coefficients (except that of X3) have the same sign and all (except that of X1) have the same size approximately. A plausible interpretation for the 1st discriminant function is that the six variables tend to discriminate almost equally and in the same direction among the three levels of reflective intelligence.

Hypotheses 2.1.1 and 2.1.1.i (i=1,...,6). Hypothesis 2.1.1 focused on the mean vector of low reflective intelligence on achievement variables as contrasted with the mean vector of medium reflective intelligence on the same variable. The results of ANOVA's for the hypotheses are reported in Table 26. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). Isolated univariate F's

Table 23

Reflective Intelligence Hypotheses

Hypothesis 2: $\mu(L;X1,\ldots,X12) = \mu(M;X1,\ldots,X12) = \mu(H;X1,\ldots X12)$

F-ratio for Multivariate Test of Equality of Mean Vectors = 6.88

df = 24 and 422 , p less than 0.00

Table 24

Reflective Intelligence Hypotheses

Hypotheses 2.1: $\mu(L;X1,\ldots,X6) = \mu(M;X1,\ldots,X6) = \mu(H;X1,\ldots,X6)$

F-ratio for Multivariate of Equality of Mean Vectors = 10.50

df = 12 and 434 p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	398.61	54.77	0.00
X2	102.36	42.08	0.00
X3	22.65	13.99	0.00
X4	57.47	31.01	0.00
X5	37.33	20.94	0.00
X6	67.37	34.30	0.00

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 25

Discriminant Functions for Hypothesis 2.1: $\mu(L;X1,\ldots,X6) = \mu(M;X1,\ldots,X6) = \mu(H;X1,\ldots,X6)$

1st Discriminant Functions

Variable	Raw Coefficient	Standardized Coefficient
X1	-.20	-.55
X2	-.15	-.24
X3	.09	.12
X4	-.17	-.23
X5	-.12	-.16
X6	-.11	-.15

2nd Discriminant Function

Variable	Raw Coefficient	Standardized Coefficient
X1	.35	.95
X2	-.30	-.46
X3	-.59	-.75
X4	-.41	-.56
X5	.17	.23
X6	.06	.09

1st Discriminant Function Accounts for 93% of canonical Variance

$\chi^2 = 111.92$ with df = 12 , p less than 0.00

2nd Discriminant Function Accounts for 7% of canonical variance

$\chi^2 = 9.84$ with df = 5 , p less than 0.08

Table 26

Reflective Intelligence Hypotheses

Hypothesis 2.1.1: $\mu(L-M;X1,\ldots,X6)=0$ and 2.1.1.i: $\mu(L-M,Xi) = 0$ $(i=1,\ldots,6)$

F-ratio for Multivariate Test of Equality of Mean Vectors=2.83

df = 6 and 217 , p less than 0.01

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	28.66	3.94	0.05
X2	15.16	6.23	0.01
X3	17.78	10.98	0.00
X4	23.68	12.78	0.00
X5	6.74	3.78	0.05
X6	10.53	5.36	0.02

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

for X2, X3, and X4 were significant at $\alpha = 0.01$ ($p < 0.00$).

Hypotheses 2.1.2 and 2.1.2.i. Hypothesis 2.1.2 focused on the mean vector formed by the average of the mean vectors of medium and low reflective intelligence on achievement variables as contrasted with the mean vector of high reflective intelligence level on the same variables. Hypotheses 2.1.2.i (i=1,...,6) focused on the corresponding components of those vectors. The results of ANOVA's for the hypotheses are reported in Table 27. The multivariate F and all the isolated univariate F's were all significant at $\alpha = 0.01$ ($p < 0.00$ for each).

Hypothesis 2.2. This hypothesis focused on the differences among the mean vectors of three reflective intelligence levels on retention variables. The results of ANOVA's for the hypothesis is reported in Table 28. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). The discriminant coefficients (raw and standardized) for the first two discriminant functions are given Table 29. The first discriminant function accounted for 95% of canonical variance and was significant at $\alpha = 0.01$ ($p < 0.00$). The second discriminant function accounted only for 5% of the variance and was not significant at $\alpha = 0.01$ ($p < 0.18$). For the first discriminant function, variables X7 (measure of C1), X8 (measure of C2), X10 (measure of C3) and X12 (measure of C4) seem to discriminate equally and in the same direction among levels of reflective intelligence.

Table 27

Reflective Intelligence Hypotheses

Hypothesis 2.1.2: μ(ML-H; X1,...,X6)=0 and 2.1.2.i: μ(ML-H;Xi)=0
(i=1,...,6)

F-ratio for Multivariate Test of Equality of Mean Vectors=20.24

df = 6 and 217 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	768.56	105.60	0.00
X2	189.55	77.93	0.00
X3	27.51	16.99	0.00
X4	91.26	49.24	0.00
X5	67.93	38.09	0.00
X6	124.22	63.23	0.00

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 28

Reflective Intelligence Hypotheses

Hypothesis 2.2: $\mu(L,X7,...X12) = \mu(M;X7,...,X12) = \mu(H;X7,...,X12)$

F-ratio for Multivariate Test of Equality of Mean Vectors = 12.01

df = 12 and 434 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X7	335.29	48.73	0.00
X8	93.84	39.19	0.00
X9	16.28	10.60	0.00
X10	92.54	48.40	0.00
X11	31.32	19.91	0.00
X12	85.03	47.17	0.00

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 29

Discriminant Functions for Hypothesis 2.2: $\mu(L;X7,\ldots,X12) = \mu(M;X7,\ldots,X12) = \mu(H;X7,\ldots,X12)$

1st Discriminant Function

Variable	Raw Coefficient	Standardized Coefficient
X7	-.14	-.37
X8	-.15	-.23
X9	.10	.12
X10	-.31	-.44
X11	.01	.02
X12	-.25	-.33

2nd Discriminant Function

Variable	Raw Coefficient	Standardized Coefficient
X7	-.06	-.09
X8	.06	.09
X9	-.65	-.81
X10	-.35	-.48
X11	.69	.87
X12	.23	.31

1st Discriminant Function Accounts for 95% of canonical variance.

$\chi^2 = 125.84$ with df = 12, p less than 0.00.

2nd Discriminant Function Accounts for 5% of Canonical Variance.

$\chi^2 = 7.54$ with df = 5, p less than 0.18

Hypotheses 2.2.1 and 2.2.1.i. Hypothesis 2.2.1 focused on the mean vector of low reflective intelligence on retention variables as contrasted with the mean vector of medium reflective intelligence on the same variables. Hypotheses 2.2.1.i (i = 7,...,12) focused on the corresponding components of these mean vectors. The results of ANOVA's for all the hypotheses are reported in Table 30. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). The isolated univariate F's for X7, X8, X9, X10 were significant at $\alpha = 0.01$ ($p < 0.00$).

Hypothesis 2.2.2 and 2.2.2.i (i = 7,...,12). Hypothesis 2.2.2 focused on the mean vector formed by the average of mean vectors of low and medium reflective intelligence on retention variables as contrasted with the mean vecgor of high reflective intelligence on the same variables. Hypotheses 2.2.2.i (i = 7,...,12) focused on the corresponding components of these two vectors. The results of ANOVA's for the hypotheses are reported in Table 31. The multivariate F and each of the isolated univariate F's were significant at $\alpha = 0.01$ ($p < 0.00$ for each).

Treatment Hypotheses on Difference Variables [Family 3]

A difference variable, it will be remembered, was defined as: $Di = Xi - X(i+6)$ (i = 1, ...,6). Table 32 shows that the interaction of treatment and reflective intelligence was not significant at $\alpha = .01$ ($p < .02$). Hence the simple main effect model was tenable.

Table 30

Reflective Intelligence Hypotheses

Hypothesis 2.2.1: μ(L-M, X7,...,X12)=0 and Hypotheses 2.2.1.i: μ(L-M,Xi)=0 (i=7,...,12)

F-ratio for Multivariate Test of Equality of Mean Vectors = 3.81
df = 6 and 217 , p less than .00

Variable	Hypothesis Mean Square	Univariate F	p less than
X7	68.45	9.95	0.00
X8	20.63	8.61	0.00
X9	11.06	7.20	0.00
X10	32.24	16.86	0.00
X11	1.48	0.94	0.33
X12	9.01	4.10	0.03

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 31

Reflective Intelligence Hypotheses

Hypothesis 2.2.2: $\mu(ML-H; X7,\ldots,X12) = 0$ and Hypotheses 2.2.2.i: $\mu(ML-H,Xi)=0$ $(i=7,\ldots,12)$

F-ratio for Multivariate Test of Equality of Mean Vectors 23.29
df = 6 and 217 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X7	602.14	87.52	0.00
X8	167.05	69.77	0.00
X9	21.49	14.00	0.00
X10	152.84	79.94	0.00
X11	61.16	38.87	0.00
X12	161.05	89.34	0.00

Degrees of Freedom for Mulitvariate Hypothesis = 1
Degrees of Freedom for Error = 222

Hypothesis 3

This hypothesis focused on the mean vector of T_1 on difference variables as contrasted with the mean vector of T_2 on the same variables. The results of ANOVA for the hypothesis is reported in Table 33. The multivariate F for the test of equality of mean vectors was not significant at $\alpha = 0.01$ ($p < 0.32$). Since this was the case, no subordinate hypothesis was tested.

Reflective Intelligence Hypotheses on Difference Variables [Family 4]

Hypothesis 4

This hypothesis focused on the differences among the mean vectors of the three reflective intelligence levels on difference variables. The results of ANOVA for the hypothesis is reported in Table 34. The multivariate F for equality of mean vectors was not significant at $\alpha = 0.01$ ($p < 0.18$); hence no subordinate hypothesis was further tested.

Treatment Hypotheses Within Reflective Intelligence [Family 5]

In this family, three contrasts were under consideration T_1-T_2/L, T_1 - T/M and T_1 - T_2/H. Estimates of the three contrasts with the associated standard error on achievement and retention variables are reported in Table 35.

Hypothesis 5

This hypothesis focused on the mean vector of T_1 on achievement

Table 32

MANOVA for the Interaction of Treatment and Reflective Intelligence on Difference Variables

F - ratio for Multivariate Test of Equality of Mean Vectors=2.04
df = 12 and 434 , p less than .02

Table 33

Treatment Hypotheses on Difference Variables

Hypothesis 3: $\mu(T_1-T_2;\ D1,\ldots,D6) = 0$
F-ratio for Multivariate Test of Equality of Mean Vectors = 1.17 df = 6 and 217 , p less than 0.32

Table 34

Reflective Intelligence Hypotheses on Difference Variables

Hypothesis 4: $\mu(L, D1,\ldots,D6) = \mu(M, D1,\ldots,D6) = \mu(H,D1,\ldots,D6) = 0$
F - ratio for Multivariate Test of Equality of Mean Vectors = 1.36
df = 12 and 434 , p less than 0.18

Table 35

Least Square Estimates of Treatment Within Reflective Intelligence Contrasts on Achievement and Retention Variable

Contrast	Variable											
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
$T_1 - T_2/L$[a]	-0.29 (0.62)	-0.92 (0.36)	0.11 (0.29)	0.47 (0.31)	-0.24 (0.31)	0.11 (0.32)	-1.00 (0.60)	-0.63 (0.35)	-0.06 (0.28)	-0.42 (0.32)	0.03 (0.29)	0.16 (0.31)
$T_1 - T_2/M$[b]	-1.76 (0.62)	-0.87 (0.36)	0.58 (0.29)	0.58 (0.31)	0.24 (0.31)	-0.16 (0.32)	-0.89 (0.60)	-0.53 (0.35)	0.45 (0.28)	0.47 (0.32)	0.05 (0.29)	0.08 (0.31)
$T_1 - T_2/H$[c]	-1.26 (o.62)	-0.68 (0.36)	0.84 (0.29	0.00 (0.31)	0.34 (0.31)	-0.08 (0.32)	-0.34 (0.60)	-0.26 (0.35)	0.37 0.28	0.45 0.32	0.26 0.29	-0.16 (0.31)

Note - Standard errors in parenthesis.

[a] $T_1 - T_2/L$: A contrast between T_1 and T_2 within low reflective intelligence level

[b] $T_1 - T_2/M$ A contrast between T_1 and T_2 within medium reflective intelligence level

[c] $T_1 - T_2/H$ A contrast between T_1 and T_2 within high reflective intelligence level.

and retention variables within each reflective intelligence level as contracted with the mean vector of T_2 on the same variables within each reflective intelligence level. The results of ANOVA for the hypothesis are reported in Table 36. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypothesis 5.1. This hypothesis focused on the mean vector of T_1 on achievement variables within each reflective intelligence level as contrasted with the mean vector of T_2 on the same variables within each reflective intelligence level. The results of ANOVA for the hypothesis are reported in Table 37. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypotheses 5.1.1 and 5.1.1.i (i = 1, ...,6). Hypothesis 5.1.1 focused on the mean vector of T_1 on achievement variables within low reflective intelligence level as contrasted with the mean vector of T_2 on the same variables within the same level. Hypotheses 5.1.1.i (i = 1,...,6) focused on the corresponding components of the two mean vectors. The results of ANOVA's for the hypotheses are reported in Table 38. The multivariate F for the test of equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.01$). Only the univariate isolated F for X2 was significant at $\alpha = 0.01$ ($p < 0.01$).

Hypotheses 5.1.2 and 5.1.2.i Ii = 1,...,6). Hypothesis 5.1.2 focused on the mean vector of T_1 on achievement variable within medium reflective intelligence level as contrasted with the mean vector of T_2

Table 36

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5: $\mu(T_1-T_2/L;\ X1,\ldots,X12)=\mu(T_1-T_2/M;X1,\ldots,X12)=\mu(T_1-T_2/H;\ X1,\ldots,X12)=0$

F-ratio for Multivariate Test of Equality of Mean Vector = 2.14

df = 36 and 624 , p less than 0.00

Table 37

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5.1: $\mu(T_1 - T_2/L; X1,\ldots,X6) =$

$\mu(T_1 - T_2/M; X1,\ldots,X6) = \mu(T_1 - T_2/H; X1,\ldots,X6) = 0$

F-ratio for Multivariate Test of Equality of Mean Vectors = 3.44

df = 18 and 614 , p less than 0.00

Table 38

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5.1.1: $\mu(T_1-T_2/L;X1,\ldots,X6)=0$ and Hypotheses 5.1.1.i: $\mu(T_1-T_2/L,Xi)=0$ $(i=1,\ldots,6)$

F-ratio for Multivariate Test of Equality of Mean Vector = 2.76
df = 6 and 217 , p less than 0.01

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	1.59	0.22	0.64
X2	16.12	6.63	0.01
X3	0.21	0.13	0.71
X4	4.26	2.30	0.13
X5	1.07	0.60	0.44
X6	0.21	0.11	0.74

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

on the same variables within the same level. Hypotheses 5.1.2.i ($i = 7, \ldots 12$) focused on the corresponding components of the two mean vectors. The results of ANOVA's for the hypotheses are reported in Table 39. The multivariate F for the equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). Only the univariate isolated F for X1 was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypotheses 5.1.3 and 5.1.3.i ($i = 1, \ldots, 6$). Hypothesis 5.1.3 focused on the mean vector of T_1 on achievement variables within high reflective intelligence level as contrasted with the mean vector of T_2 on the same variables within the same level. Hypotheses 5.1.3.i ($i = 1, \ldots, 6$) focused on the corresponding components of the two mean vectors. The results of ANOVA's for the hypotheses are reported in Table 40. The multivariate F for equality of mean vectors was significant at $\alpha = 0.01$ ($p < 0.00$). Only the univariate F of X3 was significant at $\alpha = 0.01$ ($p < 0.00$).

Hypothesis 5.2. The hypothesis focused on the mean vectors of T_1 on retention variables within each reflective intelligence level as contrasted with the mean vectors of T_2 on the same variables within the same levels. The results of the ANOVA for the hypothesis is reported in Table 41. The multivariate F for the test of equality of mean vectors was not significant at $\alpha = 0.01$ ($p < 0.11$). Consequently no subordinate hypothesis was tested.

Table 39

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5.1.2: $\mu(T_1-T_2/M; X1,\ldots,X6)=0$ and Hypothesis 5.1.2.i: $\mu(T_1-T_2/M;X1)=0$ $(i=1,\ldots,6)$

F-ratio for Multivariate Test of Equality of Mean Vectors = 4.20
df = 6 and 217 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	59.06	8.12	0.00
X2	14.33	5.89	0.02
X3	6.37	3.93	0.05
X4	6.37	3.44	0.07
X5	1.07	0.60	0.44
X6	0.47	0.24	0.62

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 40

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5.1.3: $\mu(T_1-T_2/H;X1,\ldots,X6) = 0$ and Hypothesis 5.1.3.i: $\mu(T_1-T_2/H;Xi)=0$ $(i=1,\ldots,6)$

F-ratio for Multivariate Test of Equality of Mean Vectors = 3.58
df = 6 and 217 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	30.31	4.17	0.04
X2	8.89	3.66	0.06
X3	13.47	8.31	0.00
X4	0.00	0.00	1.00
X5	2.22	1.25	0.27
X6	0.12	0.06	0.81

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 222

Table 41

Treatment Hypotheses Within Reflective Intelligence

Hypothesis 5.2: $\mu(T_1-T_2/L,X7,\ldots,X12)=\mu(T_1-T_2/M;X7,\ldots,X12)=\mu(T_1-T_2/H;X7,\ldots,X12)=0$

F-ratio for Multivariate Test of Equality of Mean Vectors = 1.44

df = 18 and 614 , p less than 0.11

Summary of Results

Data was collected to examine the five families of hypotheses described in Chapter IV. Interest was focused on each family as an entity as well on hypotheses within each family. Hypotheses in each family were formulated as contrasts of relevant mean vectors on relevant variables. In any family, of a hypothesis was not rejected then all subordinate hypotheses were not tested on the assumption that differences in the subordinate hypotheses were chance differences. Since the total number of hypotheses is large, a summary of the findings is presented in Tables 42 and 43. Table 42 gives a summary of the findings on achievement and retention variables (i.e. Families 1, 2, 5) and Table 43 gives a summary of the findings on difference variables (i.e., Families 3 & 4). Significance at $\alpha = 0.01$ for each contrast as well as marginal significance ($0.01 < p < 0.05$) are included.

Supplementary Analyses

Blocking on Concomitant Variables

It is to be remembered that students were not strictly randomly assigned to treatments. However, samples in the two treatments were balanced on such factors as reflective intelligence, sex and socio-economic status. Moreover, teacher effect was balanced by having each of the five teachers teach two sections one according to T_1 and another according to T_2

A supplementary analysis was made to check whether blocking on reflective intelligence (three levels), sex (two levels) and socio-economic status (two levels: high and low) would result in statistical

Table 42

Summary of Results for Achievement and Retention Variables

Source	Multivariate F on			Univariate F on											
	(X1,...,X12)	(X1,...,X6)	(X7,...,X12)	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
	A & R	A	R	C1,A	C2,A	C3,A	C3,A	C4,A	C4,A	C1,R	C2,R	C3,R	C3,R	C4,R	C4,R
$T_1 - T_2$	**	**	ns	**	**	**	*	ns	ns	*	*	ns	ns	ns	ns
Reflective Intelligence	**	**	**												
L-M		**	**	*	**	**	**	*	*	**	**	**	**	ns	*
ML-H		**	**	**	**	**	**	**	**	**	**	**	**	**	**
Treatment within Reflective Intelligence	**	**	ns												
T_1-T_2/L		**	-	ns	**	ns	ns	ns	ns	-	-	-	-	-	-
T_1-T_2/M		**	-	**	*	*	*	ns	ns	-	-	-	-	-	-
T_1-T_2/H		**	-	*	*	**	ns	ns	ns	-	-	-	-	-	-

** $p < 0.01$
* $0.01 < p < 0.05$
ns $p > 0.05$
\- not tested

C1: criterion 1.
C2: criterion 2
C3: criterion 3
C4: criterion 4

A: achievement
R: retention

Table 43

Summary of Results for Difference Variables

Source	Multivariate Form on	Univariate F on					
	(D1,...,D6	D1	D2	D3	D4	D5	D6
$T_1 - T_2$	ns	–	–	–	–	–	–
Reflective Intelligence	ns	–	–	–	–	–	–
M-L	–	–	–	–	–	–	–
ML-H	–	–	–	–	–	–	–

ns: $p > .05$

–: not tested

$Di = Xi - X(i + 6) \quad (i=1,\ldots,6)$

decisions which differ from those obtained as far as treatment was concerned. The results of ANOVA's for treatment hypotheses in this supplementary analysis appear in Tables 44, 45 and 46. A summary of those results is given in Table 47. By comparing the results in Table 47 with treatment results obtained earlier (first line of Table 41), one can see that all statistical decisions were the same except for one case. The multivariate F on retention variables was non-significant earlier but marginally significant in the supplementary analysis. However, this discrepancy between the two decisions is of no practical consequences since in both cases the multivariate F was not significant at $\alpha = 0.01$ and since the pattern of decisions based on the univariate F was exactly the same in both cases.

Contrasting Subpopulations

Samples in this investigation came from a low socio-economic sub-population (UNRWA/UNESCO) and a high one (I.C.). Implicit in the former statement was the assumption that the group of students formed by pooling the two distinct socio-economic samples can be viewed as a sample from a population formed by pooling the two sub-populations. The possibility arises that different statistical decisions would result, had the two sub-populations been considered separately. A supplementary analysis was carried out to check whether considering each sub-population separately would result in different statistical decisions as far as treatment was concerned. The least square estimates and standard errors of treatment contrast for each sub-population appear

Table 44

Treatment Hypotheses with Blocking on Reflective Intelligence, Socio-economic Status and Sex.

Hypothesis 1: $\mathrm{H}(T_1 - T_2;\ X1, \ldots, X12) = 0$
F-ratio for Multivariate Test of Equality of Mean Vectors = 4.40
df = 12 and 199 , p less than 0.00

Table 45

Treatment Hypotheses with Blocking on Reflective Intelligence, Socio-economic Status and Sex

Hypothesis 1.1: $\mu(T_1-T_2;X1,\ldots,X12)=0$ and Hypothesis 1.1.i(i = 1,...,6): $\mu(T_1-T_2;Xi)=0$

F-ratio for Multivariate Test of Equality of Means Vectors = 8.46
df = 6 and 205 , p less than 0.00

Variable	Hypothesis Mean Square	Univariate F	p less than
X1	69.63	15.19	0.00
X2	38.75	22.38	0.00
X3	14.75	9.52	0.00
X4	7.02	4.75	0.03
X5	0.74	0.41	0.52
X6	0.11	0.08	0.73

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for Error = 210

Table 46

Treatment Hypotheses with Blocking on Reflective Intelligence, Socio-economic Status and Sex

Hypotheis 1.2: $\mu(T_1-T_2;\ X7,\ldots,X12)=0$ and Hypotheses 1.2.i(i=7,...,12):

$$\mu(T_1-T_2;\ Xi)=0$$

F-ratio for Multivariate Test of Equality of Mean Vectors = 2.54
df = 6 and 205 , p less than 0.02

Variable	Hypothesis Mean Square	Univariate F	p less than
X7	31.69	5.90	0.02
X8	12.79	5.99	0.02
X9	3.69	2.55	0.11
X10	1.58	0.87	0.35
X11	0.74	0.49	0.48
X12	0.04	0.02	0.88

Degrees of Freedom for Multivariate Hypothesis = 1
Degrees of Freedom for error = 210

Table 47

Results of Supplementary Analysis with Blocking on Concomitant Variables

Source	Multivariate F			Univariate F											
	(X1,...,X12)	(X1,...,X6)	(X7,...,X12)	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
	A & R	A	R	C1,A	C2,A	C3,A	C3,A	C4,A	C4,A	C1,R	C2,R	C3,R	C3,R	C4,R	C4,R
$T_1 - T_2$	**	**	*	**	**	**	*	ns	ns	*	*	ns	ns	ns	ns

** $p < 0.01$
* $0.01 < p < 0.05$
ns $p > 0.05$

C1:criterion one
C2:criterion two
C3:criterion three
C4:criterion four

A:achievement
R:retention

GPO 804-930-10

in Table 48. The results of ANOVA's for treatment hypotheses in this supplementary analysis appear in Tables 49, 50 and 51. A summary of those results is given in Table 52. By comparing the results in Table 52 with treatment results for the pooled population (first row of Table 41), one can see that all statistical decisions were the same except for three cases. First, the multivariate F on retention variables was non-significant in two cases (UNRWA/UNESCO and pooled population) and marginally significant in one case (I.C); however, this discrepancy is of no practical consequences since in both cases the multivariate F was not significant at $\alpha = 0.01$ and since the pattern of decisions based on the univariate F's (X7 & X8) is the same in the three cases. Second, the univariate F for X1 was only marginally significant ($P \leq 0.02$) in one case (UNRWA/UNESCO) but significant ($\alpha = 0.01$) in two cases (I.C & pooled population). Third, the univariate F's on measures of C3 (X3 and X4) was not consistent. The later remark should temper any interpretation of treatment differences in C3.

Table 48

Least Square Estimates and Standard Errors of Treatment
Contrast for the Two Sub-populations

Contrast		Variable											
		X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
T_1-T_2	UNRWA/ UNESCO	-1.08 (0.47)	-1.18 (0.27)	0.35 (0.20)	0.54 (0.23)	-0.08 (0.17)	-0.02 (0.13)	-0.93 (0.49)	-0.58 (0.30)	0.01 (0.20	0.07 (0.25)	-0.20 (0.17)	-0.11 (0.17)
	I.C	-1.75 (0.44)	-0.68 (0.27)	0.61 (0.31)	-0.08 (0.29)	0.25 (0.36)	-0.33 (0.35)	-0.98 (0.49)	-0.57 (0.30)	0.54 (0.29)	0.14 (0.35)	0.42 (0.33)	-0.00 (0.37)

Table 49

Treatment Hypotheses with Sub-populations Contrasted

Hypothesis 1: $\mu\ (T_1 - T_2;\ X1, \ldots, X12) = 0$

UNRWA/UNESCO:

F-ratio for Multivariate Test of Equality of Mean Vectors = 3.23

df = 12 and 114 , p less than 0.00

I.C:

F-ratio for Multivariate Test of Equality of Mean Vectors = 3.10

df = 12 and 88 , p less than 0.00

Table 50

Treatment Hypotheses with Sub-populations Contrasted

Hypothesis 1.1: $\mu(T_1-T_2; X1,\ldots,X6)$ and Hypotheseses 1.2.i: $(i=1,\ldots,6)$

$\mu(T_1-T_2;Xi) = 0$

F-ratio for Multivariate Test of Equality of Mean Vectors = 6.31(5.19)

df = 6(6) and 120(94) , p less than 0.00 (0.00)

Variable	Hypothesis Mean Squares	Univariate F	p less than
X1	37.03(76.97)	5.42 (16.15)	0.02 (0.00)
X2	44.17(11.74)	19.17 (6.58)	0.00 (0.01)
X3	3.97(9.47)	3.27 (4.00)	0.07 (0.05)
X4	9.19(0.15)	5.56 (0.07)	0.02 (0.79)
X5	0.21(1.55)	0.22 (0.49)	0.64 (0.49)
X6	0.02(2.69)	0.03 (0.86)	0.86 (0.36)

Degrees of Freedom for Multivariate Hypothesis = 1(1)
Degrees of Freedom for Error = 125(99)

Note: Numbers in parentheses are I.C sub-population values; numbers not in parentheses are UNRWA/UNESCO sub-population values.

Table 51

Treatment Hypotheses with Sub-populations Contrasted

Hypothesis 1.2: $\mu(T_1 - T_2;\ X7, \ldots mX12) = 0$ and Hypotheses 1.2.i: (i = 7,...,12)

$$\mu(T_1 - T_2;\ Xi) = 0$$

F-ratio for Multivariate Test of Equality of Mean Vectors = 1.29

df = 6(6) and 120(94) , p less than 0.27(0.03)

Variable	Hypothesis Mean Squares	Univariate F	p less than
X7	27.65 (24.03)	3.57 (4.02)	0.05 (0.05)
X8	10.60 (8.17)	3.83 (3.63)	0.05 (0.05)
X9	0.01 (7.26)	0.01 (3.34)	0.95 (0.07)
X10	0.14 (0.51)	0.07 (0.16)	0.79 (0.69)
X11	1.21 (4.40)	1.37 (1.58)	0.24 (0.21)
X12	0.40 (0.00)	0.44 (0.00)	0.51 (0.99)

Degrees of Freedom for Multivariate Hypothesis = 1(1)
Degrees of Freedom for error = 125 (99).
Note: Number in parentheses are I.C sub-population values; numbers not in parentheses are UNRWA/UNESCO sub-population values.

Table 52

Summary of Results with the Two Sub-populations Contrasted

		Multivariate F			Univariate F											
		(X1,...,X12)	(X1,...,X6	(X7,...,X12)	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
		A & R	A	R	C1,A	C2,A	C3,A	C3,A	C4,A	C4,A	C1,R	C2,R	C3,R	C3,R	C4,R	C4,R
T_1-T_2	UNRWA/ UNESCO	**	**	ns	*	**	ns	*	ns	ns	*	*	ns	ns	ns	ns
	I.C	**	**	*	**	**	*	ns	ns	ns	*	*	ns	ns	ns	ns

** $p \leq 0.01$
* $0.01 < p \leq 0.05$
ns $p > 0.05$

C1: criterion one
C2: criterion two
C3: criterion three
C4: criterion four

A: achievement
R: retention

Chapter VI

CONCLUSIONS AND DISCUSSION

In this chapter, conclusions pertaining to the five questions with which this study dealt are drawn from results of data analysis. In addition the conclusions are discussed in terms of expectations motivating this study and in terms of related research. A brief overview of the study is given first.

Overview

This study had two general aims:

1. To compare, on predetermined criteria, the effectiveness of two teaching methods having two distinct levels of emphasis on mathematical structure in organizing and presenting the same mathematical content.

2. To identify the effect of a cognitive ability known as reflective intelligence on four cognitive levels of learning mathematical structure.

Integral powers of 2 and 3, as models of infinite cyclic group, were chosen as a suitable mathematical content for 8th graders in Lebanon where the investigation was carried out. Two treatments T_1 and T_2 were constructed in such a way that T_2 tended to emphasize explicitly the structural properties of the models in developing operations and algorithms and in manipulating isomorphisms whereas T_1 attempted a direct approach with no explicit emphasis on structural properties.

T_1 and T_2 were piloted and then administered to 5 intact 8th grade classes each (a total of 114 students for each). Each of the five teachers, following the specifications of T_1 and T_2, taught two sections, one according to each. T_1 lasted for six 40-minutes lessons and T_2 for seven 40-minutes lessons.

The sample was divided, according to the sum score of two parts of Skemp test of reflective intelligence, into three categories: (1) low reflective intelligence level (L); (2) medium reflective intelligence level (M) and (3) high reflective intelligence level (H).

Outcomes were evaluated against four predetermined criteria:

C1: Ability to solve mathematical sentences of the form $ab = x$, $ax = b$ and $xa = b$ in the taught models.

C2: Ability to solve the same type of mathematical sentences in an isomorphic model.

C3: Ability to select and solve mathematical sentences which "model" decisions in a physical model on which an isomorphic structure is imposed.

C4: Ability to select and solve mathematical sentences which "model" decisions in a generalized model of the taught model, *i.e.*, contains an isomorphic copy of it.

Measurements of the four criteria were taken at two occasions: (1) immediately following the conclusions of the treatments (achievement) and (2) two weeks later (retention). X1, X2, {X3, X4}, {X5, X6} were achievement measures for C1, C2, C3, and C4 respectively. X7, X8, {X9, X10}, {X11, X12} were retention measures of C1, C2, C3, and C4 respectively.

Five questions were generated from the general aims of the study:

1. Are there treatment differences? For which criteria? For which measures?
2. Are there reflective intelligence differences? For which criteria? For which measures?
3. Are there treatment differences on differences variables? (a difference variable Di is defined as Xi - X (i+1); i = 1,...,6).
4. Are there reflective intelligence differences on difference variables? For which variables?
5. Within reflective intelligence, are there treatment differences? For which criteria? For which measures?

The five questions generated five families of hypotheses which were respectively: (1) treatment hypotheses [family 1] ; (2) reflective intelligence hypotheses [family 2] ; (3) treatment hypotheses on difference variables [family 3] ; (4) reflective hypotheses on difference variables [family 4] and (5) treatment hypotheses within reflective intelligence [family 5]. The particular hypotheses in each family were given in Chapter IV.

Conclusions

The conclusions to be drawn from this study are believed to be valid for the samples investigated. Generalization to samples from other populations is subject to limitations which are rather restrictive.

Generalization to other samples should be subject to limitations pertaining to population, sampling procedure and base-line data:

(1) The population of 8th graders in this study, beside having distinct cultural characteristics was a hypothetical one constructed by pooling two almost extreme sub-populations in terms of both socio-economic class and academic aptitude, i.e. UNRWA/UNESCO and I.C sub-populations. The questions arise as to the usefulness of such a hypothetical and may be non-representative population and also as to the possibility of having confounded treatment effects by pooling two extreme sub-populations. The hypothetical nature of the population is to be taken as one limitation of this study. The results of the supplementary analysis (contrasting sub-populations) suggested that the treatment results hold for each and both samples from the two sub-populations - an observation which strengthens the possibility of generalizing (subject to other limitations) to either sub-population. Again, this last observation greatly reduces the possibility of confounding, since the patterns of treatment effects were comparable for both sub-populations.

(2) A second limitation of this study is the fact that sampling was not a random sampling, for which, admittingly, there is no substitute. Given what was available, some measures were taken to make up partly for this shortcoming. For one thing, the samples were balanced on teacher effect. For another, the possibility that treatment effect was due to such concomitant variables as sex, socio-economic class and reflective intelligence was ruled out as suggested by the results of the supplementary analysis (blocking on concomitant variables).

(3) A third limitation of this study is the lack of base-line data. Although there was no reason to believe that students were familiar with the mathematical content of the experiment, the availability of base-line data such as levels of mathematical knowledge and maturity would have possibly allowed stronger inferences.

Conclusions are drawn from data analysis presented in Chapter V and, in particular, from the summary presented in Tables 41 and 42. The direction of a particular difference is judged from the direction of the observed values of the contrasts in Tables 16 and 35. Conclusions are presented in five sections corresponding to the five families of hypotheses.

Treatment Conclusions

Subject to the limitations of this study, there is evidence for each of the following statements :

1. Across reflective intelligence, there were overall treatment differences on achievement and retention measures of the four criteria.

2. Across reflective intelligence, there were overall treatment differences on achievement measures. The pattern of differences was as follows:

   a. The difference in estimated means of T_1 and T_2 favored significantly ($\alpha = 0.01$) T_2 on measures of criteria C1 and C2.

   b. The difference in estimated means of T_1 and T_2 favored significantly ($\alpha = 0.01$) T_1 on one measure of C3, i.e. ability to select mathematical sentences which model decisions in a physical model on which an isomorphic (to the taught model) structure is imposed. For the second measure of C3, i.e. solving such sentences, the difference in estimated means of T_1 and T_2, although marginally significant ($0.01 < p < 0.05$), favored T_1.

   c. T_1 and T_2 were comparable on measures of criterion four(C4). Moreover, T_1 and T_2 were both characterized by low achievement on measures of C4.

3. Across reflective intelligence, T_1 and T_2 were comparable on retention measures of the four criteria.

Reflective Intelligence Conclusions

Subject to: (1) the limitations of this study; (2) the way reflective intelligence was measured by the sum score of two parts of Skemp test: Operation Formation and Reflective Activity with Operations; and (3) the way in which the three levels of reflective intelligence were identified, there is evidence for the following statements:

1. Across treatment, there were overall reflective intelligence differences on achievement and retention measures of the four criteria.

2. Across treatments, there were overall reflective intelligence differences on achievement measures of the four criteria. The pattern of differences was as follows:

   a. The difference in estimated means between low and medium reflective intelligence levels favored significantly ($\alpha = 0.01$) the medium level on measures of criterion two (C2) and criterion 3 (C3). The same difference, although only marginally significant, favored the medium level on measures of criterion one (C1) and criterion four (C4).

   b. The difference in estimated means between the average of low and medium reflective intelligence levels and high reflective intelligence level favored signifcantly (α = 0.01) the high level on all measures of C1, C2, C3, and C4.

3. Across treatments there were overall reflective intelligence differences on retention measures of the four criteria. The pattern of differences was as follows:

a. The difference in estimated means between low and medium reflective intelligence levels favored significantly (α = 0.01) the medium level on measures C1, C2 and C3. The same difference favored the medium level on C4, although the difference nonsignificant on one measure and marginally significant on the other.

b. The difference in estimated means between the average of low and medium and that of high reflective intelligence levels favored significantly (α = 0.01) the high level on all measures of C1, C2, C3, and C4.

Treatment Conclusions within Reflective Intelligence

Subject to:(1) the limitations of this study; (2) the way reflective intelligence was measured by the sum score of Operation Formation and Reflective Activity with Operations; and (3) the way by which the three levels of reflective intelligence were identified, there is evidence for the following statements:

1. Within reflective intelligence, there were overall treatment differences on achievement and retention measures of the four criteria.

2. Within reflective intelligence, there were overall treatment differences on achievement measures of the four criteria. The pattern of differences was as follows:

a. Within low reflective intelligence level, the difference in estimated means between T_1 and T_2 favored significantly ($\alpha = 0.01$) T_2 on measure of criterion two (C2). Otherwise T_1 and T_2 were comparable.

b. Within medium reflective intelligence level, the difference in estimated means between T_1 and T_2 favored significantly ($\alpha = 0.01$) T_2 on measure of criterion one (C1) and T_1 on measures of criterion two (C2) and criterion three (C3) but marginally significant ($0.05<p<0.01$).

c. Within high reflective intelligence level, the difference in estimated means between T_1 and T_2 favored significantly ($\alpha = 0.01$) T_1 on one measure of criterion three (C3), i.e. ability to select sentences which model decisions in a physical model on which an isomorphic structure is imposed. The same difference, although marginally significant ($0.01 < \rho < 0.05$), favored T_2 on measures of criteria C1 and C2. Otherwise T_1 and T_2 were comparable.

3. Within reflective intelligence levels, T_1 and T_2 were comparable on retention measures of the four criteria.

Treatment Conclusions on Difference Variables

Subject to:(1) the limitation of this study; (2) the way reflective intelligence was measured by two subtests of Skemp test; and (3) the definition of a difference variable as: $Di = Xi - X(i + 6)$, $i = 1,...6$; there is evidence for the following statement:

Across reflective intelligence levels, T_1 and T_2 were comparable on difference variables.

Reflective Intelligence Conclusion on Difference Variables

Subject to (1) the limitation of this study; (2) the way reflective intelligence was measured by two subtests of Skemp test; and (3) the definition of a difference variable as $Di = Xi - x(i+1)$, $i = 1,...,6$; there is evidence for the following statement:

Across treatments, the three reflective intelligence levels were comparable on difference variables.

Discussion

Treatments

One interesting result of this study is the relative superiority of T_2 as contrasted with T_1 in producing better performance on criteria C1 and C2, i.e. ability to solve mathematical sentences in the taught models and in isomorphic models. The discriminant analysis showed that measures of C1 and C2 discriminate rather highly and in the same direction between T_1 and T_2. A close examination of C1 and C2 reveal that (1) both deal with models of the same mathematical theory (i.e., same language and same postulate system), (2) models involved in C1 and C2 are isomorphic but different interpretations (i.e., different sets) of the same mathematical theory. On purely logical level, it does not seem surprising that the emphasis in T_2 on the structure of the models reinforced the ability to operate within the taught models and the ability to "translate" the operations to untaught different but isomorphic interpretations. However, one should not be tempted to read in this statement more than it conveys. For example, this study does not (and was not designed to) provide evidence as to whether the relatively better performance of T_2 (as contrasted with T_1) on C1 and C2 implies more awareness of the underlying structural properties of the models.

C3 and C4, however, involve a different kind of "modelling" and cognitively higher order processes. C3 and C4, unlike C1 and

and C2, involve "modelling" .. w. ich the interpretation, i.e. the set and the function which relates the mathematical theory to the set, is not readily availa' but rather has to be conceived or actually constructed. C3 and C4 involve "modelling" from a physical model on which an isomorphic structure (to that in C1) or a generalized structure is imposed. Cognitively, C3 and C4 require decision-making in the sense that they call for selecting an algorithm rather than performing a well-practised algorithm (as in C1) or translating the algorithm into an analogous symbolic system (as in C2). C3 requires the process of application and C4 the process of analysis (a fuller discussion is given in Chapter III).

The results of this study suggest that the better performance associated with T_2 relative to T_1 on criteria C1 and C2 failed to carry over to criteria C3 and C4. This failure raises at least two possibilities: (1) the scales for C3 and C4 were not sufficiently sensitive for differences between T_1 and T_2; or (2) T_2 is not actually better than T_1 on measures of C3 and C4. If one accepts the argument that there is no reason why students should learn what they are not specifically taught, one tends to find the second possibility more plausible since neither T_1 nor T_2 included the objective of a "modelling" from a physical situation as an intended learning. Results of many studies in different contexts indicate that sensitivity towards the use or application of a structural property (or rule) is not a necessary consequence of learning it.

This has been supported by conclusions of studies mentioned in Chapter II (Gray, 1965; Scandura, 1967; Weaver, 1973). In this light, this study does not seem to provide evidence to the claim that emphasizing structural properties in teaching some models is necessarily conducive to better performance on problems involving "modelling" from physical models, even though an isomorphic (to the taught models) is imposed on them.

One rather unexpected result was the better performance in the achievement phase associated with T_1 as contrasted with T_2 on measure X3 of C3, i.e. ability to select mathematical sentences which model decisions in a physical model on which an isomorphic structure is imposed. Since this result was not consistent for the two sub-populations and since it was not stable on the corresponding retention measure (in the retention phase the treatment contrast on X3 was nonsignificant) one is inclined to interpret it as a chance result. In comparison, differences which favored T_2 in the achievement phase tended to be rather stable in the retention phase (the treatment contrast on each of the measures of C1 and C2 in the retention phase was marginally significant ($0.01 < p < 0.05$). Whether the emphasis on mathematical structure in T_2 was conducive to this stability of achievement should be regarded as a tenuous conjecture.

Reflective Intelligence

The results of this study, as far as reflective intelligence is concerned, provide some evidence to the validity of the construct of reflective intelligence as defined by Skemp (1962). Reflective intelligence involves the functioning of a second order system on concepts and operations of a sensori-motor system. The system of concepts and operations involved in the learning of powers of 2 and 3 with the multiplication and division operations is a second-order system since it builds on a first-order system, i.e., the system of integers and the operations of addition and subtraction on them. Consequently, it is expected that a better performance on concepts and operations of a second-order system should be associated with higher reflective intelligence. This expectation was substantiated rather well by the results of this study: Across treatments, differences between reflective intelligence levels on all measures of the four criteria and in both achievement and retention phases, favored (significantly in most cases and marginally significant in the rest) the higher level. The fact that differences between medium and low levels were less pronounced than differences between high and either medium or low levels should not be taken to indicate differences in the rate of learning between the three levels. A rather more reasonable interpretation lies in the way in which the three levels were identified: the sample was divided into three categories according to the sum score of two subtests of Skemp test (Appendix E). In the process, the three levels did not come out to be equidistant from each other as can be judged from the range of scores in each level: low (4 - 20), medium (21 - 27) and high (28 - 67). Obviously, the medium level is nearer to the low level than the high to the medium level.

The results concerning treatment effect within reflective intelligence levels suggest two patterns: (1) significant differences between treatments were mainly confined to the achievement phase and to the lower criteria C1, C2 and partially C3, and (2) differences between treatments were more frequent within the middle and high reflective intelligence levels. Whether those two patterns are valid in general cannot be concluded from the results of this study since a good number of the relevant treatment differences within reflective intelligence levels were marginally significant ($0.01 < p < 0.05$).

Achievement Versus Retention

A mean score on a difference variable might be viewed as indication of the rate of forgetting (time was fixed, *i.e.*, two weeks) associated with the criterion under consideration. The non-existence of significant differences between treatments suggests that they are comparable, on the four criteria, in their rates of forgetting the initial learning. The same interpretation may be given to the non-existence of significant differences between reflective intelligence levels.

NLSMA's Model

The results reported in this study support the hypothesis that mathematics achievement is a multivariate phenomenon as strongly suggested by the achievement model developed by National Longitudinal Study of Mathematical Abilities (NLSMA) (Romberg and Wilson, 1969). In particular, students who did well on one criterion did not necessarily do well on another; for example, students in T_2 did better than those in T_1 on C1 (classified as computation) and C2 (classified as comprehension) but not on C3 (classified as application) and C4 (classified as analysis). Moreover, discriminant analysis showed that a contrast between the measures of the two lower criteria (C1 and C2) and of the two higher criteria (C3 and C4) provided maximum discrimination between treatments. This last remark is in line with the pattern suggested by NLSMA (Mcleod and Kilpatrick, 1969):

> "For one thing, it is clear that, although we can easily divide the goals of junior high mathematics instruction into 'computational facility' and 'higher processes,' these higher processes have yet to be clearly delineated." (p. 82).

Further Research

This study being an exploratory study, needs to be replicated with variations. In particular, replications in different cultural contexts and in different grade levels are possible variations. Another variation would be to vary the way emphasis on structure is conceived and implemented. This study attempted a global approach in the sense that observed experimental effects were the net result of emphasizing mathematical structural properties in at least three contexts: (1) developing operations; (2) developing algorithms; and (3) manipulation of isomorphisms. A promising line of approach would be to design studies in order to isolate the effect of emphasizing mathematical structure in each context or in a sequence of two or more of the three contexts.

With respect to reflective intelligence, the identification of more functional relationships between reflective intelligence and problem solving in mathematics is needed. Results so far suggest that the existence of reflective intelligence is a necessary condition for learning second-order systems in mathematics. However, little is known how different teaching strategies affect the development of reflective intelligence and how reflective intelligence is related to different strategies of problem solving in mathematics.

At last one final remark to put this investigation in proper perspective. There is no doubt that a mathematical theory is a source of economy in mathematics in the sense that it applies to infinitely many models. However, in school education, the situation is different because developmental and social dimensions together with the mathematics dimension come into play. This study attempted to explore the extent to which "economy" feature is attained if teaching emphasized structural properties of the taught models in developing operations, algorithms and isomorphisms. The results suggest that the economy involved is of very restricted scope. Identification of the best strategies which produce maximum economy in school mathematics is still a challenging question.

APPENDICES

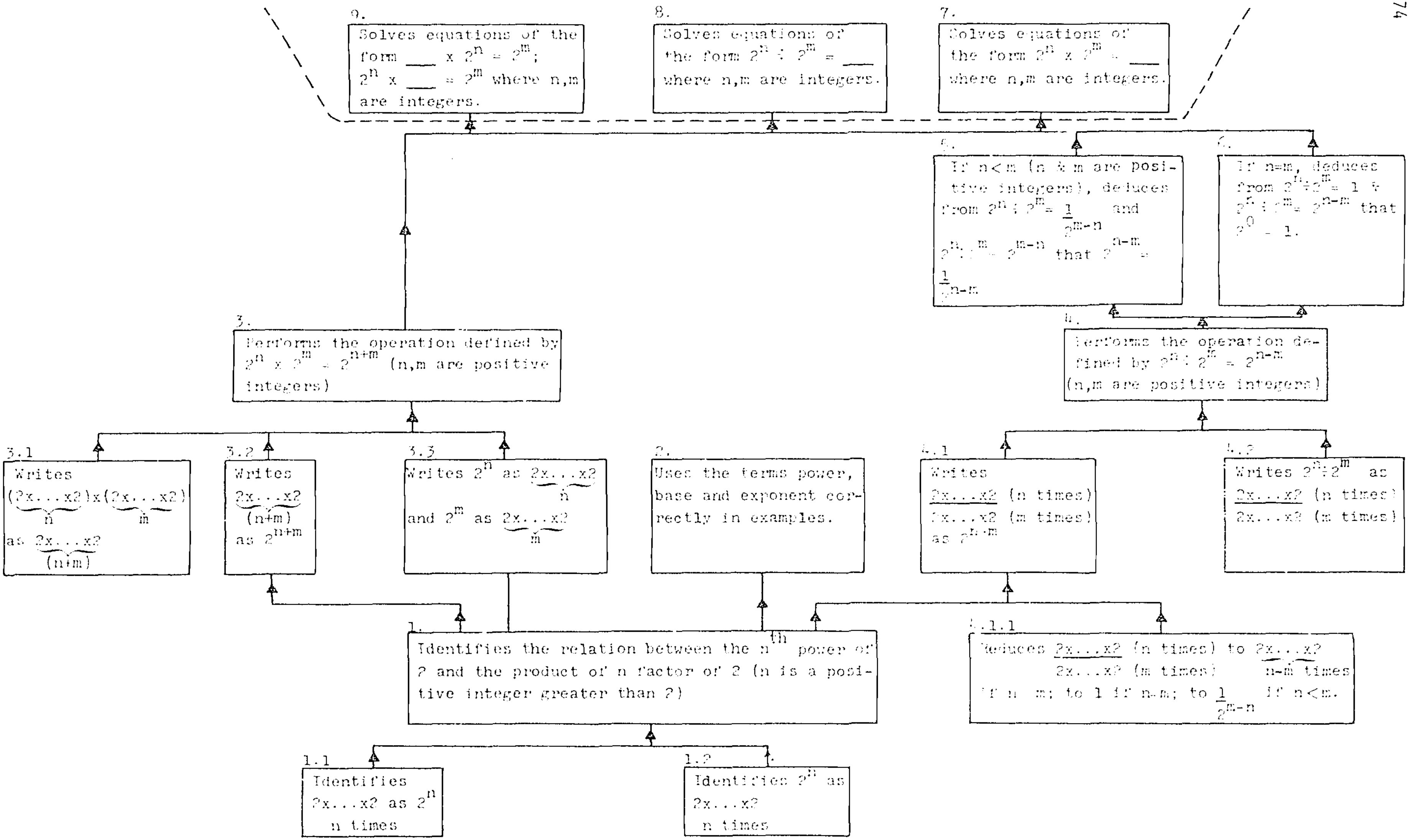
Appendices B, C, D, and E have been omitted from this publication, but are available on microfilm from Memorial Library, University of Wisconsin, Madison, Wisconsin.

Appendix A

SCHEMATIC DIAGRAMS FOR BEHAVIOURS OF

T_1 and T_2

Schematic Diagram for Behaviours of T_1



REFERENCES

Begle, E. & Wilson, J. Evaluation of programs. In Mathematics Education. National Society for the Study of Education 69th Yearbook. Chicago: University of Chicago Press, 1970.

Bock, R. D. Multivariate analysis of variance. In D.K.Whitla (Ed.), In Handbook of Measurement and Assessment in Behavior Sciences. Reading: Addison-Wesley, 1968.

Branca, A. N. & Kilpatrick, J. The consistency of strategies in the learning of mathematical structures. Journal for Research in Mathematics Education, 1972, 3, 132-140.

Bruner, J. S. The Process of Education, Massachusetts: Harvard University Press, 1963.

Carry, L. R. Patterns of Mathematics Achievement in Grades 7 and 8: X-Population. NLSMA Report No. 11, Stanford: Stanford University, 1970.

Carry, L. R. and Weaver, J. F. Patterns of Mathematics Achievement in Grades 4, 5, 6: X-Population. NLSMA Report No. 10. Stanford: Stanford University, 1969.

Cassell, R. & Jerman, M. Preliminary evaluation of SMSG instruction in arithmetic and algebra for 7th, 8th and 9th grade pupils. California Journal of Educational Research. 1963, 14, 202-207.

Coxford, A. F. The effects of two instructional approaches on the learning of addition and subtraction concepts in grade one. Dissertation Abstracts, 1966, 26, 6543-44.

Dienes, Z. P. & Jeeves, M. A. Thinking in Structures. London: Hutchinson Educational Ltd., 1965.

Dienes, Z. P. & Jeeves, M. A. The Effects of Structural Relations on Transfer. London: Hutchinson Educational Ltd., 1970.

Eves, H. & Newson, C. An Introduction to the Foundations and Fundamental Concepts of Mathematics. New York: Holt, Rinehart and Winston, 1965.

Games, P. A. Multiple comparisons of means. American Educational Research Journal, 1971, 8, 531-565.

Gray, R. F. An experiment in the teaching of introductory multiplication. Arithmetic Teacher. 1965, 12, 199-203.

Harrison, D. B. Reflective intelligence and mathematics learning. Unpublished Ph.D. dissertation, University of Alberta, 1967.

Lebanese Ministry of Education, School Curricula for Middle School. Beirut, 1970.

Mastain, R. K. A comparison of the effect of two mathematics curricula on the mathematical understanding of 4th grade pupils. Dissertation Abstracts. 1964, 25, 3439.

McLeod, G. K. and Kilpatrick, J. Patterns of Mathematics Achievement in Grades 7 & 8: Y-Population. NLSMA Report No. 12. Stanford: Stanford University, 1969.

Osborne, A. R., The effects of two instructional approaches on the understanding of subtraction by grade 2 students. Dissertation Abstracts, 1967, 28, 158A.

Peel, E. A. Psychological and educational research bearing on mathematics teaching. In W. Servais and T. Varga (Eds.), Teaching School Mathematics. Penguin Books - Unesco, 1971.

Piaget, J. The Psychology of intelligence (English Edition). London: Routledge and Kegan Paul, 1950 (5th edition).

Romberg, T. A. and Wilson, J. W. The Development of Tests. NLSMA Report No. 7. Stanford: Stanford University, 1969.

Romberg, T. A. Curriculum, development and research. In Mathematics in Secondary School. National Council of Teachers of Mathematics 33rd Yearbook. Washington, D. C., 1970.

Rosenbloom, P. SMSG evaluation, grades 7-12. Reports Newletter No. 10. New Haven, Conn. SMSG, 1961, 12-26.

Roy, S. N., and R. Granadeskian. Further contribution to multivariate confidence bounds. Biometrika, 1957, 44, 399-410.

Ruddell, A. K. The results of a modern mathematics program. Arithmetic Teacher. 1962, 9, 330-335.

Scandura, J. M., Woodward, E., & Lee, F. Rule generality and consistency in mathematics learning. American Educational Research Journal, 1967, 4, 303-320.

SMSG Newsletter No. 1. School Mathematics Study Group, 1958.

SMSG. Mathematics For Junior High School. Vol. I, Part II. Stanford, 1965.

SMSG. Intermediate Mathematics part I. Stanford: 1965.

Scott, Lloyd. Trends in Elementary School Mathematics. Chicago: Rand McNally, 1966.

Skemp, R. R. Reflective intelligence and mathematics. British Journal of Educational Psychology. 1961, 31, 45-54.

UNESCO. Cairo Seminar Report. UNESCO Mathematics Project for Arab Countries. Paris: 1969 (limited distribution).

Weaver, J. F., Student achievement in SMSG classes. Reports Newsletter No. 15, Stanford, SMSG: 1963: 3-8.

Weaver, J. F. Some factors associated with pupils' performance on examples involving selected variations of the distributive idea. A paper presented at 1973 annual meeting of AERA, New Orleans, February 27.

Wilson, J. W. Secondary School Mathematics. In Bloom et al (Eds.), Handbook of Summative and Formative Evaluation. McGraw-Hill, 1970.

Wrigley, J. The factorial nature of ability in elementary mathematics. British Journal of Educational Psychology. 1958, 28, 61-78.

GPO 804-930-12

National Evaluation Committee

Helen Bain
Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
University of Colorado

Sue Buel
Dissemination and Installation Services
Northwest Regional Educational Laboratory

Francis S. Chase
Professor Emeritus
University of Chicago

George E. Dickson
College of Education
University of Toledo

Chester W. Harris
Graduate School of Education
University of California

Hugh J. Scott
Consultant
National Evaluation Committee

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Joanna Williams
Professor of Psychology and Education
Columbia University

Executive Committee

William R. Bush
Director, Program Planning and Management
Deputy Director, R & D Center

M. Vere DeVault
Professor
School of Education

Herbert J. Klausmeier
Principal Investogator
R & D Center

Joel R. Levin
Principal Investigator
R & D Center

Donald N. McIsaac
Associate Dean, School of Education
University of Wisconsin

Richard A. Rossmiller, Committee Chairman
Director
R & D Center

Len VanEss
Associate Vice Chancellor
University of Wisconsin–Madison

Dan Woolpert
Director, Management Systems
R & D Center

Faculty of Principal Investigators

Vernon L. Allen
Professor
Psychology

B. Dean Bowles
Associate Professor
Educational Administration

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
V. A. C. Henmon Professor
Educational Psychology

Gisela Labouvie
Assistant Professor
Educational Psychology

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

James Lipham
Professor
Educational Administration

Wayne Otto
Professor
Curriculum and Instruction

Robert Petzold
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Dennis W. Spuck
Assistant Professor
Educational Administration

Richard L. Venezky
Associate Professor
Computer Science

Larry M. Wilder
Assistant Professor
Communication Arts