DOCUMENT RESUME

ED 035 593 | 24 | SP 003 462

AUTHOR	Sowder, Larry
TITLE	Discovery Learning: A Status Study, Grades 4-7, and an Examination of the Influence of Verbalizing Mode on Retention. Report from the Project on Analysis of Mathematics Instruction.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
REPORT NO	TR-99
BUREAU NO	BR 5-0216
PUB DATE	Jul 69
CONTRACT	OEC-5-10-154
NOTE	154p.
EDRS PRICE	EDRS Price MF-$0.75 HC-$7.80
DESCRIPTORS	*Discovery Learning, *Elementary School Mathematics, Mathematics Instruction, *Retention, *Stimulus Generalization, Verbal Learning

ABSTRACT

A study had two aims: to explore the ability of pupils of grades 4-7 to give operational evidence of generalizing in selected numerical situations, and to study the effects of differing manners of verbalizing a generalization on the retention of the ability to use the generalizations. Pupils (18) from each of grades 4-7 in a public school were randomly chosen and given an individually administered discovery test consisting of the stimulus portions of instances of generalizations. For the exploratory part of the study, the number of instances required before the pupil gave correct responses, as well as the number of generalizations apparently formed, were recorded and analyzed by a grade-by-IQ level-by-sex analysis of variance and with respect to a linear model with independent variables age, IQ, arithmetic achievement, and mathematical interests. Performance on a follow-up test (one week later) based on instances of the generalizations on the discovery test was to provide information for the influence-of-verbalizing study. Indications are that most pupils can form generalizations of the type encountered, although pupils of lower IQ require more instances. With the number of instances needed as a criterion, the optimal grade level at which to offer generalizing tasks appear to be grade 6 or after. The plateau at grade 6 supports Piagetian thought, although it may be due to a plateau on computational proficiency. (Instruments and bibliography are included.) (Author/JS)

BR-5-0216 · TR 99
PA-24
OE/BR

Technical Report No. 99

DISCOVERY LEARNING: A STATUS STUDY, GRADES 4-7, AND AN EXAMINATION OF THE INFLUENCE OF VERBALIZING MODE ON RETENTION

Report from the Project on
Analysis of Mathematics Instruction

By Larry Sowder

Thomas A. Romberg, Assistant Professor of Curriculum and Instruction
Chairman of the Examining Committee, and Principal Investigator

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

July 1969

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Henry Chauncey
President
Educational Testing Service

Elizabeth Koontz
President
National Education Association

Patrick Suppes
Professor
Department of Mathematics
Stanford University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Roderick McPhee
President
Punahou School, Honolulu

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

Francis S. Chase
Professor
Department of Education
University of Chicago

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

G. Wesley Sowards
Director, Elementary Education
Florida State University

UNIVERSITY POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

John Guy Fowlkes
Director
Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

M. Crawford Young
Associate Dean
The Graduate School

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Donald J. McCarty
Dean
School of Education

***James W. Cleary**
Vice Chancellor for Academic Affairs

H. Clifton Hutchins
Chairman
Department of Curriculum and Instruction

Ira Sharkansky
Associate Professor of Political Science

Leon D. Epstein
Dean
College of Letters and Science

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Max R. Goodson
Professor of Educational Policy Studies

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Ronald R. Allen
Associate Professor of Speech and of Curriculum and Instruction

Gary A. Davis
Associate Professor of Educational Psychology

Max R. Goodson
Professor of Educational Policy Studies

Richard G. Morrow
Assistant Professor of Educational Administration

Vernon L. Allen
Associate Professor of Psychology
(On leave 1968-69)

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Warren O. Hagstrom
Professor of Sociology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Nathan S. Blount
Associate Professor of English and of Curriculum and Instruction

Frank H. Farley
Assistant Professor of Educational Psychology

John G. Harvey
Associate Professor of Mathematics and Curriculum and Instruction

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Robert C. Calfee
Associate Professor of Psychology

John Guy Fowlkes (Advisor)
Professor of Educational Administration
Director of the Wisconsin Improvement Program

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Thomas A. Romberg
Assistant Professor of Mathematics and of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

Burton W. Kreitlow
Professor of Educational Policy Studies and of Agricultural and Extension Education

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

MANAGEMENT COUNCIL

***Herbert J. Klausmeier**
Director, R & D Center
Acting Director, Program 1

Thomas A. Romberg
Director
Programs 2 and 3

James E. Walter
Director
Dissemination Section

Dan G. Woolpert
Director
Operations and Business

Mary R. Quilling
Director
Technical Section

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from Phase 2 of the Project on Prototypic Instructional Systems in Elementary Mathematics in Program 2. General objectives of the Program are to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to develop assessment procedures for those concepts and skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to the Program objectives, the Mathematics Project, Phase 1, is developing and testing a televised course in arithmetic for Grades 1-6 which provides not only a complete program of instruction for the pupils but also inservice training for teachers. Phase 2 has a long-term goal of providing an individually guided instructional program in elementary mathematics. Preliminary activities include identifying instructional objectives, student activities, teacher activities materials, and assessment procedures for integration into a total mathematics curriculum. The third phase focuses on the development of a computer system for managing individually guided instruction in mathematics and on a later extension of the system's applicability.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

ACKNOWLEDGMENTS

Acknowledgment pages in dissertations take on added meaning once one has written a dissertation. Although I had formerly viewed such pages as token apple-polishing, it is now clear that the guidance and assistance of others necessitate public acknowledgment and expressions of appreciation.

Especial appreciation must be extended to my major professor, Dr. Thomas A. Romberg, who despite his heavy schedule could always find time to discuss problems. His grasp of what needed to be done and the calming effect of his knowledgeable, common sense approach were invaluable.

Appreciation must also be extended to the other members of the reading committee, Drs. John G. Harvey and David C. Johnson, for their comments and their efforts in expediting the processing of the manuscript.

Recognition must be made of Dr. Henry Van Engen for his support and counsel in the early part of my graduate study, and Dr. J. Fred Weaver for reading part of the manuscript.

Irvin King and Jack Shepler, who interrupted their doctoral studies to assist in the testing, merit deep appreciation.

Mention of others' contributions cannot be omitted:

--Mrs. Carol Gaines, who reacted to the instruments used in the classrooms and gave much-needed advice,

--George Glasrud, who located the schools used in the pilot testing and the study,

--the Middleton school officials and staff, in particular S. S. Angell, who assisted in the pilot testing,

--the school officials and staff of the school in which the study took place, for their assistance and cooperation, and finally,

--the Wisconsin Research and Development Center for Cognitive Learning computing and statistical consultants, duplication room personnel, and especially the clerical staff, for their advice, service, and assistance in the preparation of the manuscript.

Larry Sowder

ABSTRACT

The Problem

The study had two aims: (1) To explore the ability of pupils of Grades 4, 5, 6, and 7 to give operational evidence of generalizing in selected numerical situations, and (2) to study the effects of differing manners of verbalizing a generalization on the retention of the ability to use the generalizations.

Procedure

Three randomly chosen pupils from each of 24 grade-IQ level-sex blocks were tested individually on a test of randomly ordered items. Each item consisted of the stimulus portions of instances of a generalization. The number of generalizations formed, as evidenced by consecutive correct responses, and the number of instances required before giving consecutive correct responses provided the dependent measures for problem (1). Immediately after the test, each pupil was treated on the items on which he was successful in one of three ways: (a) he reviewed the items with no verbalization, (b) he was required to give a correct verbalization of his version of the generalization, or (c) the tester verbalized a correct statement of the generalization. A retention test containing instances of the items was given one week later. The analyses for problem (1) consisted of univariate grade x IQ level x sex analyses of variance of the dependent measures, and univariate multiple linear regression analyses of these measures, with independent variables chronological age, IQ, arithmetic achievement scores, and mathematical interests scores. For problem (2), retention test scores on selected items were to be analyzed by a one-way analysis of variance.

Results

When pupils did form generalizations, grade means from 4 to 5 instances were required. Over all items, means of 6 to 8 instances resulted. Performance in both the number of instances required and the number of generalizations formed seemed to reach a plateau at Grade 6. There were

statistically significant (.01) differences only among IQ level effects and among grade effects for both the total number of generalizations and the total number of instances. A post hoc analysis of the grade effects showed significant (.05) differences only between grade four and each of the other grades. On a score which combined instances and number of generalizations, the Grade 5 - Grade 6 difference approached significance at the .05 level. The regression equations for these two variables accounted for slightly more than half the respective variances, with age, IQ, and a computation achievement quotient being the most important variables. Retention data indicated no treatment differences although retention was so scanty that no practical significance could have been attached to significant differences, had they appeared.

Conclusions

Indications are that most pupils can form generalizations of the type encountered, although pupils of lower IQ require more instances. With the number of instances needed as a criterion, the optimal grade level at which to offer generalizing tasks appears to be Grade 6 or after. The plateau at Grade 6 supports Piagetian thought, although it may be due to a plateau in computational proficiency.

Chapter I

THE PROBLEMS AND THEIR BACKGROUND

Introduction

The advent of programmed materials, of computer-assisted instruction and of task- and instructional-analyses, and the reappearance of discovery learning have all resulted in recent attention to instructional strategies and techniques. This study explores how well pupils perform under one of these strategies--discovery. In addition, the important issue of the effect on retention of verbalizing a recently-formed generalization is examined.

Claims for Discovery Learning

The importance of generalizing is reflected in statements like ". . . It is only when we attempt to generalize that we fully realize the potentialities of our subject" (Allen, 1950, p. 245). The necessity of generalizing in concept formation and transfer of training is not questioned (transfer and generalization mean the same thing to many writers). The mathematics education community has long recognized the value of generalizing. For example, the 1916 National Committee on Mathematical Requirements of the Mathematical Association of America lists under "disciplinary aims" the ability to discover and formulate a general law (1923, p. 9). More recently Hartung noted, "Many discussions of curriculum issues assume implicitly or explicitly that education is better when it seeks to develop the 'higher' as well as the 'lower' mental processes. One process commonly accepted as 'higher' is the discovery and formulation of generalizations from particular instances" (Husén, 1967, p. 144).

With the advent of (or reemphasis on) discovery teaching, however, the ability to generalize takes on added importance, particularly in light of all the various claims of discovery proponents. Davis (1967), for example, indicates some of these claims when he presents these rhetorical questions:

> Does the essence of discovery lie in giving children a well-developed realization that in science no answers are ever final; all answers are tentative; there is no knowable absolute truth? . . . Is it a teaching style that gives the teacher better feedback from the students on a micro-second to micro-second basis? . . . Do we value discovery because it is a better way to teach a specific mathematical concept? Or because it teaches the student the reason for the concept by letting him experience the need before providing the gratification? Or because it gives the child an authentic realization that mathematics can be discovered (which the man on the street probably does not genuinely believe)?
>
> Do we value discovery classes because they give the child an opportunity to make an accurate assessment of his own personal ability to discover mathematics: that is to say, to learn about himself in one out of . . . a wide variety of meaningful contexts?
>
> Do we value discovery because it is an attempt to sustain a process approach in a school setting where nearly every process seems to gravitate quickly into merely the rote memorization of fact? Because it gives children a better idea of what mathematics is, of where it comes from and of what mathematicians do for a living? . . . Is this why some teachers value discovery--because it can be used to give the student a very genuine feeling of autonomy? . . . (Do) some teachers value discovery as a way of treating the child with respect? . . . Does discovery teaching represent an effort to let the child learn in his own way? . . . Is this why some teachers like discovery teaching, because it keeps alive (if, indeed, it does) the child's creativity and

> curiosity? . . . Is this why some teachers favor discovery, because it does not tend to regiment children (if, in fact, it makes any difference at all in this regard)? . . . Is this a reason for advocating discovery: it limits our ability to make our children grow up in our own image, and releases them to see the world with their own eyes? . . . Does the essence of discovery lie in not allowing the student to verbalize for himself on the grounds that he is inarticulate and will phrase his ideas badly, with the result that his originally clear idea will become contaminated by his inaccurate verbalization? . . . Does the essence of discovery lie in allowing a child to state what he has discovered, with the inevitable consequence that the other children in the class hear a higher percentage of wrong statements than they would if the teacher explained, so that, perhaps, the students learn to listen and to read more critically and more sceptically? (pp. 60-63)

Other distinguished educators and researchers sum up the pro-discovery statements similarly. Max Beberman, of the University of Illinois Committee on School Mathematics, believes that the use of discovery leads to greater understanding by actively involving the student:

> A second major principal which has guided us in developing the UICSM program is that the student will come to understand mathematics if he plays an active part in developing mathematical ideas and procedures. To us this means that after we have selected a body of subject matter to be learned we must design both exposition and exercises in such a way that the student will discover principles and rules (Heath, 1964, p. 23).

Bruner hypothesizes four benefits of discovery learning: increased intellectual potency, intrinsic motivation, the learning of the heuristics of discovery, and enhanced use of memory (1961), and gives his view of discovery teaching:

> Discovery teaching generally involves not so much the process of leading students to discover what is 'out there,' but, rather, their discovering what is in their own heads. It involves encouraging them to say, Let me stop and think about that; Let me use my head; Let me have some vicarious trial-and-error. There is a vast amount more in most heads . . .

> than we are usually aware of, or that we are willing to try to use. You have got convince students . . . of the fact that there are implicit models in their heads which are useful (Shulman and Keislar, 1966, p. 105).

Another summary of the case for discovery, this one by Wittrock (Shulman and Keislar, 1966), reads:

> . . . learning by discovery produces knowledge which transfers to new situations. Through practice at problem solving it develops problem solving ability. It is intrinsically motivating and is its own reward. By being taught to solve problems, to behave in a scientific and inductive fashion, and to go beyond the data, a student is helped to become a mature person. It is a useful conceptualization for the teaching of many subjects in schools. Left to his own resources, the student's individual history will determine the proper sequence of learning activities. It is an important end in its own right. It deserves attention, and students should have some practice at discovering answers for themselves. One must learn to produce rather than to reproduce answers and knowledge. . . (p. 36).
>
> . . . By discovery a student is supposed to learn regularities and concepts within a discipline. But more importantly, he is supposed to learn how to solve problems, to go beyond the data, and to behave as a junior scientist. He is supposed to become motivated and enthusiastic about the discipline. He is to know personal satisfaction because he has selected his own sequence of problems and, through active responses of his own, has succeeded at these problems (p. 42).

Others are emphatic; Hawkins, for example, "takes the position that there are certain kinds of things that can be learned only by discovery" (Shulman and Keislar, 1966, p. xi). Even Ausubel, who has been a strong critic of discovery learning, admits, "Learning by discovery has its proper place among the repertoire of accepted pedagogic techniques available to teachers. For certain designated purposes and for certain carefully specified learning situations, its rationale is clear and defensible" (1963, p. 139). And, "In the early, unsophisticated stages of learning any abstract subject matter, particularly prior to adolescence, the discovery method is invaluable. It is also indispensable for teaching scientific method and effective problem solving skills. Furthermore, various cognitive and motivational factors undoubtedly enhance the learning, retention, and transferability of meaningful material learned by discovery" (1961, p. 22). Hence, there is a great deal of support for, and interest in, a type of teaching (or learning) the essence of which lies in inductive generalization.

Status of Research

What is the research status of the topic of discovery? At a conference on learning by discovery (reported in Shulman and Keislar, 1966), the conclusion was reached that the research evidence was "relatively impoverished" (p. 105), based in part on Wittrock's observation that "almost none of these claims has been empirically substantiated or even clearly tested in an experiment" (p. 33) and on Cronbach's indictment that "there is precious little substantiated knowledge about what advantages it (discovery) offers, and under what conditions these advantages accrue" (p. 76).

Clearly, the consensus is that there is much room for research in discovery learning. The dilemma of where to begin, whether with rather specific, highly focused questions or with more loosely formulated and more general questions, is reviewed by Shulman and Keislar (1966, pp. 195f.). The present study followed the former course by examining in a limited context two important problems: how well <u>do</u> students discover, and how does verbalizing a discovery affect its later use.

Problem 1--A Status Study
Generalization is an influx of divinity
into the mind.--Emerson

It is clear that some children can discover; the many claims obviously are not entirely ivory-tower hypotheses. However, a survey of the literature revealed little related to how well children do discover, in the sense of how much information is needed before the child generalizes. Although the term "discovery" carries many meanings, it would seem well-advised to ascertain children's skill at discovery-with-minimum-guidance for benchmark reasons at least. For example, most discovery teaching takes place in groups; it should be worthwhile to establish whether, and in some quantitative sense how well, each individual can be expected to discover. Benchmarks having been established, one could then compare discovery-with-more-guidance to see whether discovery performance changes. Or, with benchmarks, one could examine any changes in the ability to discover as the child grows older to seek implications for the theory of cognitive growth.

Statement of the Status Study Problem

Part of the present study is devoted to examing the ability of children to discover. Specifically,

Problem 1--A Status Study

a. In selected numerical situations, what is the mean number of instances that need to be presented before boys and girls of grades 4, 5, 6, and 7 and of high, middle, and low intelligence levels show operational evidence of having attained generalizations in the situations?

b. Are there significant grade level differences, IQ level differences, sex differences, grade-IQ interaction, IQ-sex interaction, grade-sex interaction, or grade-IQ-sex interaction for the total number of instances required in these situations or for the number of generalizations formed?

c. If one postulates linear models with independent variables chronological age, IQ, arithmetic achievement scores, and mathematical interests scores, and with dependent variables which reflect the number of instances required or the number of generalizations formed, what portion of the variance is accounted for?

d. What form does the total-learning curve for the pupils in the study take?

Definitions

For the sake of clarity, "numerical situation" means a general expression or formula involving numbers which admits several specific cases, each case of which is an "instance." For example, each of 8, 12, and 32 is an instance of 4 x n. "Operational evidence of having attained the generalization" is defined to mean giving two consecutive correct instances of a general statement (after proper stimuli). To avoid possible confusion, "generalization" refers to that which is generalized, not to the process of generalizing.

Cronbach contrasts what he calls "big-D discovery" and "little-d discovery" (Shulman and Keislar, 1966, p. 78):

> . . . In these studies the learner is nearly always to discover some simple connection or, at best, a formula or inductive generalization. When a writer argues that

> discovery is a thrilling personal experience he seems to have in mind the sort of startling reorganization of interpretation illustrated on the grand scale by Kepler, and on a lesser scale by Kekulé. These "retroductions" are Discoveries that appear to be quite different psychologically from discoveries of simple regularities. Big-D discoveries are infrequent even in the life of the scientist. I doubt that the pupils in today's innovative classroom are having many big-D experiences, and I doubt that the psychologist will be able to arrange conditions so that Discovery will occur while the subject is under his eyes. Hence my account is limited to research on little-d discovery. We should not, however, allow ourselves to think that in these studies we are learning about the effect of retroductive discovery.

Note that the type of discovery involved in the present study is definitely the little-d sort, consisting of giving evidence (as defined above) of having inductively discovered a formula or the technique of a formula.

Related Research

A search of the literature for reports of studies related to the ability of children in grades 4, 5, 6, and 7 to generalize in numerical situations yielded little. In an only-remotely similar status study, Ebert (1946) studied generalization abilities in mathmatics for eighth graders by means of a set of three tests on the material deemed important in elementary mathematics. The first test measured the ability to write an additional instance of a generalization after having seen several instances of the generalization. The second test examined the ability to write a word statement of a generalization after seeing several instances of the generalization, and the third test required the student to write an instance, given a word statement of the generalization. The same 54 generalizations were involved in each of the three untimed tests; tests 1 and 2 were given at the same time with a two-week interval between the administration of these tests and the last test. Each item on each test was scored on a 3 (right and adequate), 2 (right but not adequate), 1 (not right, yet not entirely wrong), 0 (wrong or omitted) scale. Total scores ranged from 55 to 470 (486 perfect), with mean 314.6. The second test, requiring the writing of a generalization, was most difficult; test 3, requiring the writing of an instance of a generalization given in word form, was next most difficult; the test requiring the writing of an instance like several given instances was easiest. Reading ability and IQ were positively correlated with total score, with respective correlations 0.56 and 0.54. Unfortunately, Ebert does not specify how many instances "several" means; in the example illustrated (1946, p. 673) "several" was equal to 10. Ebert's generalizations were chosen from material which in many cases (e.g., quotient times divisor equals dividend) almost certainly had been covered in the classroom, so the generalizations may not have been "discovered" in the sense of the present study.

Only isolated aspects of other studies are relevant to the present status study. The mathematical items most commonly involved in discovery studies have been simple series (e.g., Gagné and Brown, 1961; Hanson, 1967; Hendrix, 1947; Kersh, 1958, 1962), although Kersh (1958) also utilized a geometric model for his series. Cronbach points out that the results of discovery studies are more nearly meaningful to school situations if the discovery tasks are rational--i.e., "as the linkage between the stimulus and the correct response becomes more rational"

(Shulman and Keislar, 1966, p. 78)--and laments those studies in which discovery "has been reduced to sheer trial and error" (p. 79). The type of task chosen for the present study is, then, neither unusual nor, in the Cronbach sense, irrational.

Age groups involved in discovery studies have almost always been eighth grade or above. Kittell (1957) and Stacey (Swenson, Anderson and Stacey, 1949) worked with sixth graders in a non-mathematical situation, one which required choosing from five words the one which did not belong. Osler and Fivel (1961) tested 6, 10, and 14 year olds in a concept-attainment-by-induction experiment in which subjects were to select the "correct" one of a pair of pictures. Up to 150 pairs of pictures (corresponding to "instances" in the present study) were available, with the concept deemed attained if the subject gave 10 consecutive correct responses. Means of the number of errors (per item) before attainment ranged from 9.4 to 70.3. One significant difference from the present study is that Osler and Fivel's subjects were not told exactly what they were to do (choosing the correct picture gave a marble which was credit toward winning a toy). Thompson (1941) found that children in grades 4, 5, and 6 performed better on a sorting task than did children in grades 1, 2, and 3, and interpreted this as representing a greater generalizing ability. The Madison Project has developed materials--and tried them--in kindergarten through grade 9 but has not reported any quantitative research data (Davis, 1964).

Kagan believes that "preadolescent children have learned the joy of discovery, and for them there is an inherent incentive in the discovery method. . . . The method is least appropriate for younger children, especially below the age of nine, who do not have high motivation to master intellectual tasks or who are prone to be impulsive" (Shulman and Keislar, 1966, pp. 160-161. Underlining added.) Ausubel indicates that he feels that discovery is most valuable prior to adolescence (see quote on p.5). The present study, then, deals with a relatively uninvestigated age group, but one for which discovery learning is recommended.

Problem 2--A Verbalizing Study

Thoughts die the moment
they are embodied by words.--Schopenhauer
(quoted in Schwartz, 1948)

A series of quotations may show the gamut of feeling about the role verbalization plays in generalizing. Judd (1927) says, "The human power of generalization is so intimately related to the evolution of language that the two cannot be thought of as existing separately . . . whatever dangers may be connected with the use of language, it still remains true that language is the chief instrument of generalization" (1927, pp. 418-419).

On the other hand, Humphrey asserts that "generalization is possible without verbalization, but verbalization apparently improves and refines the process" (1951, p. 254). Smoke's work on concept formation led him to assert, ". . . it frequently happened that the subjects, though able to discriminate with accuracy and consistency" on instances of a concept, "could not give an acceptable verbal formulation of what they had learned" (1932, p. 20). Heidbreder confirms Smoke's statement in reporting on generalizing in concept formation by observing that "concepts were often used with consistent correctness though the subject was unable to formulate them verbally" (1934, p. 673). Katona joins the Smoke-Heidbreder school of thought when, as a result of his work on problem solving (card tricks, match stick puzzles), he posits that ". . . formulating the general principle in words is not indispensable for achieving applications" (1940, p. 89).

Hendrix offers a divergent possibility by hypothesizing, "Verbalizing a generalization immediately after discovery may actually decrease transfer power" (1947, p. 198). Finally, in a statement about as far from that of Judd's as one can get, Schwartz observes, as subjective evidence of his thesis that verbalization may destroy a generalization, ". . . the experimenter can frequently empathize with the subject, watching him grasp at conceptual threads only to have them vanish when he attempts to formulate them more precisely" (1948, p. 30).

Since the studies of Hendrix and Schwartz, in particular, are most relevant here, they will be examined more closely. Hendrix's subjects learned a generalization (the sum of the first n odd numbers is n^2) either by the usual teacher-taught method (Method I) or by self-discovery by looking at several instances, the discovery being evidenced by some physical cue (smile, gasp, tension) followed by rapid writing of answers. She further treated the discovery subjects by requiring no verbalization of their generalizations (Method II) or by requiring an accurate statement of the generalization (Method III). Her results indicated that of the three groups, the Method II subjects, who were not required to verbalize their generalizations, retained the generalization best (after about two weeks). However, as Ausubel points out (1963, p. 169), some aspects of the measurement, evaluation, and controls used by Hendrix, as well as the 12% significance level obtained, make her conclusions somewhat unsatisfactory in the usual statistical sense. It is disappointing that, although her 1947 article refers to some additional ongoing research, a subsequent article (1961) refers only to the study reported in 1947.

Schwartz's study (1948) dealt with verbalizing effects on sorting tasks with a volunteer adult group who either were college graduates or had IQ's greater than 120. The parts of his study most relevant here consisted of two experiments. In experiment 1, a subject sorted blocks into four categories according to a principle derived from the subject's initial choices of blocks. The sorting was repeated until the subject made no categorizing errors. The subject was then instructed to categorize, in the same way as for the first set of blocks, a second set of blocks, different from the first but still amenable to the same sorting principle. After one attempt with this second task, the subject was then asked to verbalize the principle he used. Experiment 2 was similar but with the verbalizing attempt between the two tasks. The experiments could be summarized as follows:

E_1: Sort first set--sort second set in same way--verbalize.

E_2: Sort first set--verbalize--sort second set as in first.

In E_1, 29 of Schwartz's 40 subjects could do the second sorting, yet 32 of the 40 could not verbalize the principle which they were using. On the other hand, in E_2, where the subjects verbalized before the second sorting task, only 18 of the 40 could perform errorlessly on that second task, a proportion significantly (0.01) different from that in E_1. (Of the 11 subjects who did verbalize correctly, all 11 were able to do the second sorting.) Hence, Schwartz concludes, "A recently formed concept may be destroyed by the unsuccessful effort to verbalize it" (p. 63).

For contrast, in more recent work Hanson (1967), as part of a discovery-vs-reception study, examined the influence of writing definitions for concepts or explanations of techniques for which several instances had been given, as opposed to not writing such statements. The material dealt with concepts and techniques in arithmetic progressions; two levels of subjects--college students in an elementary mathematics course and higher ability eighth graders--provided two sub-studies. In comparing the writers and non-writers, learning scores, retention scores, and transfer scores were found to be significantly higher for the non-writers ($p < .05$) only for the college group, although, with one exception, the results favored the group which did not write the definitions and explanations. (The exception was learning score for the eighth graders, for which group the writers did better--but not significantly--than the non-writers.)

Hendrix's subjects were high school juniors and seniors and college students. Schwartz's subjects, as noted above, were college graduates or had IQ's greater than 120. If the generalizations of such subjects are adversely affected by verbalization, then it would seem certain that younger people of varying levels of intelligence should be similarly influenced. Yet, Hanson's eighth graders did not reveal such an influence when the verbalization was written. The type of verbalization demanded by Hendrix seems to have been quite precise (Hendrix, 1947, pp. 199-200), and Schwartz's sorting tasks may have represented a rather high-level type of problem which was probably different from subject to subject. In any case, the role of verbalizing in discovery learning is unsettled enough to warrant further investigation. Hence, particularly since the status study portion of the present study provided subjects who have attained generalizations (in the sense defined earlier), it was appropriate to examine what influence verbalizing the generalizations seems to play with pupils of a younger age, of differing intelligence levels, and in a school-related task.

Statement of the Verbalizing Problem

Explicitly, the verbalizing problem examined in the present study may be stated as follows:

Problem 2--A Verbalizing Study

Is there a difference in the ability to use numerical generalizations (for which operational evidence of attainment has been given) after having undergone a no-verbalization treatment, a verbalization-by-student treatment, or an experimenter-verbalization treatment on those generalizations?

Definitions

The phrases "numerical generalization" and "operational evidence of attainment" have the same meanings discussed in connection with the status study (p. 7). "No verbalization" means that the subject will not be required to verbalize the generalizations he gives evidence of having attained, whereas "verbalization-by-student" means that he will. "Experimenter-verbalization" means that, after the student has given evidence of having attained a generalization, the experimenter will give a verbal statement of the generalization.

Chapter II

THE DESIGN OF THE STUDY

Cronbach provides a broad design for investigations involving discovery when he proposes that

> . . . we search for limited generalizations of the following form:
>
> With subject matter of this nature, inductive experience of this type, in this amount, produces this pattern of responses, in pupils at this level of development (Shulman and Keislar, 1966, p. 77).

This chapter contains a discussion of these elements in the present study, since they furnished the basic design employed.

The Population and Sample

The public school of a small (population 2000) community in south-central Wisconsin provided the subjects for the study. Located about 15 miles from Madison, the community serves as a shopping and marketing center for farms in the area. Some Madison government and university commuters live there.

The population consisted of those pupils in grades 4 through 7 for whom intelligence test and achievement test records were complete enough to provide the data needed for classification of the pupils and for statistical analyses. Pupils come from both the town and the nearby rural areas. The pupil population has been moderately stable. However, a recent area reorganization did result in some seventh graders being transferred to the school from some rural schools. The seventh grade pupils had, in that sense, more heterogeneous backgrounds, although most of the transferred pupils were not in the population since insufficient testing data were available for them.

Intra-school organization involved departmentalization for the seventh grade and partial departmentalization for grades four through six. The four class sections of seventh grade mathematics were taught by two teachers; all four sixth grade mathematics classes were taught by the same teacher; each of the four fifth grade and four fourth grade sections was taught by a different teacher.

Since the status study investigated intelligence level and sex differences as well as grade differences, pupils in the population were classified into appropriate grade-sex-intelligence level categories, with the intelligence levels being labeled "high," "middle," and "low" to correspond to the approximate high-third, middle-third, and low-third of the distribution of intelligence quotients at each grade level (see below for a more precise description of these levels). Table 1 shows the number of pupils from the population in each category.

To obtain the sample, the pupils in each of these 24 grade-sex-IQ level categories were numbered. Since the verbalizing study involved three treatments, for each category three numbers were obtained from a table of random numbers to determine the sample. Two additional random numbers for each category provided two alternates for use in case of absence at the time of testing. (One exception was the grade 6-boy-high

Table 1

Number of Pupils in Each Grade-Sex-IQ Level Category

	4			5			6			7			Totals		
	B	G		B	G		B	G		B	G		B	G	
High	9	11	(20)	12	12	(24)	4	6	(10)	12	12	(24)	37 +	41 =	78
Middle	17	18	(35)	13	12	(25)	12	11	(23)	16	13	(29)	58 +	54 =	112
Low	12	9	(21)	15	7	(22)	18	10	(28)	5	6	(11)	50 +	32 =	82
Totals	38	38	(76)	40	31	(71)	34	27	(61)	33	31	(64)	145 +	127 =	272
% of pupils in that grade			76%			73%			56%			56%			

category which contained only four pupils; there was only one alternate for that category.) These 72 pupils constituted the sample, 3 in each cell. More pupils in the sample would have unduly extended the testing time required.

Naturally-occurring Data

Data collected for each pupil included grade, sex, and chronological age. These were needed for the status study, chronological age entering into the regression analysis. This grade, sex, and age information was obtained from school records and is included in Appendix I.

Data Requiring Instruments

Intelligence Tests

Intelligence quotients were used as a classification index in one part of the status study and as an independent variable in the linear model. Part of the school's testing program consisted of administration of the Kuhlmann-Anderson Tests (1960, 1963). Booklet CD of the test was administered to pupils in the fall of their third grade year; Booklet EF, in the fall of their sixth grade year. However, school officials did not plan to administer the fall sixth grade testing for the sixth graders in the present study.

The intelligence quotients used for the pupils were the most recent; the third grade testing results were used for grades four, five, and six, and the sixth grade testing results for grade 7. Interpretation of the results, then, must recognize the possible change in test IQ from the third grade testing to the time of the study, in particular for the sixth grade. However, The Sixth Mental Measurements Yearbook reports that, for the Kuhlmann-Anderson Tests, "Test-retest coefficients, with as much as two grades between testings, range from .83 to .92 . . ." (Buros, 1965, p. 737). Hence, IQ scores from these tests do remain moderately stable. In addition, "testing with adjacent forms produces correlations from .77 to .89" (Buros, 1965, p. 737). Moreover, Bloom, after an analysis of several longitudinal studies of IQ growth, concludes that under ideal measurement, "After age 8, the correlations between repeated tests of general intelligence should be between +.90 and unity" (1964, p. 61). Pupils in the study ranged from age 9.2 through 13.5 years. Hence, categorizing the pupils on the basis of scores from the administrations of two different forms (CD and EF) is moderately defensible.

In an effort to make the intelligence classifications mean approximately the same thing from grade to grade, the high, middle, and low categories were delineated by IQ scores rather than determined by taking the appropriate one-third of the pupils at each grade. In addition, as a token recognition of the fact that test measurements have errors, no pupils were chosen from a "buffer" zone of the boundary scores (one exception occurred in that sparsely populated sixth grade cell).

The "high" category was defined by those pupils with Kuhlmann-Anderson IQ's of 116 or higher; the "middle" category included those pupils with IQ's between and including 115 and 103; those pupils with IQ's less than 103 made up the "low" group. Mean IQ's in the various categories for the pupils in the sample are given in Table 2. IQ's of individual pupils in the sample are recorded in Appendix I.

Table 2

Mean IQ's, Pupils in Sample*

	4		5		6		7		IQ
	B	G	B	G	B	G	B	G	Level
High	130.0	126.7	130.7	119.0	122.0	127.0	124.3	122.0	125.2
Middle	111.3	111.3	109.3	109.7	111.7	110.7	111.0	108.7	110.5
Low	98.7	99.7	97.7	93.3	92.3	87.3	93.3	92.3	94.3
Grade-sex	114.7	112.6	112.6	107.3	108.7	108.3	109.6	107.6	Overall
Grade	113.7		110.0		108.5		108.6		110.2

*Three pupils in each grade-sex-IQ level cell, 9 in each grade-sex category, 24 at each IQ level, and 18 in each grade.

Arithmetic Achievement

It is well-known that IQ, although positively correlated with school achievement, is not an infallible guide. Some measure of school achievement was sought as an additional variable to explain variations in performance on the discovery tasks. Since the tasks were numerical, it was plausible that such a variable could be based on the usual standardized arithmetic test scores.

Part of the school's testing program consisted of administration of the Stanford Achievement Tests in the previous spring. Fourth graders in the population had taken Primary II Form W (1964); fifth graders, Intermediate I Form X (1964); sixth graders, Intermediate II Form Y (1965); and seventh graders, Intermediate II Form Y (1965). Each of these tests was given in April previous to the study. Each of these batteries contained two sub-tests entitled "Arithmetic Computation" and "Arithmetic Concepts." In addition, each of the forms taken by the sample's fifth, sixth, and seventh graders included an "Arithmetic Applications" sub-test.

The computation sections are aptly named, consisting of a selection of addition, subtraction, multiplication, and division computations. "Arithmetic Concepts" includes items on numeration, terminology, meanings of fractions, and occasionally an item in which three or four numbers of a pattern are given and the next number is to be chosen. (This last type of item is of interest since it represents to some extent the type of task on the discovery test used in the study. There are, however, only a few items of this type in the sub-tests of the battery.)

"Arithmetic Applications" is roughly translatable as "Story Problems," whereas the concepts section of the Primary II test contains story problems of a simple sort. Split-half reliability coefficients (corrected by the Spearman-Brown formula) for these sub-tests range from .77 to .93 (Kelley and others, 1964).

Scores on these sub-tests were given as grade-equivalents in the school records. Since arithmetic achievement was to be used in the regression analysis, these grade-equivalent scores were transformed by the investigator to "achievement quotients" (by dividing a pupil's grade equivalent score by his grade level when he took the test) to obtain a measure which was less correlated with age. Hence, there were a computation achievement quotient and a concepts achievement quotient for each pupil and, for the pupils in grades 5, 6, and 7, a third quotient, an applications achievement quotient. These quotients are recorded in Appendix I.

Mathematical Interests

The Mathematical Interests Questionnaire (School Mathematics Study Group, 1965) provided measures of the pupil's mathematical interests. This questionnaire yields three scores, one indicative of the pupil's interest in creative sorts of activities in mathematics, a second indicative of his interest in routine sorts of mathematical tasks, and a third for his interest in non-mathematical exercises. The complete questionnaire, with instructions for administering and scoring, is attached as Appendix II.

The questionnaire was administered by the pupil's teachers during the week before the discovery test (described below) to all pupils in the grades involved. Scoring was done by the investigator. Occasionally pupils mismarked the questionnaire, giving ratings, for example, of 1, 2, 2, instead of 1, 2, 3. In cases such as these (there were few in the sample), the ratings were adjusted so that ties retained equal ratings but so that the sum of the ratings was 6. In a 1, 2, 2 case, for example, ratings were recorded as 1, 2.5, 2.5, respectively. Pupils' scores are given in Appendix I.

The scores on the questionnaire are sums of sets of 1-2-3 rankings (greatest interest, 9, to least interest, 27) and as such are not properly suited for most statistical manipulations. Further, the three scores have sum 54 so only two of the scores may be used with even a plausible hope for independence. An underlying normal distribution was assumed, however, so that the creative and routine scores could be used and interpreted in the regression analysis.

Like most interest instruments, the validity and reliability of this questionnaire may be questioned. No information on reliability is available, but face validity can be judged by examining the items which purportedly are samples of creative mathematical activities, routine mathematical activities, or non-mathematical activities (these are indicated on the sample questionnaire in Appendix II).

The Discovery Test

The discovery test provided the dependent measures for the status study and also provided the pupils an opportunity to form some generalizations on which the verbalizing treatments could be given. The discovery test consisted of two warm-up items and eight data-producing items. The test was administered individually so that the experimenter could more objectively tell at what instance the pupil formed his generalization; for contrast, Hendrix inferred the point at which generalizations were formed by noting gasps, smiles, tension (1947, p. 199). The warm-up items served to familiarize the pupil with the task, the experimenter, and the testing situation.

As illustrations, two of the eight items used on the test follow ("item" refers to the generalization and the eleven instances of the generalization). Formulations of all the items used in the discovery test are given in Table 3. A sample complete discovery test is contained in Appendix III.

Sample 1 (Item 2 on the test)

These problems involve multiplying a number ending in 5 by itself. See whether you can find a short-cut to get the answers.

65 x 65 = 4225
25 x 25 =
105 x 105 =
55 x 55 =
45 x 45 =
85 x 85 =
15 x 15 =
75 x 75 =
35 x 35 =
95 x 95 =
205 x 205 =

Sample 2 (Item 8 on the test)

Here four numbers are matched with a fraction by a secret rule. See whether you can figure out the secret rule and give the answers.

2,3,4,5 ---> 6/8
1,3,5,6 --->
2,4,5,9 --->
3,4,5,2 --->
6,4,2,3 --->
8,9,2,1 --->
7,4,5,8 --->
3,6,9,2 --->
2,5,8,3 --->
4,6,7,5 --->
6,8,9,2 --->

Table 3

General Forms of Items on Discovery Test

Warm-up items

a. $(\overset{n\ 1\text{'s}}{1\cdots 1})^2 = 123\cdots n\cdots 321 \qquad (n \leq 9)$

b. n ---> 2n

Short-cut items

1. $1 + 2 + \ldots + (n-1) + n + (n-1) + \ldots + 2 + 1 = n^2$

2. $(10n + 5)^2 = \underline{n(n+1)}25$

3. $\sum_{k=1}^{n}(2k-1) = n^2$

4. $(10n - 1)(10n + 1) = \underline{(n^2-1)}99$

Secret-rule items

5. n ---> n + 4

6. a,b,c ---> ab/(a+c)

7. n ---> n(n+1)

8. a,b,c,d ---> (a+c)/(b+d)

Items for the test were obtained from a pool of items which were tested with 24 fourth and seventh graders at the E. G. Kromrey School, Middleton, Wisconsin, in March and April, 1968. As a result of the pilot, items which were too difficult were discarded. Difficulties of the items used in the discovery test ranged from .30 (portion correct) to .92 in the pilot testing. The pilot testing suggested not ordering the stimulus portion of the instances in increasing order since pupils were prone to search vertically through the instances already presented and to extrapolate rather than to search horizontally (within instances) for clues. The pilot testing also served to indicate the suitability of the vocabulary used by the experimenter, timing information, and the nature of the verbalizations given by the pupils.

The items used in the discovery test could be classified as short-cut items or secret-rule items. A short-cut item involves a situation which is meaningful to the pupil and for which he is to discover a shorter method of obtaining the answer. Secret-rule items, on the other hand, consist of discovering an arbitrary rule (see Table 3).

One warm-up of each of these two types was used, as indicated in Table 3; the eight test-proper items consisted of four of each type. Since a secondary aim of the status study was to examine what, if any, change in performance took place during the discovery test, the eight items were arranged in a quasi-random order, as follows. The four items of one type (short-cut or secret-rule) permit 24 orderings, each used three times in composing the 72 tests needed. To obtain one of the 72 tests, an ordering of the short-cut items was randomly assigned (by means of a table of random numbers) to an ordering of the secret-rule items, with the items interlaced, secret-rule, short-cut, secret-rule, short-cut, etc., or short-cut, secret-rule, short-cut, etc. (Thirty-six of the 72 tests started with secret-rule items, the other 36 with short-cut items.) Doing this resulted in each item appearing the same number of times (9) in the sequence, first item of someone's test-second item of someone's test-etc., and hence allowed considering the composite performances on everyone's first item, second item, etc. Since one of the secret-rule items (Item 7) and one of the short-cut items (Item 2) were partly similar in response, orderings which placed these two items adjacent to each other were not used. This was not an overly severe restriction, leaving 648 of the 1152 possible interlacings from which to select.

A detailed description of how the discovery test was administered is contained in the flow-chart in Figure 1 and in the instructions included with the test in Appendix III. Briefly, the first instance of an item was exposed for 7-10 seconds; then a cardboard mask was moved to reveal the second instance (which was incomplete--see samples above). The pupil was given 15 seconds in which to respond. If he responded, the correct answer was written immediately and he was permitted to study all previous work on the item for 15 seconds; if he did not respond, the correct answer was written and he was permitted to study all previous work on the item for 15 seconds. The next instance was then exposed, he again was allowed 15 seconds in which to respond, after which (or after his response) the correct answer was written and he was given 15 seconds in which to study previous work, etc.

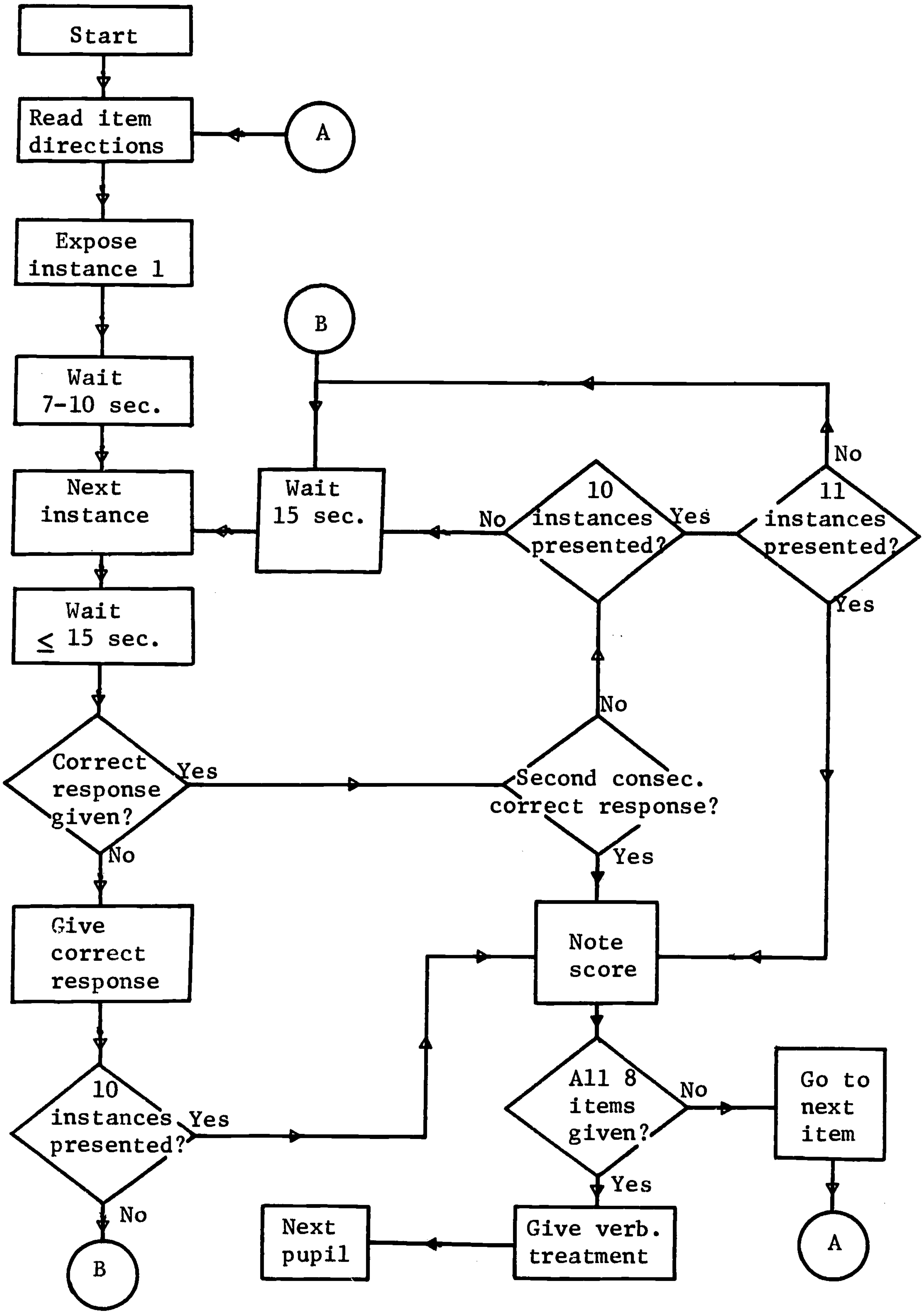


Fig. 1. Flow-chart for Discovery Test Administration

If the pupil gave two consecutive correct responses, he was adjudged to have formed a generalization and proceeded to the next item. On each item, a pupil was given an item-instances score as follows:

$$\text{item-instances score} = \begin{cases} k \text{ if pupil first gave two correct responses after } k \text{ instances, } 1 \leq k \leq 9 \\ 10 \text{ if pupil did not.} \end{cases}$$

The eleventh instance in each item was used only if the pupil gave a correct response on the tenth instance without giving a correct response on the ninth instance.

Hence, the pupil received eight item-instances scores. These were summed to give a total-instances score. Total-instances scores could range from 8 to 80, inclusive. Note that lower scores indicate that generalizations were formed with fewer instances; higher scores indicate that more instances were required to form the generalizations.

". . . the problem of setting the objectives in problem-solving and conceptualizing experiments is not a nuisance to be got rid of by 'cute' instructions that conceal the purpose of the research," as Bruner and others put it (1956, p. 243). As the instructions for the items in the present study indicate, the pupil was informed of precisely what he was to do--discover a short-cut or a secret-rule. Providing this information, of course, meant that the tasks were not pure discovery tasks, but giving the information at least made the pupil aware of what he was to do. In addition, underlining was used in some of the more difficult items (see Sample 1 above) as a visual cue to the pupil; in others, some vocal emphasis was given by the experimenter as he wrote correct responses ("the sum of the first nine odd numbers is 81"). Since the task was to involve relatively unguided discovery, no specific hints (e.g., "what if you added 4?") were given, nor, as Wills found to be effective (1967), was a "target task" given.

Opportunity to Learn

Romberg has pointed out the importance of knowing the subject's prior experience with research tasks in interpreting performance on the tasks (1968). Hence, a two-fold attempt was made to establish what previous discovery experiences the pupils in the present study had had by (a) examining the text-series and (b) questioning the pupils' present teachers.

A search through the textbooks used, however, revealed few tasks that could be called discovery oriented (McSwain and others, 1965abc, 1963). Hence, inasmuch as all the pupils in the sample had used this text series through whatever part of grades 4-7 they had covered, the texts they used gave inconsequential discovery experiences. The grade 4 text has seven exercise sets which consist of column a, column b exercises which are related and could provoke a discovery even though such is not the intent of the exercises. These lists include commutativity of addition, related addition-subtraction facts, related multiplication-division facts, or related multiplication and/or division equations (McSwain and others, 1965a, pp. 19, 34, 35, 36, 78, 98, 103, 157, 191). The grade 5 text, however, does not even provide such opportunities such as these. The grade 6 text includes three situations in which discovery in encouraged--with locating decimal points when multiplying by powers of ten or one-tenth and when dividing by powers of ten (McSwain and others, 1965c, pp. 158-9, p. 164, p. 182). These pages had not been covered by the sixth graders in the sample.

The grade 7 text (McSwain and others, 1963) includes two questions, both geometric, which could be remotely regarded as permitting discovery. The first asks the pupils to conclude something about the measures of vertical angles by measuring two pairs of vertical angles, comparing the measurements, and then comparing results with classmates (p. 265). The second case consists of asking the pupils whether they notice anything about the measures of corresponding angles of parallel lines (immediately before stating the proposition and immediately after noting the converse). In summation, the texts used by the pupils in the sample provided only a negligible amount of activity even remotely related to the tasks of the discovery test.

At the initial meeting with the teachers whose classes were to provide the population of the study, no teacher claimed to make use of discovery methods in the classroom. In view of the texts used, the teacher would certainly have had to make a conscious effort to provide discovery experiences, so it is most likely that the pupils received negligible, if any, amounts of practice at discovery. Unfortunately, however, a conversation with the sixth grade teacher, who taught all four sections of that grade, revealed that she had shown three of her sections short-cuts similar to that of item 2 (Table 3). Hence, although as a group the teachers did not provide discovery tasks, the sixth graders performance on the discovery test must be weighed carefully before reaching any conclusions, not only because of the practice with short-cuts similar to that of the one item but also because this experience with short-cuts might have given them indications of the sort of thing to seek.

Verbalizing Treatments

The verbalizing treatments were administered at the conclusion of the discovery test with the items for which the pupil gave evidence of having generalized. The three pupils in each category were assigned at random to the three verbalizing treatments. The guidelines used by the experimenters in administering the verbalizing treatments are included in Appendix III. About one minute was devoted to each item.

The no-verbalizing treatment was the easiest to keep uniform from subject to subject, for obvious reasons. The only difficulty anticipated was that pupils undergoing the no-verbalizing treatment might spontaneously verbalize. Such spontaneous verbalizing did not materialize.

In the experimenter-verbalizing and the subject-verbalizing treatments, giving or requiring precise, fully general verbalizations of the generalizations was not done, mainly because at these ages the pupil could not be expected to understand or produce such statements. Rather loose verbalizations were to be given and accepted. This practice contrasts sharply with that of Hendrix, who apparently required accurate, general statements of the generalization (1947, p. 199). On the other hand, this giving and requiring the verbalization only in less formal, more familiar terminology might enhance the pupil's retention of the generalization. Kendler, for example, feels that "when a person discovers something, he is able to formulate it in his own language so that it fits in--that is, meshes--with his linguistic network. This allows him to retain and apply the idea he has discovered more effectively because it becomes part of a well-practiced and highly integrated habit system" (Shulman and Keislar, p. 172).

As evidenced by his statement of principles, Davis (1964) also supports accepting the child's natural response:

> Principle 1. The teacher should always use "clean language."
> Principle 2. We do not expect this of the student.
>
> * * *
>
> The child has the right idea, and deserves credit for understanding it rather than censure for not yet having mastered all the intricacies of language.
> Principle 4. As a matter of fact, we prefer an answer in the unhesitating, genuine language of the child's own words, rather than a glib repetition of what the teacher said (pp. 15-16).

The types of verbalizing treatments in the present study are not the only difference from Hendrix's work (1947) and perhaps other differences should be made explicit to dispel any impression that the present study is a replication of Hendrix's study with a different age group and more items. (1) Pupils in the present study were told to look for a short-cut or secret rule; Hendrix's subjects apparently were not. (2) In the present study, a pupil knew whether his response was correct (the experimenter told him), whereas Hendrix's subjects received no corroboration from someone else. That her senior high and college subjects should place enough faith in a discovered rule to write several responses solely on the basis of the discovery is of interest, especially if the subjects were not told such a rule existed. (3) In the present study, for a given item the number of instances to which a pupil was exposed might vary from pupil to pupil. Hendrix's subjects all were exposed to the same number of instances. (4) On the other hand, after pupils in the present study did make a discovery, the amount of practice on new instances was constant; depending on where in the course of their examination they formed a generalization, Hendrix's subjects might practice on varying numbers of instances.

Administration of the Discovery Test

The discovery test was administered on consecutive days, November 11-14, 1968. All pupils at a given grade were tested on the same day: Monday--grade 4, Tuesday--grade 5, Wednesday--grade 6, and Thursday--grade 7. The decision to test all pupils at a given grade on the same day instead of attempting to control for possible day-effects facilitated the administration of the follow-up test (see below).

Three male interviewers were used, two research trainees in mathematics education at The Wisconsin Research and Development Center for Cognitive Learning and the investigator. Each of the interviewers was familiar with discovery learning literature and the mathematics involved in the discovery test items. The week before the testing, the interviewers underwent a half-day training period, during which the study and testing protocol were discussed, with the interviewers performing trial administrations of the test with pupils of the ages to be tested.

Attempts to control effects other than those studied took various forms. Each interviewer administered the discovery test to three pupils in the morning and three pupils in the afternoon. Tests were assigned to pupils randomly. In an effort to control for possible interactions between interviewer and IQ level, sex, or verbalizing treatment, equal numbers of pupils in these categories were assigned to interviewers by means of a Greco-Latin square for the morning pupils and for the afternoon pupils. Testing was scheduled to control for morning-afternoon testing and IQ level or verbalizing treatment interactions; over all grades, sexes were balanced between morning and afternoon, but not at each grade. The testing schedule is included in Appendix IV.

The Follow-up Test

The influence of verbalizing mode on retention of ability to use a generalization for which evidence of attainment had been given was to be determined by a follow-up test containing instances of the generalizations. The follow-up test contained two parts relevant to the study. Part 1 included instances of the four short-cut items without any of the underlining or cueing contained in the discovery test; this part, then, tested a strong form of retention--recognize an opportunity to use a previously formed generalization and then remember how to use the generalization. Part 2 included as its first four items instances of the four secret-rule items of the discovery test. Since the teachers were asked not to relate the test to the earlier interview-test and since there were no identifying marks on the test which would relate it to the experiment, these four items were the first open reminders of the discovery test. The next four items presented instances of the short-cuts (different instances from those of Part 1). Instances in both parts were chosen from the last four instances in the discovery test items; pilot testing indicated that few pupils attained a generalization after that point. Hence, Part 2, containing the reminder of the interview test, provided a measure of how well the pupil could use the generalization when the stimulus was related to the generalization-forming situation.

In both Part 1 and Part 2, instances of a particular type--short-cut or secret-rule--were presented in the same order as the pupil encountered the related items during the discovery test. Cueing ($\diamond$, Φ , $\triangle$, or Σ) over the arrow in the secret-rule items was used to enable the pupil to distinguish the secret-rules since their visual appearances are quite similar. These symbols were not described verbally in the course of the discovery test but were treated as important labels for the secret-rules. The symbols were associated with the secret-rule items in the same order, regardless of the order of the secret-rule items.

Deciding how much time to leave between the discovery test and the follow-up test was difficult. Hendrix (1947) used two weeks with the sum-of-first-n-odd-numbers generalization; however, her subjects were high school upperclassmen or college students and her follow-up test apparently allowed for re-discovering the generalization since 27 of her 40 subjects were correct more often than they applied the generalization (p. 207). Kittell found that his sixth grade minimum-guidance subjects remembered slightly less than half of the principles they had learned four weeks earlier (1957, p. 401). Since a one-week interval was administratively convenient and semed quite consonant with Kittell's choice of waiting-period, the follow-up test was administered one week after the pupils of a particular grade level were exposed to the discovery test. The test was administered by the classroom teacher, with pupils who were not in the sample receiving a test similar in form to the follow-up test. The relevant parts of a sample follow-up test are included in Appendix V.

Summary of Instrument Data Collected for Each Pupil

Data	Source
IQ score	Kuhlmann-Anderson (school records)
Arithmetic achievement quotients a. computation b. concepts c. applications (grades 5, 6, 7)	Stanford Achievement Tests (school records)
Mathematical interests scores a. creative b. routine c. non-mathematical	Mathematical Interests Questionnaire
Item-, total-instances scores, number of generalizations attained	Discovery test
Retention scores a. short-cut: no reminder of discovery test b. secret-rule and short-cut: reminder of discovery test	Follow-up Test

Table 4

Summary of Grade Means on Independent Variables

Grade	Age	IQ	Creative	Routine	Comp AQ	Conc AQ	App AQ
4	9.7	112.9	18.7	19.2	0.98	1.08	--
5	10.8	109.9	18.9	18.2	1.09	1.07	1.05
6	11.7	108.5	18.1	18.1	1.05	1.15	1.21
7	12.7	108.7	21.0	17.6	0.97	1.00	0.96
4-7	11.2	110.0	19.2	18.3	1.02	1.07	--

Summary of Standard Deviations on Independent Variables

Grade	Age	IQ	Creative	Routine	Comp AQ	Conc AQ	App AQ
4	0.4	13.7	2.7	2.7	0.14	0.33	--
5	0.4	13.5	2.9	2.8	0.22	0.29	0.24
6	0.5	15.7	3.7	3.0	0.19	0.23	0.32
7	0.4	13.6	3.4	3.7	0.18	0.22	0.28
4-7	1.2	14.0	3.3	3.1	0.19	0.27	--

Statistical Analyses

The Status Study

Wittrock cites inadequate statistical analysis as being a common shortcoming in many discovery studies (Shulman and Keislar, 1966, p. 43). Fortunately, most of the questions investigated in the present study suggest standard statistical treatments. The status study questions stated in Chapter 1 (pp. 6-7), for example, were handled as follows. Problem 1a (Ch. 1, p. 6), on the mean number of instances required for generalization by pupils in the various categories, was answered by the calculation of appropriate means. Problem 1b (Ch. 1, p.7) was answered by three-way analyses of variance of the total-instances scores and the number of generalizations formed. These analyses were performed by a University of Wisconsin CDC 1604 computer using a program prepared by Houston (1967).

The regression analysis used for Problem 1c (Ch. 1, p. 7) warrants special explanation. The type of analysis employed is described by Bock as follows:

> Yet another approach to tests of multivariate hypotheses is the so-called "step-down" analysis of Roy and Bargmann. Computationally this procedure is just a sequence of analyses of covariance. It leads to a univariate test of significance of the second variate eliminating the first, of the third eliminating the first and second, and so on down to the last variate eliminating all previous variates. Roy and Bargmann have shown that under the hypothesis of no group differences, these tests of significance, including one for the first variate, are statistically independent (Cattell, 1966, p. 828).

The computer program used for this analysis was due to Finn (1967).

Total-learning curves were plotted in response to the secondary Problem 1d (Ch. 1, p. 7) on what improvement seemed to take place during the discovery test.

The Verbalizing Study

Whether the three verbalizing treatments differed in their effects on retention was to be answered by one-way analyses of the scores on the follow-up test. Since the items on which pupils formed generalizations could vary widely from pupil to pupil, retention was to be analyzed on only a subset of the items chosen by studying the item difficulties and choosing items of low enough difficulty so that several pupils would have formed generalizations for them.

Chapter III

ANALYSIS OF TEST RESULTS

This chapter contains the results of the pupils' performances on the discovery test and the follow-up test, along with the statistical analyses chosen to attempt to answer the questions of the status study and the verbalizing study.

The Discovery Test

Item Analysis

An item analysis of the eight items of the discovery test is given in Table 5. The analysis was performed on a University of Wisconsin Control Data Corporation 1604 computer using a program prepared by Baker and Martin (1968) and included for each item a report of its difficulty (as the portion giving evidence of having formed a generalization), an item-test correlation coefficient R, X_{50}, and β. These last three statistics being less familiar, they will be described briefly below. Recall that the content of the items is described in Table 3, p. 27.

The item-test correlation R reported is the point biserial correlation, which is appropriate for use when the scoring of an item is dichotomous (Lord and Novick, 1968, pp. 335-336). Item scores (not item-instances scores) on the discovery test were regarded as dichotomous (0 = no generalization and 1 = generalization). For items such

Table 5
Item Analysis, Discovery Test

Item	Grade	Difficulty	R	X_{50}	β
1	4	.278	.707	.624	2.888
1	5	.611	.627	-.354	1.320
1	6	.611	.532	-.417	.920
1	7	.500	.680	0	1.628
1	4-7	.500	.653	0	1.423
2	4	.111	.425	1.730	.996
2	5	.278	.550	.802	1.084
2	6	.389	.515	.431	.867
2	7	.278	.469	.941	.803
2	4-7	.264	.533	.878	1.035
3	4	.222	.725	0	0
3	5	.611	.445	-.499	.686
3	6	.667	.758	-.438	5.356
3	7	.667	.596	-.557	1.218
3	4-7	.542	.696	-.120	1.796
4	4	0	0	0	0
4	5	.222	.047	11.552	.066
4	6	.389	.620	.358	1.280
4	7	.333	.673	.494	1.787
4	4-7	.236	.532	.981	1.077
5	4	.778	.349	-1.570	.558
5	5	.889	.251	-2.932	.458
5	6	.889	.271	-2.717	.503
5	7	.944	.363	-2.150	1.104
5	4-7	.875	.339	-2.115	.648
6	4	.444	.661	.168	1.494
6	5	.722	.704	-.627	2.773
6	6	.889	.677	0	0
6	7	.889	.615	0	0
6	4-7	.736	.716	-.654	3.670
7	4	.167	.123	5.276	.187
7	5	.333	.627	.530	1.398
7	6	.500	.689	0	1.716
7	7	.556	.681	-.163	1.660
7	4-7	.389	.633	.351	1.357
8	4	.556	.515	-.216	.850
8	5	.667	.627	-.530	1.398
8	6	.833	.651	-.997	4.042
8	7	.889	.529	-1.391	1.830
8	4-7	.736	.614	-.763	1.475

Items scores 0 (no generalization) or 1 (generalization formed)
R based on total score for test (number of generalizations formed)

Time required for the discovery test ranged from 1-4 minutes for the warm-up items and from 25-40 minutes for the test itself.

Test reliability coefficients (Hoyt) were as follows: grade 4, .56; grade 5, .60; grade 6, .75; grade 7, .74; and all grades, .75. Taking into consideration that the discovery test contained only eight items, these are reasonably adequate. Using the Spearman-Brown formula (Lord and Novick, 1968, p. 112) to estimate what the reliabilities would have been had there been 20 items on the test, one gets these results: grade 4, .76; grade 5, .79; grade 6, .88; grade 7, .88; and all grades, .88.

Summary of Group Performances on Items

Recall that an item-instances score less than 10 indicates that the pupil gave two correct responses immediately after that number of instances (i.e., "formed a generalization"); an item-instances score of 10 means the pupil did not give such responses ("had not formed a generalization"). Appendix VI contains tables giving the mean number of instances by each group of interest for each item and recording the number of pupils in each category who formed a generalization for each item.

It should be repeated that each pupil received the items in a different order. Differences in mean item-instances (or number of pupils generalizing) from category to category, then, may be partially attributable to differences in item-order. A given sequence of items may have been perfectly ordered for one pupil's best performance; the opposite may have been true for another pupil.

Results--The Status Study

Problem 1a

In selected numerical situations, what is the mean number of instances that must be presented before boys and girls of grades 4, 5, 6, and 7 and of high, middle, and low intelligence levels show operational evidence of having attained the generalizations? The data relevant to problem 1a are given in Table 6, which contains the mean total-instances scores on the discovery test for the various groups in the study.

Table 6
Mean Total-instances Scores

Grade	Sex	IQ Level High	IQ Level Middle	IQ Level Low	Grade Boy	Grade Girl	Grade
4	Boy	67.7	68.7	71.0	69.1		67.4
	Girl	57.0	73.0	67.0		65.7	
5	Boy	49.0	50.3	60.7	53.3		56.4
	Girl	54.7	57.0	67.0		59.6	
6	Boy	47.7	55.7	56.3	53.2		48.6
	Girl	33.7	47.0	51.0		43.9	
7	Boy	31.0	57.0	55.0	47.7		49.6
	Girl	46.0	48.3	60.0		51.4	
	Boy	48.8	57.9	60.8	55.8		Grand mean 55.5
	Girl	47.8	56.3	61.3		55.1	Overall std. dev. 14.4
IQ level		48.3	57.1	61.0			
4		62.3	70.8	69.0			
5		51.8	53.7	63.8			
6		40.7	51.3	53.7			
7		38.5	52.7	57.5			

As might be expected, pupils in a higher intelligence category required fewer instances on the average than pupils in a lower category. Since they were not so experienced with arithmetic--particularly multiplication--it is not surprising that the fourth graders required more instances, with the 4H (grade 4, high IQ level) group only slightly better than the 5L group. The interesting thing is that the trend of improvement from grade to grade seems to stop at grade 6, with the seventh graders performing slightly worse, although this is not true at the high IQ level. It may be that (1) the sixth graders' classroom experience mentioned in chapter 2 exaggerated the sixth graders' performance, (2) there is a plateau in performance on discovery tasks of this sort once a certain proficiency in arithmetic is attained, or (3) the change in cognitive structure believed to exist by Piaget (Flavell, 1963), which could influence performance on such tasks, might account for the pattern. Whether the differences noted above are statistically significant is examined under problem 1b below.

Tables 7-9 contain results not specifically referred to in problem 1a but which are related to the problem. Table 7 records mean performance on the discovery test in terms of the mean number of generalizations formed (out of 8).

Since both the above statistics--number of generalizations formed and total-instances score--are of interest, a single score which reflected both was defined by the author as follows:

$$G/I = 10 \times \frac{\text{number of generalizations formed}}{\text{total-instances score}}.$$

The G/I score may be viewed as the number of generalizations formed per 10 instances.

The G/I score differentiates pupils who had equal total-instances scores but formed different numbers of generalizations, or who formed equal non-zero numbers of generalizations but had different total-instances scores. Greater G/I scores indicate better performance on the discovery test. G/I scores for individual pupils are given in Appendix VI. G/I scores for categories are given in Table 8.

Table 7
Mean Number of Generalizations (possible 8)

Grade	Sex	IQ Level High	IQ Level Middle	IQ Level Low	Grade Boy	Grade Girl	Grade
4	Boy	2.7	2.3	2.2	2.2		2.6
	Girl	4.7	2.0	2.0		2.9	
5	Boy	5.7	5.3	3.7	4.9		4.3
	Girl	4.7	4.3	2.3		3.8	
6	Boy	5.0	4.3	3.7	4.3		5.2
	Girl	7.7	5.7	4.7		6.0	
7	Boy	7.7	4.3	4.3	5.4		5.1
	Girl	5.7	5.0	3.3		4.7	
	Boy	5.3	4.1	3.3	4.2		Grand mean 4.3
	Girl	5.7	4.3	3.1		4.3	
IQ Level		5.5	4.2	3.2			Overall std. dev. 2.2
4		3.7	2.2	1.8			
5		5.2	4.8	3.0			
6		6.3	5.0	4.2			
7		6.7	4.7	3.8			

Table 8

G/I Scores by Category

$$\left(G/I = 10\text{x} \frac{\sum_{\text{category}} \text{no. of generalizations}}{\sum_{\text{category}} \text{no. of instances}}\right)$$

Grade	Sex	IQ Level High	IQ Level Middle	IQ Level Low	Grade Boy	Grade Girl	Grade
4	Boy	.39	.34	.23	.32		.38
	Girl	.82	.27	.30		.44	
5	Boy	1.16	1.06	.60	.92		.77
	Girl	.85	.76	.35		.63	
6	Boy	1.05	.78	.65	.81		1.06
	Girl	2.28	1.21	.92		1.37	
7	Boy	2.47	.76	.79	1.14		1.02
	Girl	1.23	1.03	.56		.91	
	Boy	1.08	.71	.55	.76		Overall
	Girl	1.18	.76	.50		.79	.77
IQ level		1.13	.73	.53			
4		.59	.33	.27			
5		1.00	.90	.47			
6		1.56	.97	.78			
7		1.73	.89	.67			

Table 7 shows much the same pattern as Table 6, with the number of generalizations formed increasing with higher intelligence level and with grade until the slight decrease from grade 6 to grade 7. A grade-IQ level-sex analysis of variance of the number of generalizations formed is given in connection with problem 1b. Table 8 also shows the same sort of pattern as Tables 6 and 7.

Finally, Table 9, which is included even though it is not based on performance on the complete test, gives the mean number of instances required on the items for which the pupils did generalize (item-instances scores $\leq$ 9).

Table 9

Mean Number of Instances on Items for which Generalizations Were Formed

Grade	Sex	IQ Level High	IQ Level Middle	IQ Level Low	Grade Boy	Grade Girl	Grade
4	Boy	5.38	5.14	4.60	5.01		
	Girl	5.07	6.50	3.50		5.04	5.06
5	Boy	4.64	4.44	4.73	4.55		
	Girl	4.57	4.69	4.43		4.59	4.56
6	Boy	3.53	4.38	3.55	3.82		
	Girl	3.95	4.02	3.79		3.98	3.92
7	Boy	3.45	4.69	4.23	4.06		
	Girl	4.00	3.67	4.00		3.88	3.98
	Boy	4.06	4.60	4.23	4.28		
	Girl	4.32	4.43	3.92		4.26	Overall 4.27
IQ Level		4.20	4.51	4.08			
4		5.19	5.76	4.00			
5		4.55	4.55	4.61			
6		3.79	4.27	3.68			
7		3.78	4.15	4.13			

Table 9 indicates that when pupils do generalize on items of the types on the discovery test, they require from 3 to 6 instances. While also indicating many of the same differences as in Tables 6-8, Table 9 contains perhaps surprising results: (1) the 4L group appears to have performed better than either the 4H or 4M group, and similarly, the 6L group seems to have been better than the 6H or 6M group, and (2) overall, the low IQ group appears to have performed slightly better than the high IQ group. These anomalies can be accounted for in part by noting that the low IQ levels formed generalizations mainly on the three easiest items whereas the higher IQ levels also formed generalizations on harder items which require more instances. Hence, the entries in Table 9 cannot be compared meaningfully.

Hence, to look at performance on items for which most of the pupils generalized, yet to keep the items the same so comparisons could be more meaningful, the mean number of instances required on the three easiest items (Items 5, 6, and 8) were totaled. Table 10 summarizes these means and gives much the same pattern as Tables 6, 7, and 8.

Problem 1b

Are there significant grade level differences, IQ level differences, sex differences, grade-IQ interactions, IQ-sex interactions, grade-sex interactions, or grade-IQ-sex interaction for the total number of instances required, or for the total number of generalizations, in these situations? This problem resulted in the analyses of variance summarized in Tables 11, 12, and 13.

Table 11 indicates that only effects due to grade or IQ level were different (.01). Means for the sources of variation in Table 11 which gave a significant F are plotted in Figure 2.

Table 10

Mean Number of Instances,
Items 5, 6, and 8

Grade	Sex	IQ Level: High	IQ Level: Middle	IQ Level: Low	Grade Boy	Grade Girl	Grade
4	Boy	7.4	6.9	8.0	7.4		6.9
	Girl	4.9	8.3	5.8		6.3	
5	Boy	5.2	5.1	8.0	6.1		5.6
	Girl	4.3	4.2	6.6		5.0	
6	Boy	3.7	4.9	4.9	4.5		4.2
	Girl	2.7	4.3	4.9		4.0	
7	Boy	3.6	5.1	5.4	4.7		4.3
	Girl	3.1	2.8	5.8		3.9	
	Boy	5.0	5.5	6.6	5.7		
	Girl	3.8	4.9	5.8		4.8	Overall 5.2
IQ Level		4.4	5.2	6.2			
4		6.2	7.6	6.9			
5		4.8	4.7	7.3			
6		3.2	4.6	4.9			
7		3.3	3.9	5.6			

Table 11

ANOVA--Total-instances Scores (n = 72)

Source	d.f.	MS	F
Grade	3	1354.8	9.5**
IQ Level	2	1011.0	7.1**
Sex	1	8.7	<1
IQ-sex	2	6.9	<1
Grade-sex	3	225.1	1.6
IQ-grade	6	77.7	<1
Sex-IQ-grade	6	106.2	<1
Within cell	48	142.4	

** $p << .01$ $F_{.99;3,48} = 4.22$ $F_{.95;3,48} = 2.80$

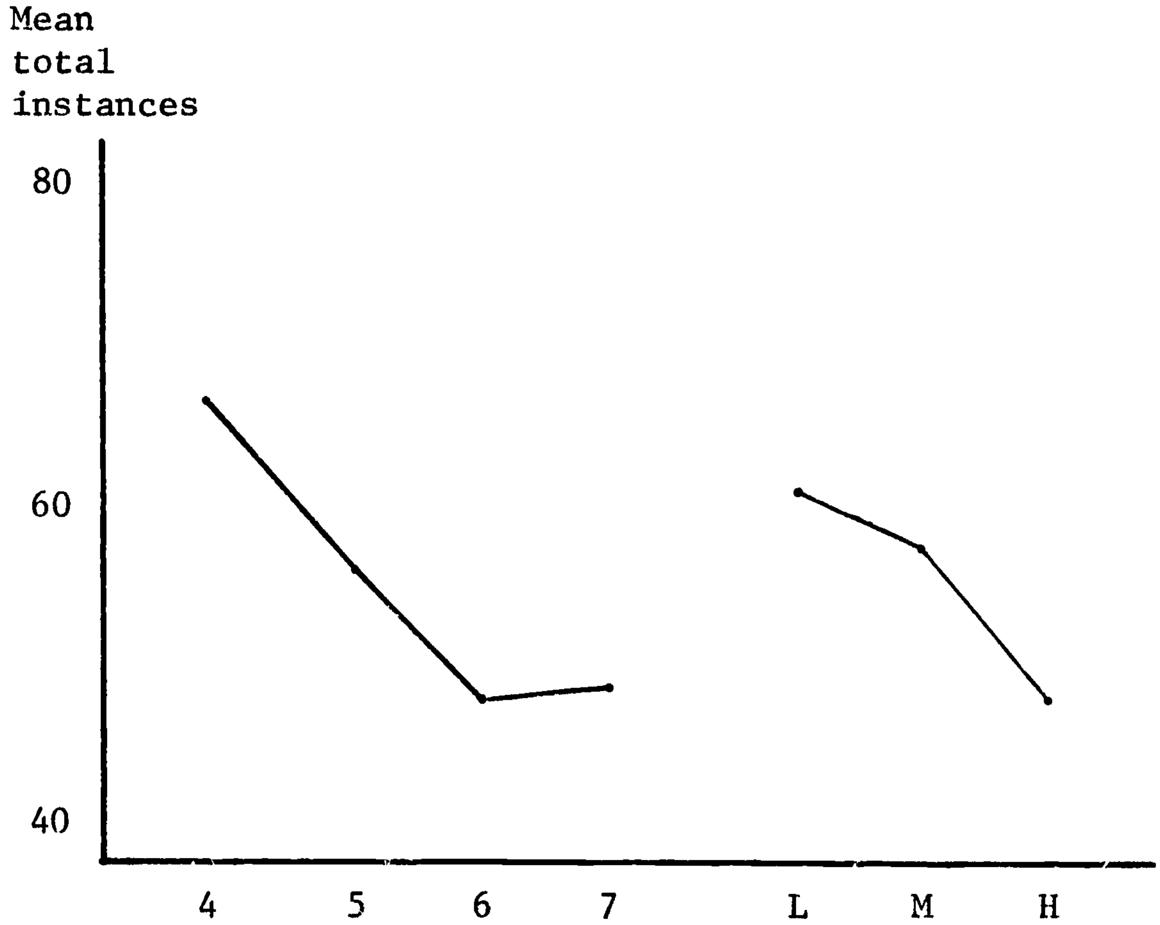


Fig. 2. Mean Total-instances vs Grades and IQ Level

Table 12 shows significantly different (.01) effects due to grade and IQ level. Means for the sources of variation in Table 12 which gave a significant F-test are plotted in Figure 3.

Table 12

ANOVA--Total Number of Generalizations
(n = 72)

Source	d.f.	MS	F
Grade	3	26.19	8.2**
IQ	2	30.60	9.6**
Sex	1	.22	<1
IQ-sex	2	.68	<1
Grade-sex	3	7.52	2.36
IQ-grade	6	1.06	<1
Sex-IQ-grade	6	1.81	<1
Within cell	48	3.19	

**$p << .01$ $F_{.99;3,48} = 4.22$ $F_{.95;3,48} = 2.80$

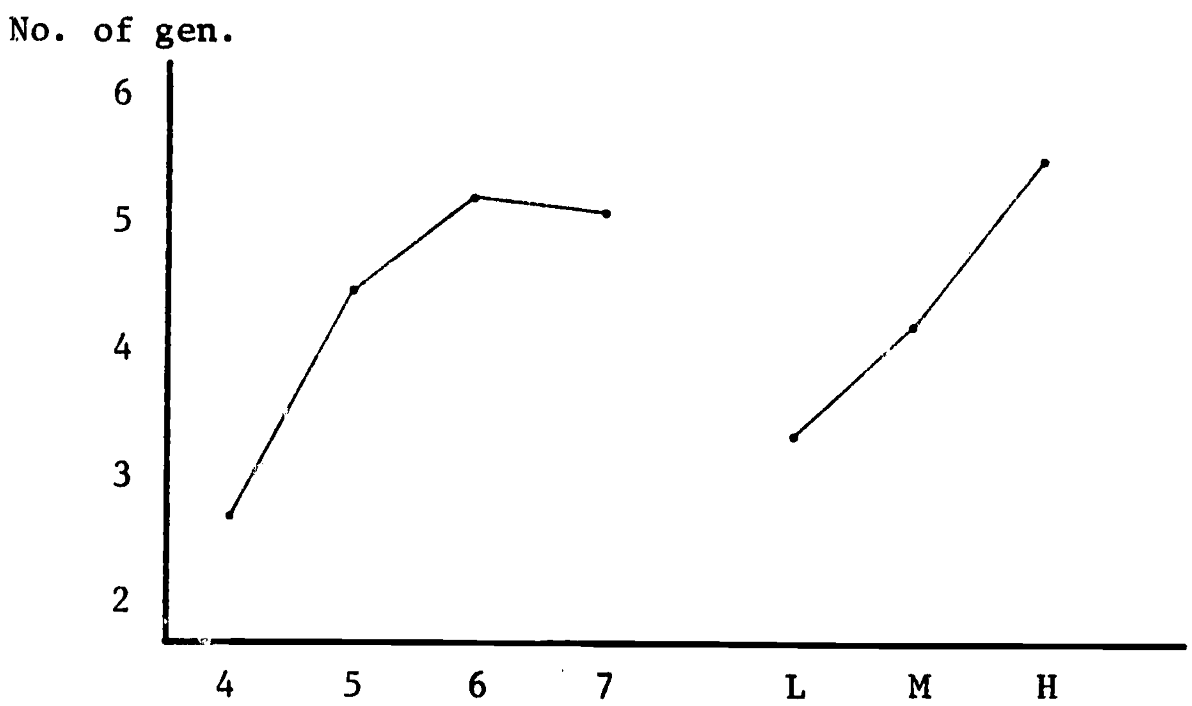


Fig. 3. Mean Number of Generalizations _vs_ Grades and IQ Levels

Table 13 gives the result of the analysis of variance of the G/I scores. Since the G/I score, being based on different numbers of instances from pupil to pupil, can be considered to be on only an ordinal scale, the arithmetic involved in calculating means and variances is not applicable (Siegel, 1956, p. 26). Hence, the usual parametric analysis of variance, which requires an interval scale, was not used. However, the Kruskal-Wallis one-way analysis of variance technique can be applied to measurements on such a scale (Siegel, 1956, pp. 184-193). Grade differences were examined by this technique; Table 13 summarizes the analysis. The statistic H is distributed as a chi-square with k-1 degrees of freedom (k = number of treatments--grades in this analysis). As the table shows, the hypothesis that the G/I scores for the grades are equal may be rejected at the .001 level. G/I scores for the four grades are plotted in Figure 4.

Table 13

Kruskal-Wallis Analysis of Variance
for Grade Effects on G/I Scores

Grade	4	5	6	7
Sum of ranks	958.5	664.5	498.5	506.5

$$H = \frac{12}{72(72+1)}\left(\frac{958.5^2}{18} + \frac{664.5^2}{18} + \frac{498.5^2}{18} + \frac{506.5^2}{18}\right) - 3(72+1)$$

$H = 17.6$ $\qquad \chi^2_{.999;3} = 16.27$

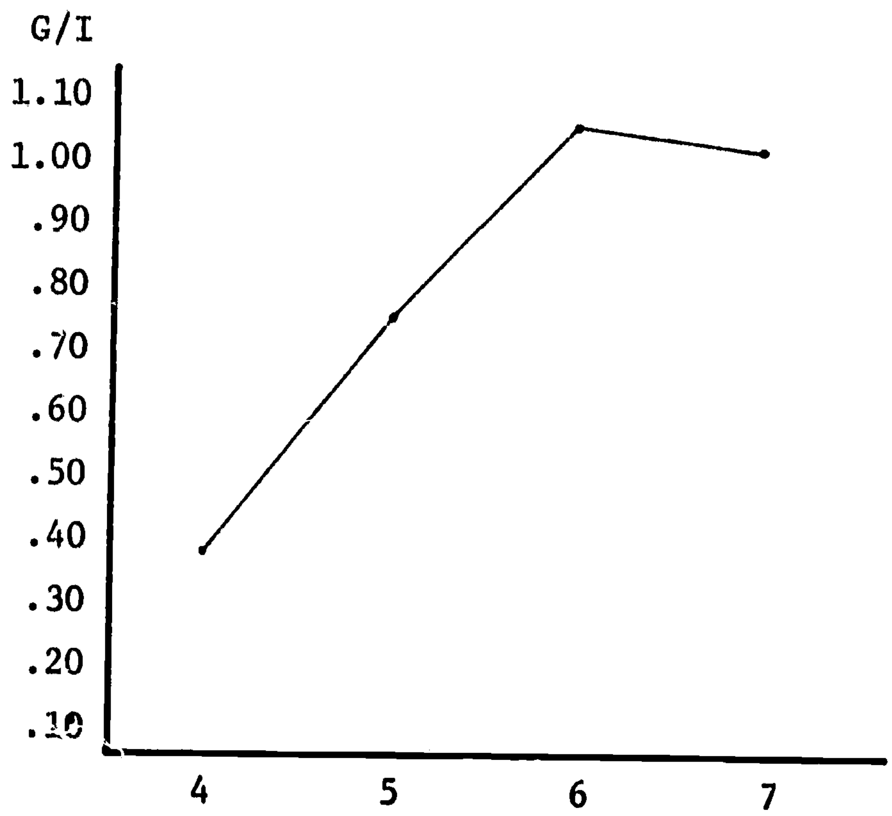


Fig. 4. G/I vs Grade

As was mentioned earlier, the fourth grade's relative inexperience with computation--multiplication particularly, in view of the items--may have unduly influenced their discovery performance. Taking into account this inexperience, the fifth, sixth, and seventh graders' established familiarity with the basic operations, and the sixth graders' inadvertent experience with short-cuts, one might "adjust" Figure 2 and hypothesize a total instances curve, as in Figure 5.

This figure and the Piagetian-proposed transition from a concrete-operational cognitive state to a formal-operational cognitive state (Flavell, 1963) suggested performing post hoc tests to attempt to ascertain whether the observed grade differences--in particular that between grades 5 and 6--were statistically significant.

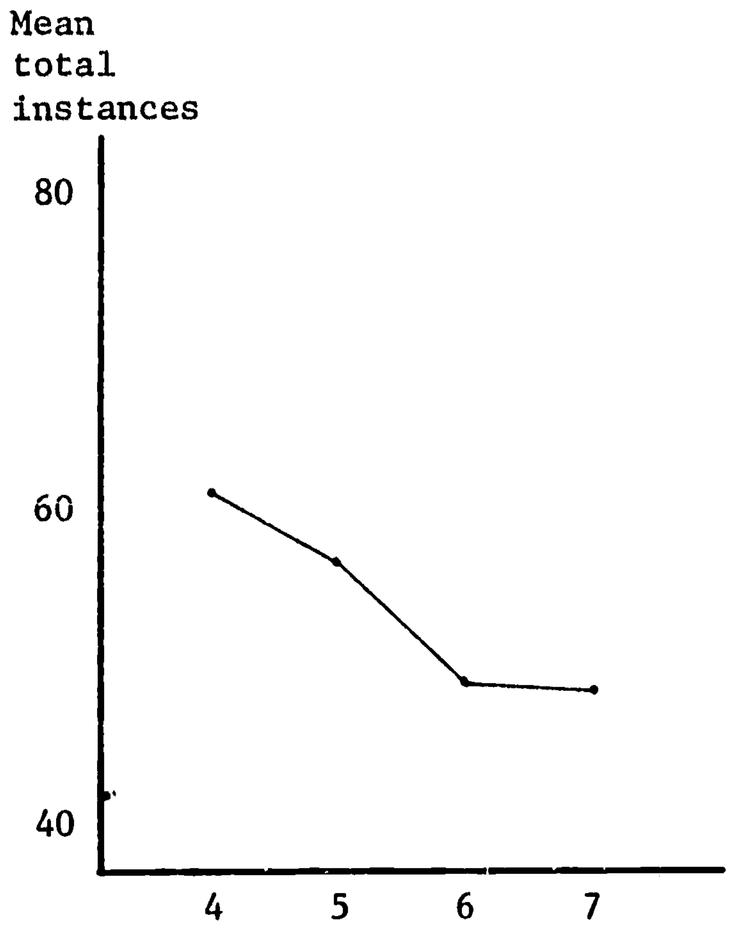


Fig. 5. Hypothetical Mean Total-instances vs Grade

Since there were equal numbers of pupils at each grade level and since the type of contrast of interest was the comparison of only two means, Tukey's method was chosen to test for grade differences (Scheffe, 1959, ch. 3). Tables 14 and 15 give summaries of the post hoc tests and show that, for both total-instances scores and number of generalizations formed, the fourth grade mean was significantly (.05) different from those of the other grades, but the grade 5 - grade 6 difference was not significant.

Table 14

Post Hoc Tests for Differences in Grade Means, Total-instances Scores

Grade	4	5	6
5	-11.0*		
6	-18.8*	-7.8	
7	-17.8*	-6.8	1.0

*significant, .05

$$t = q_{.05;4,48} \sqrt{\frac{MS_{error}}{n}}$$

$$= 3.77 \sqrt{\frac{142.4}{18}} = 10.6$$

Table 15

Post Hoc Tests for Differences in Grade Means, Number of Generalizations Formed

Grade	4	5	6
5	1.7*		
6	2.6*	0.9	
7	2.5*	0.8	-0.1

*significant, .05

$$t = q_{.05;4,48} \sqrt{\frac{MS_{error}}{n}}$$

$$= 3.77 \sqrt{\frac{3.19}{18}} \doteq 1.58$$

Table 16, summarizing the normal approximation for a Mann-Whitney U test (Siegel, 1956) for differences in G/I scores between grade 5 and grade 6, gives a z-value corresponding to a probability of about .08; hence, the hypothesis of no difference in the grade 5 and grade 6 G/I scores approached rejection at the .05 level.

Table 16

Mann-Whitney U Test for Differences in G/I Scores, Grade 5 vs Grade 6

Grade	5	6
Sum of ranks (lowest G/I--rank 1)	288.5	377.5

$$U = n_1 n_2 + \frac{n_i(n_i + 1)}{2} - R_i$$

(where i corresponds to group with greater sum of ranks)

$$U = 18(18) + \frac{18(18 + 1)}{2} - 377.5$$

$$U = 117.5$$

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

$$z = \frac{117.5 - 162}{\sqrt{\frac{18(18)(18 + 18 + 1)}{12}}}$$

$z = -1.41$; $p = .0793$

Problem 1c

If a linear model with independent variables chronological age, IQ, arithmetic achievement scores, and mathematical interests scores and with dependent variables total-instances score or total number of generalizations is postulated, what portion of the variance is accounted for? As was mentioned in Chapter 2, the regression technique utilized a "step-down" analysis. The computer program used for this analysis was due to Finn (1967) and required specifying the order of the independent variables in the analysis; according to Bock,

> . . . Unlike the other tests of multivariate hypotheses, the step-down test is not invariate under rearrangements of the variates. The order in which the variates are eliminated must be determined beforehand. It is convenient to choose this ordering so that the step-down F statistics can be used to judge the partial contribution of successive variates to discrimination between groups. . . . Variates believed important to discrimination should appear earlier in the ordering and more dubious variables later. If the latter make no appreciable contribution, there is some empirical ground for omitting them from further consideration (in Cattell, 1966, p. 828).

Since grade and IQ level seemed to be important variables (see above results from analyses of variance), age and IQ were entered first into the all-grades analysis. However, under the assumption that at a given grade level age would not play so important a role, it was entered later in the analysis within a grade. To assist in ordering the other variables, correlation coefficients (product moment) were calculated, using a University of Wisconsin Computing Center program (Wetterstrand, 1966). These correlation coefficients are recorded in Table 17 and provided the basis for ordering the independent variables after age and IQ.

Some of these correlations should be pointed out. For example, for these pupils IQ is negatively correlated with age. IQ and routine mathematical interests score are negatively correlated (-.24). IQ is most highly correlated with the concepts achievement quotient. Also of interest is the -.65 correlation between creative- and routine-interests scores although this is no doubt due in part to the method of scoring the interests questionnaire.

Table 17

Correlation Coefficients (Decimal Points Omitted)

a. Grades 4-7 Combined

IQ	Cre	Rou	Non	Com	Con	T-in	Tot. Gen.	
-24a	15	-12	-10	-14	-19	-39b	33b	Age
	12	-25a	16	41b	59b	-35b	42b	IQ
		-65b	-44b	04	-10	-14	15	Cre(ative interests)
			-38b	-12	-10	15	-15	Rou(tine interests)
				07	26a	05	-04	Non(math'l int.)
					65b	-48b	50b	Com(putation AQ)
						-34b	43b	Con(cepts AQ)
							-95b	Total instances (T-in)

a--Significantly different from 0 at .05 level

b--Significantly different from 0 at .01 level

Table 17 (continued)

b. By Grade

IQ	Cre	Rou	Com	Con	T-in	T-ge	App	
07	-61	08	-33	-22	-02	-04	a	4Age
-40	42	-14	08	-26	32	-47	-28	5Age
-53	-18	20	-30	-51	25	-28	-39	6Age
-68	-18	36	-58	-63	23	-24	-67	7Age
	09	-19	40	64	-46	-04	a	4IQ
	01	-29	62	68	-49	55	76	5IQ
	31	-40	22	52	-46	53	52	6IQ
	09	-23	55	66	-57	59	69	7IQ
		-47	26	24	-41	55	a	4Creative
		-63	27	-22	-06	03	-14	5Creative
		-70	-20	-27	-47	40	-30	6Creative
		-76	17	08	39	-35	-11	7Creative
			-31	-45	22	-30	a	4Routine
			-20	-01	06	-06	-03	5Routine
			26	29	26	-19	37	6Routine
			-28	-23	-17	14	-05	7Routine
				76	-43	49	a	4Computation AQ
				66	-68	61	69	5Computation AQ
				57	-50	47	42	6Computation AQ
				72	-62	64	70	7Computation AQ
					-50	55	a	4Concepts AQ
					-56	61	85	5Concepts AQ
					-18	33	87	6Concepts AQ
					-58	60	82	7Concepts AQ
						-93	a	4Total-instances
						-94	-54	5Total-instances
						-94	-02	6Total-instances
						-97	-62	7Total-instances
							a	4Total Gens.
							55	5Total Gens.
							21	6Total Gens.
							68	7Total Gens.

a--No Applications AQ for Grade 4.
Correlation coefficients $\geq .40$ in absolute value are significantly different from zero at .05 level.

As a result of the above considerations, independent variables were entered into the all-grades analyses in the following order: age, IQ, computation achievement quotient (AQ), concepts AQ, creative mathematical interests, and routine mathematical interests. For the analysis at each grade, the variables were entered in the order IQ, computation AQ, concepts AQ, creative mathematical interests, routine mathematical interests, age, and applications AQ (when available). Results of these analyses are given in Tables 18-21.

Table 18

Regression Analysis for
Total-instances Scores--Grades 4-7

Variable	Raw regression coefficients	Standardized regression coefficients	Significance*
Age	-6.31	-0.53	.0007
IQ	-0.32	-0.31	.0001
Computation AQ	-33.46	-0.44	.0001
Concepts AQ	0.61	0.01	.8527
Creative interests	-0.26	-0.06	.9623
Routine interests	-0.40	-0.09	.4642

*Probability of type I error for hypothesis that the regression coefficient for the variable equals zero.

$R = 0.72$ $R^2 = 0.52$ Probability of such an R, if R actually equals zero, is less than .0001.

Table 19

Regression Analysis for
Number of Generalizations--Grades 4-7

Variable	Raw regression coefficients	Standardized regression coefficients	Significance*
Age	0.87	0.48	.0045
IQ	0.05	0.33	.0001
Computation AQ	4.10	0.36	.0001
Concepts AQ	0.96	0.12	.4688
Creative interests	0.07	0.11	.7057
Routine interests	0.08	0.11	.3288

*Probability of type I error for hypothesis that the regression coefficient for the variable equals zero.

$R = 0.73$ $R^2 = 0.53$ Probability of such an R, if R actually equals zero, is less than .0001.

Tables 18 and 19 show that for either total-instances scores or number of generalizations, the variables age, IQ, and computation achievement quotient contribute most to the linear model. In either case, the linear model accounts for slightly more than half the variance of the dependent variable.

Table 20

Regression Analyses for
Total-instances Scores, by Grade

	Standardized reg. coeff.							
Variable	4	5	6	7	4*	5*	6*	7*
IQ	-.04	.21	-.60	-.50	.054	.038	.054	.014
Computation AQ	-.20	-.81	-.57	-.59	.235	.024	.066	.070
Concepts AQ	-.46	-.14	-.16	-.15	.618	.559	.084	.732
Creative interests	-.86	-.06	-.53	.54	.185	.775	.053	.005
Routine interests	-.41	-.01	-.35	.10	.614	.876	.555	.944
Age	-.68	.48	-.11	-.49	.033	.079	.709	.030
Applications AQ	--	.11	.60	0	--	.812	.178	.989
R	0.78	0.79	0.84	0.90	.059[a]	.110[a]	.038[a]	.005[a]
R^2	0.61	0.62	0.71	0.82				

*Type I error for hypothesis that the regression coefficient for the variable equals zero.

[a] Probability of such R, even if true R equals zero.

Table 21

Regression Analyses for
Number of Generalizations Formed, by Grade

Variable	Standardized reg. coeff. 4	5	6	7	4*	5*	6*	7*
IQ	.18	-.12	.53	.48	.016	.017	.024	.010
Computation AQ	.27	.58	.42	.53	.163	.106	.089	.057
Concepts AQ	.31	.38	.29	.06	.769	.342	.460	.681
Creative interests	.97	.23	.52	-.47	.029	.972	.102	.008
Routine interests	.36	.08	.30	-.14	.656	.695	.519	.965
Age	.67	-.61	.16	.56	.007	.021	.567	.023
Applications AQ	--	-.22	-.38	.25	--	.583	.446	.394
R	.89	.83	.77	.92	.003[a]	.049[a]	.134[a]	.003[a]
R^2	.80	.69	.60	.84				

*Type I error for hypothesis that the regression coefficient for the variable equals zero.

[a]Probability of such R, even if true R equals zero.

Tables 20 and 21 indicate that IQ is uniformly an important variable in the linear models investigated in the two tables. With the exception of the sixth grade, age--even when relegated to a later position in the list of entering variables--is also important. The signs of the regression coefficients for age in grade 5 and for IQ in grade 5 in the two tables are puzzling. The computation achievement quotient variable also contributes fairly well, except at the fourth grade; again, this is probably because fourth grade computation AQ's do not directly reflect ability with multiplication, an important operation in many of the generalizations. The concepts achievement quotient variable does not play so important a role, perhaps because of its .65 and .59 correlations with computation AQ and IQ, respectively. Concepts AQ does appear to contribute substantially more to the model for total-instances in grade 6. The creative mathematical interests variable behaves, in a broad sense, about the same across grades except for the sharp contrast in grade 5. The creative interests score seems to be more important to the linear model for the total-instances score (except for grade 4); perhaps greater creative mathematical interest is related to offering more hypotheses and would as a result enable the pupil to hit upon the correct generalization in fewer instances.

Problem 1d

For the pupils in the study, what form does a total learning curve take? Of interest here was whether pupils would form generalizations in fewer instances in the course of the short experience of the discovery test. Recall that items were presented in quasi-randomly chosen orders so that each item was the first one encountered by 9 pupils, the second one met by 9 others, the third one given to 9 others, etc. The graph given in Figure 6 is in terms of the total number of instances required; hence, a descent in the graph means better performance. Since the 53 instances improvement from first item to last item represents the improvement over 72 pupils, not much claim for improvement can be made.

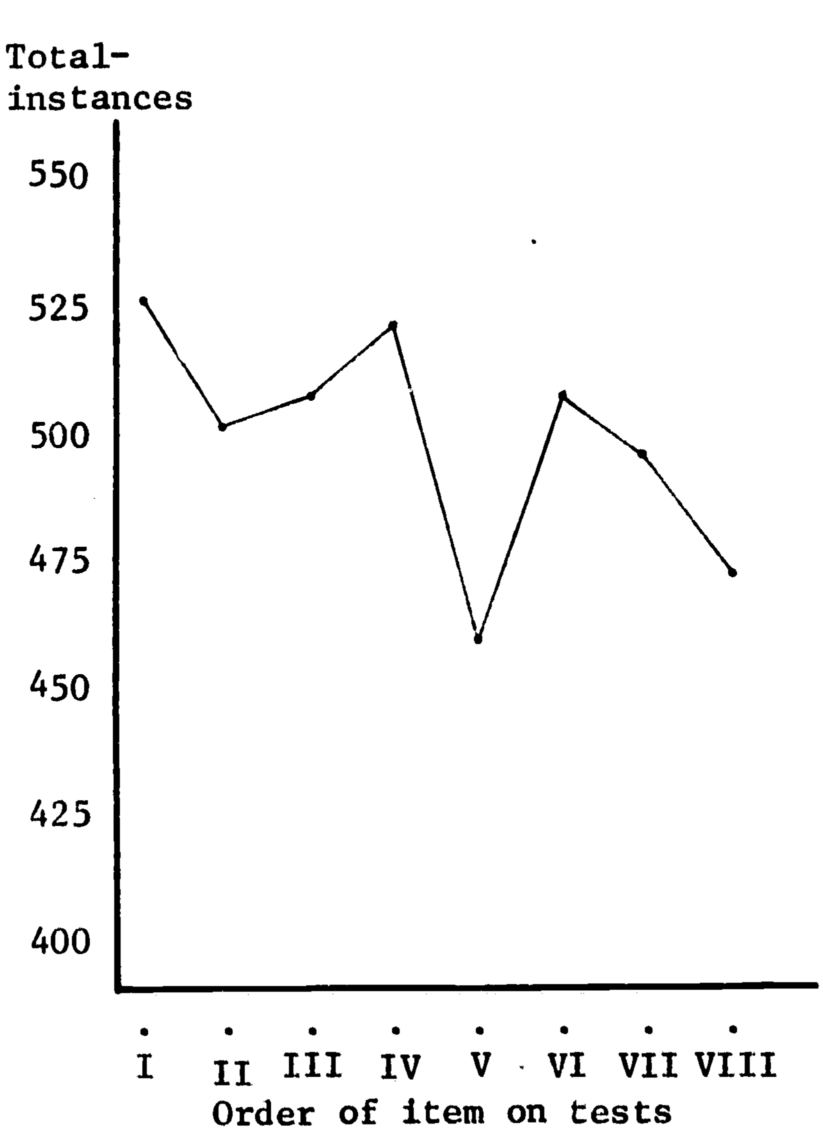


Fig. 6. Total Learning Curve
(72 pupils)

Interpretation of total performance with respect to various dimensions (grade, IQ level, etc.) is not trustworthy since each of the eight test items is not represented an equal number of times at each position in the administration of the test for these dimensions. Nonetheless, Figures 7 and 8 for grades and IQ levels are presented for examination.

The short-cut items as a whole were more difficult than the secret-rule items. Total performances on the two types are plotted in Figure 9. Some improvement does take place with the short-cut items, but secret-rule performance is fairly stable.

Results--The Verbalizing Study

The Problem Restated

Is there a difference in the ability to use numerical generalizations (for which operational evidence of attainment has been given) after having undergone a no-verbalization treatment, a verbalization-by-pupil treatment, or an experimenter-verbalization treatment on those generalizations? Ability to use the generalizations was determined by giving a two-part follow-up test one week after the discovery test. Part I contained instances of the short-cut itmes without mention of the discovery test; Part II contained instances of both the short-cut and the secret-rule items with the secret-rule symbols serving as a direct reminder of the discovery test. Although ostensibly a speed test, no limit was actually placed on the parts of the follow-up test. The three treatment groups were not different with respect to IQ ($F < 1$). Pupil performance on the follow-up test is summarized by treatment in Tables 22-24.

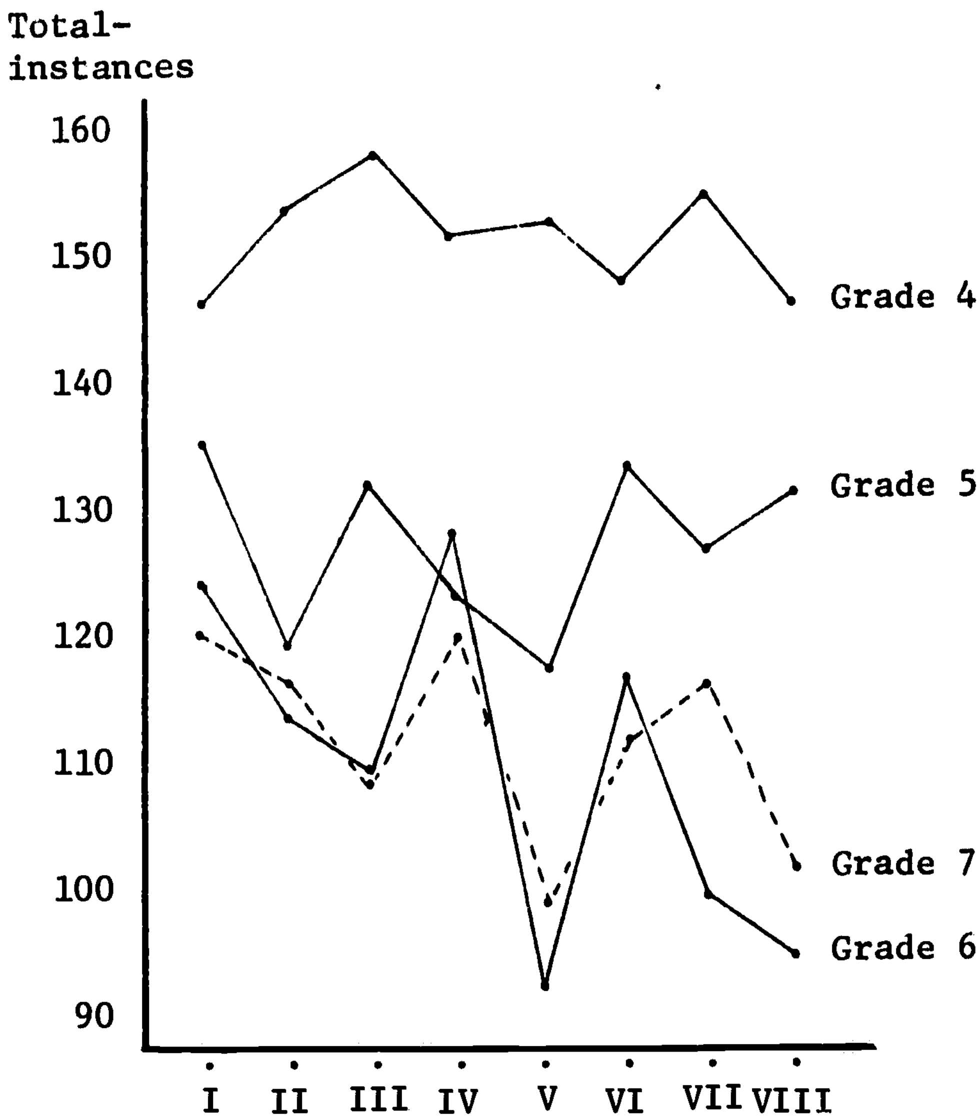


Fig. 7. Total Learning Curves, by Grades

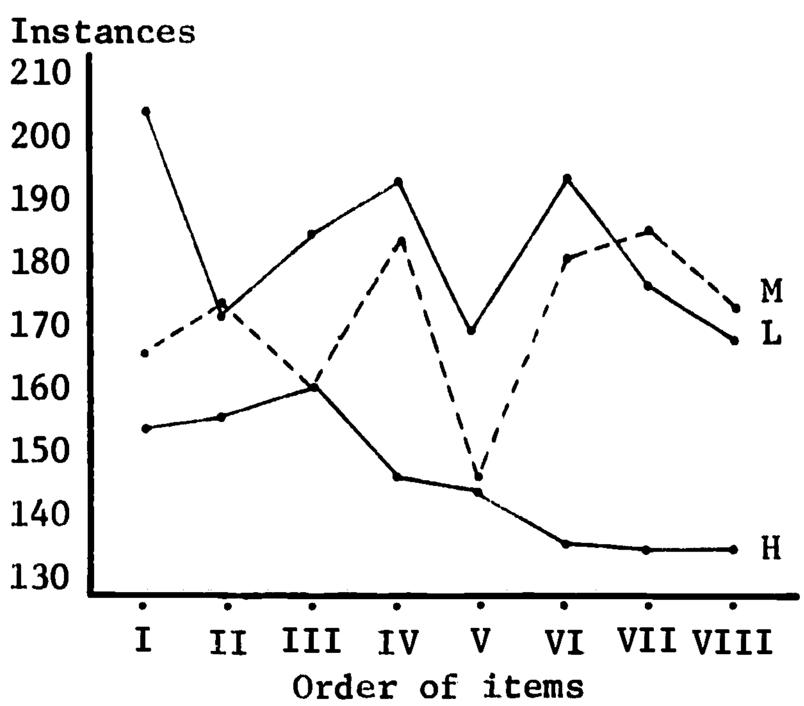


Fig. 8. Learning Curve, by IQ Level

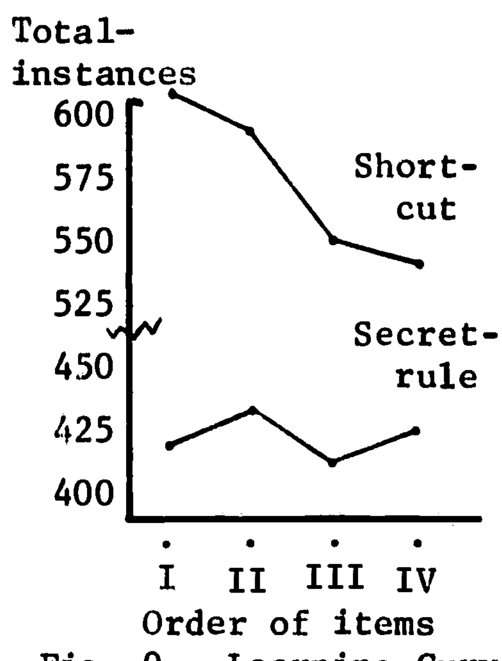


Fig. 9. Learning Curves, Short-cut and Secret-rule Items (all grades, 72 pupils)

Tables 22-24

Follow-up Performance

Key PQR: P--grade; Q--IQ level; R--sex

a-b: Entry for items 1-4 tells whether instance on the follow-up test was correctly (1) or incorrectly (0) completed. a gives performance on Part I, b on Part II.

a-b/c: a = sum of short-cut instances correct on Part I and secret-rule instances correct on Part II
b = sum of short-cut and secret-rule instances correct on Part II
c = number of generalizations formed on discovery test

Short-cut instances in Parts I and II were all based on the same items (1-4) but were different instances of these items.

Table 22

Follow-up Performance, No-verbalization Group

	1	2	3	4	5	6	7	8	Total
4HB	X	X	X	X	0	X	X	X	0/1
4HG	X	X	X	X	0	0	X	1	1/3
4MB	X	X	X	X	0	X	X	0	0/2
4MG	0-0	X	X	X	X	0	X	X	0-0/2
4LB	X	X	0-0	X	0	X	0	X	0-0/3
4LG	X	X	X	X	0	X	X	0	0/2
5HB	0-a	0-a	0-a	X	a	a	a	X	a-a/3
5HG	X	0-0	X	X	0	0	X	0	0-0/4
5MB	X	X	X	0-0	0	X	X	X	0-0/2
5MG	0-1	X	0-0	Z	0	0	X	0	0-1/5
5LB	1-0	X	0-0	X	0	X	0	0	1-0/5
5LG	X	X	X	0-0	0	0	0	1	1-1/5
6HB	X	X	X	X	0	0	0	0	0-0/4
6HG	1-1	1-0	1-0	0-0	0	0	0	0	3-1/8
6MB	0-0	X	X	X	0	0	X	0	0-0/4
6MG	0-0	X	X	X	X	0	X	0	0-0/3
6LB	0-1	X	0-0	X	0	0	X	0	0-1/5
6LG	X	X	0-0	0-0	0	0	0	0	0-0/6
7HB	0-1	X	0-0	0-0	0	0	0	0	0-1/7
7HG	1-1	X	0-0	X	0	0	1	0	2-2/6
7MB	0-0	X	X	0-0	0	0	0	0	0-0/6
7MG	X	0-0	1-0	0-0	0	0	X	0	1-0/6
7LB	X	X	X	X	0	X	X	1	1/2
7LG	X	X	X	X	0	X	X	X	0/1
	3/11	1/4	2/10	0/7	0/21	0/15	1/9	3/18	10/95
	5/10	0/3	0/9	0/7					9/92[a]

[a]Data missing or incomplete.

Table 23

Follow-up Performance, Subject-verbalization Group
(Key on page 72)

	1	2	3	4	5	6	7	8	Total
4HB	0-0	X	X	X	X	0	0	0	0-0/4
4HG	0-0	X	0-0	X	0	0	X	0	0-0/5
4MB	X	X	X	X	X	X	X	X	0/0
4MG	X	X	X	X	0	X	X	0	0/2
4LB	X	X	X	X	1	X	X	1	2/2
4LG	X	X	X	X	0	X	0	X	0/2
5HB	0-0	X	0-0	X	0	0	X	0	0-0/5
5HG	X	X	0-0	0-1	X	0	X	0	0-1/4
5MB	1-1	1-1	X	0-0	0	0	0	0	2-2/7
5MG	X	X	X	X	0	0	X	0	0/3
5LB	0-0	X	0-0	X	0	0	X	X	0-0/4
5LG	X	X	X	X	1	X	X	X	1/1
6HB	0-0	0-0	0-0	X	0	0	0	0	0-0/7
6HG	0-0	X	0-1	0-1	1	0	0	0	1-3/7
6MB	0-0	X	0-0	X	X	0	X	0	0-0/4
6MG	0-0	0-0	0-1	X	0	0	X	0	0-1/6
6LB	X	X	X	X	0	X	X	X	0/1
6LG	1-0	X	0-0	0-0	0	0	0	0	1-0/7
7HB	1-1	0-0	0-0	0-0	0	0	0	1	2-2/8
7HG	X	0-0	0-0	X	0	0	1	0	1-1/6
7MB	0-0	X	0-0	X	0	0	X	1	1-1/5
7MG	X	X	X	X	0	0	0	1	1/4
7LB	X	X	0-0	X	0	0	X	0	0-0/4
7LG	X	X	0-0	X	0	0	X	X	0-0/3
	3/12	1/5	0/14	0/5					12/101
	2/12	1/5	2/14	2/5	3/20	0/18	1/9	4/18	15/101

Table 24

Follow-up Performance,
Experimenter-verbalization Group
(Key on page 72)

	1	2	3	4	5	6	7	8	Total
4HB	X	0-0	X	X	0	0	X	X	0/3
4HG	0-0	0-0	0-0	X	0	0	X	1	1/6
4MB	0-0	X	0-0	X	0	0	X	0	0/5
4MG	X	X	X	X	0	0	X	X	0/2
4LB	X	X	X	X	X	X	X	X	0/0
4LG	X	X	X	X	0	X	X	0	0/2
5HB	1-1	0-1	0-0	X	0	0	X	0	1-2/6
5HG	0-0	X	0-0	X	0	0	0	0	0/6
5MB	0-0	0-0	0-0	X	0	0	1	1	2/7
5MG	0-1	X	0-0	X	0	0	X	1	1-2/5
5LB	0-0	X	0-0	X	X	X	X	X	0/2
5LG	X	X	X	X	1	X	X	X	1/1
6HB	X	0-0	X	0-0	0	0	X	X	0/4
6HG	0-1	0-0	0-0	0-1	1	0	0	0	1-3/8
6MB	X	X	0-0	X	0	0	0	0	0/5
6MG	0-0	1-1	0-0	0-0	0	0	1	0	2-2/8
6LB	X	0-0	0-0	X	0	0	X	0	0/5
6LG	X	X	X	X	a	X	X	X	a
7HB	1-1	0-1	0-0	0-1	0	0	0	0	1-3/8
7HG	1-1	0-0	X	X	0	0	X	0	1/5
7MB	X	X	X	X	X	0	X	1	1/2
7MG	X	X	0-0	X	0	0	0	0	0/5
7LB	0-0	X	0-0	0-0	0	0	0	0	0/7
7LG	0-0	X	0-0	X	0	0	0	0	0/6
	3/13	1/10	0/15	0/5	2/20	0/19	2/9	4/17	12/108
	5/13	3/10	0/15	2/5					18/108

[a]Data missing.

Examination of the entries for items generalized on the discovery test shows that they are predominately zeros. The question was whether to dignify these results with the intended analysis of variance since even if statistically significant treatment differences were detected, they could scarcely be practically significant.

However, merely to see whether any treatment differences appeared to be indicated, three analyses of variance were performed with pupils all of whom had generalized on the items examined in the particular analysis. Criteria used to arrive at subsets of the eight discovery test items to be considered in an analysis were as follows:

(1) the subset of items was to contain more than two items,

(2) each treatment was to be represented by at least 50% of the pupils (or within 1 of 50%) under that treatment who formed at least as many generalizations as in the subset being considered, and

(3) there were to be at least three pupils representing each treatment.

Eleven subsets of items met these criteria (135678, 145678, 345678, 15678, 13568, 5678, 3568, 1568, 158, 358, 568). Note however the great deal of overlap among the items in these subsets. Furthermore, the pupils involved for each of these subsets of items were much the same. The three ANOVA's performed (on item subsets 145678, 13568, and 1568) all gave F's less than 1. Hence, no case at all can be made for any treatment differences.

Product-moment correlations for the number of generalizations retained, the portion of generalizations attained on the discovery test which were correct on the follow-up test, and several of the other variables in the study are given in Table 25. The correlations with the portion retained are, strictly speaking, not appropriate since the portion retained is based on different numbers of items.

Table 25

Correlation Coefficients for Follow-up Data

	Age	IQ	Cre	Rou	T-in	Com	Con	Fnum
Number retained on follow-up (Fnum)	13	28*	15	-13	-42*	06	14	
Portion retained	02	-03	12	-08	03	-25*	-17	69*

*Significantly different from 0 at .05 level.

Miscellany--Interviewer Differences

After the first day of administering the discovery test, casual examination of the data indicated that the pupils examined by one interviewer seemed to be performing uniformly better than those of the other interviewers. Consequently, an audio recording of a sample of his interviews was made and examined. The only departure from the standard protocol was that this interviewer on occasion permitted more time (up to 15 seconds more) before giving the next instance and allowed less time (as much as 15 seconds less) before giving the correct answer to an instance. Schematically, Standard protocol: Instance........answer........next instance.

Interviewer 1: Instance..answer..............next instance.

As a result, the total time exposure was virtually the same as with the standard protocol but distributed differently with respect to the exposing of instances and the giving of answers. Although subsequent analysis of the performances of the interviewers' pupils showed no significant differences among interviewers at the .05 level (total-instances $F = 2.16$, number of generalizations $F < 1$), this possible role of the distribution of study time was interesting. Performances of the pupils of different interviewers are summarized in Table 26.

Table 26

Discovery Test Performance, By Grade and Interviewer

Interviewer	Total instances 4	5	6	7	Total	Total generalizations formed 4	5	6	7	Total
1	387	326	263	274	1250	19	27	33	35	114
2	417	344	339	307	1407	13	27	26	29	95
3	409	346	282	311	1348	14	24	34	27	99

The observed F-value of 2.16 for the total-instances scores (corresponding to a significance level of about .12) suggests that getting more instances before the pupil perhaps helps him by giving him more clues sooner.

Chapter IV

SUMMARY AND CONCLUSIONS

Summary

Purpose

The purpose of this study was two-fold: (1) to explore the ability of boys and girls of different IQ levels and in grades 4 through 7 to give evidence of having discovered a short-cut or rule in selected numerical situations, and (2) to investigate the influence on the retention of the ability to use these generalizations after verbalizing, not verbalizing, or listening to verbalizations of, the generalizations.

Method

Eighteen pupils from each of grades 4-7 in the public school of a small south-central Wisconsin town were randomly chosen and given an individually administered discovery test consisting of the stimulus portions of instances of generalizations. For the exploratory part of the study, the number of instances required before the pupil gave correct responses, as well as the number of generalizations apparently formed, were recorded and analyzed (a) by a grade-by-IQ level-by-sex analysis of variance and (b) with respect to a linear model with independent variables age, IQ, arithmetic achievement, and mathematical interests. Performance on a follow-up test (one week later) based on instances of the generalizations on the discovery test was to provide information for the influence-of-verbalizing study.

Conclusions

It must be mentioned, of course, that any results of the study cannot be generalized either beyond the population sampled in the one school system of the study or beyond the content and style of the discovery items. The conclusions will be examined in terms of the specific problems stated in chapter I.

The Status Study--Conclusions

Problem 1a

In selected numerical situations, what is the mean number of instances that must be presented before boys and girls of grades 4, 5, 6, and 7 and of high, middle, and low intelligence levels show operational evidence of having attained the generalizations? Tables 6, 7, and 9 of chapter III indicate that all such groups can "form" some generalizations with the items of the discovery test, with the number of instances required per generalization ranging from 3 to 6 when the pupils did generalize (*i.e.*, not including those items for which they did not generalize). Figure 2 of chapter III pictures the not unexpected improvement in performance with increasing IQ level. Figure 2 also shows the interesting pattern of improvement from grade 4 through grade 6, with the slight decrease in performance from grade 6 to grade 7. Whether the observed differences were statistically significant is discussed in connection with problem 1b.

Problem 1b

Are there significant grade level differences, IQ level differences, sex differences, grade-IQ level interactions, IQ-sex interactions, grade-sex interactions, or grade-IQ level-sex interaction for the total-instances scores, or for the number of generalizations formed, in these situations? There were significant (.01) differences among grade levels and among IQ levels, but not between sexes. No interactions were significant at the .05 level.

Post hoc tests were used to attempt to isolate the grade level differences: grade 4 was significantly (.05) different from each of grades 5, 6, and 7, but no other significant (.05) grade differences were detected. Similar results were obtained for grade level differences on the G/I score (= 10 x number of generalizations/total-instances score).

Problem 1c

If one postulates a linear model with independent variables chronological age, IQ, arithmetic achievement scores, and mathematical interests scores and with dependent variable total-instances score or the number of generalizations formed, what portion of the variance does the model account for? Tables 18 and 19 of chapter III record the results for grades 4-7 combined: slightly more than half the variance is accounted for in either the total-instances model or the number of generalizations model. Age, IQ, and computation achievement quotient contribute significantly to accounting for the variance, with concepts achievement quotient, creative mathematical interests, and routine mathematical interests adding little. The type of regression analysis performed might yield different regression coefficients from those of the study if a different order of entering the variables were used (Cattell, 1966, p. 828).

The results of the separate analyses for each grade are not so easy to summarize (Tables 20 and 21 in chapter III). The models for these grades accounted for from .60 to .84 of the variance. IQ was still an important part of the models. Age, entered in a later position than in the all-grades analysis, was also important except at grade 6. Computation achievement quotient also contributed well when the type of computation it was based on included multiplication. The creative mathematical interests score seemed important in grades 4, 6, and 7, but not in grade 5. The concepts achievement quotient seemed important only once (grade 6, total-instances model); routine mathematical interests and the application achievement quotient added little to any model. Again, a different order of entering the independent variables might have given different results.

Problem 1d

For pupils in the study, what form does a total learning curve take? Figure 6 (chapter III) shows that over all 72 pupils only 53 fewer instances were required for last items than for first items, as received on the discovery test. Figure 7 in chapter III indicates that, in terms of last item performance *vs* first item performance, grades 6 and 7 profited more from the practice than did grades 4 and 5; these data may be specious however since there may have been unequal interaction between items and ability levels, grade, or sex (each grade level did receive the same number of short-cut and secret-rule items). Furthermore, the irregular nature of the graphs in Figures 6 and 7 would make one reluctant to claim unequivocally that any indicated improvement was reliable. However, Figure 9, which shows performance on the short-cut and secret-rule items separately, suggests that performance on secret-rule items stabilizes quickly but that performance on short-cut items improves noticeably. In either case, of course, the small number of items makes any claim risky.

The Verbalizing Study-Conclusions

The problem

Is there a difference in the ability to use numerical generalizations for which operational evidence of attainment has been given, after having undergone a no-verbalization treatment, a verbalization-by-pupil treatment, or an experimenter-verbalization treatment on these generalizations? As was noted in chapter III, pupils under all treatments gave scanty evidence of having retained the ability to use the generalizations. Hence, no statement can be made about the influence of the three types of treatments on such retention.

Implications of the Studies

The main implication of the status study is that virtually all pupils of grades 4-7 in the sampled population can make some "little-d" discoveries. Part of Davis' work (1964) has shown that groups of underprivileged children can succeed at some of his tasks. More to the point, however, Friedlander warns,

> Though the matter of non-participation by significant and possibly substantial portions of a class group is primarily an issue of classroom management rather than a psychological problem, the topic deserves some mention here in view of the prospect that only some members of a class are likely to be fruitful discoverers. It is a particularly insidious danger that a teacher might fall into the pattern of rushing buoyantly along on sequences of interlocking cascades of thought and inquiry with just the handful of students who are able to fly along on the same journey. The classroom observer sees what the teacher seldom seems to notice--that only a few students appear to be attentive and involved when the teacher is working with his greatest enthusiasm. What may be an exhilarating experience for the teacher and a few responsive students might also be a bleak disservice to the majority of the class. . . . The real danger here lies in the risk of overestimating the advantage of the benefits for the few while underestimating the cost in lost communication with the many (1965, p. 33).

The present study indicates that lower ability individuals can make some discoveries (at least of the type on the discovery test), albeit at a slower rate. Hence, if the claimed intellectual and motivational benefits of discovery do accrue from such tasks and are to come to all in a classroom, individuals or homogeneous groups could be given some tasks of the sort in this study with the expectation of all experiencing some success in discovering.

Pupils' performance with grade improves through grade 6 and then seems to reach a plateau. Hence, other objectives being equal (motivation, interest, drill, etc.), the most efficient use of such discovery tasks will appear after grade 5 (under the conditions of the study).

The study also indicates the number of instances required by pupils of these grades to form a generalization of the type studied and under time strictures like those of the study. As very rough guidelines, since performance varies so much with item and pupil, at least three instances of an item such as those studied should be presented. On the other hand, presenting more than six to eight instances is a waste of time for most pupils. An additional consideration, if the item is a short-cut item, seems to be the quantity of previous experience with such items since performance on them seems to improve noticeably with experience. Secret-rule items, on the other hand, are easy to construct and yield to fairly quick discoveries. Hence, they might serve as warm-up or motivational activities.

As might be expected with items of this type, pupils with higher computation achievement do better. If success at such discovery tasks is desired, then computational skills should be brought to as high a level as possible.

Further Study

The Romberg-DeVault paradigm, Figure 10, (from 1967, p. 96) provides a framework for possible further studies. It is recognized that suggested studies most often intersect more than one area.

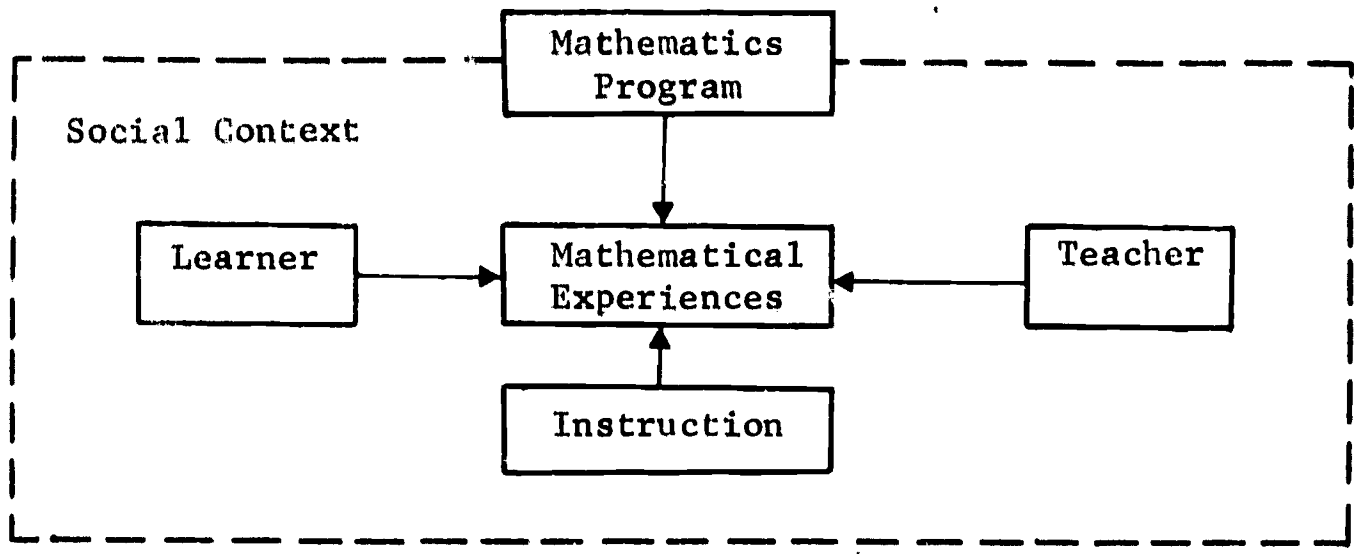


Fig. 10. Mathematical Curriculum Research Model

Instruction

1. How would pupils perform on the discovery test items after moderate or extensive practice with items (a) of the same type or (b) of dissimilar types?

2. What test protocol, if any, gives the best results in terms of total-instances, say? Is it best with all pupils? (These questions are suggested by the results obtained by the one interviewer in the present study who used different time intervals between instance-presentation and answer-giving.) Does provision of a "target task" (Wills, 1967) enhance performance?

3. Does requiring the pupil to verbalize during these discovery tasks, or telling the pupil he will be required to verbalize his discovery, influence his test or follow-up performance? Gagné and Smith (1962) asked this question for a towers-of-Hanoi task and reported that the former enhanced task performance but the latter did not.

Mathematics content

4. How does performance on non-numerical items--e.g., geometric items--compare? Whether pupil interest or motivation is greater with one type of item would also be of interest.

5. Can a testing procedure be designed so that, with a carefully selected group of secret-rule items of different difficulties, one can distinguish different pupil strategies of discovery? Pupils in the pilot study particularly seemed to fall into two categories: between-item scanners who searched the stimulus column and/or the response column vertically, and within-item scanners who apparently hypothesized on the basis of the most recent instance, with or without testing the hypotheses on earlier instances.

Teacher

6. Do pupils' performances on a discovery task differ significantly when their teachers are autocratic as opposed to permissive? Are certain personality characteristics more common to teachers who are successful users of discovery?

Learner

7. Would the results of this study be the same for populations different from that of the study? Would the results have been different if the testing had been done in the second semester instead of the first?

8. Does the apparent plateau in performance at grades 6 and 7 continue into grades 8 and 9? Is the grade 4-5-6 growth due only to greater arithmetic experience?

9. Is transfer, as Haslerud suggests (1958), anticipative rather than perseverative? The design of the present study was predicated mainly on perseverative retention and did not result in any worthwhile retention by the subjects. Whether a change in the test protocol could improve this is moot; an interval of one week may be too long for retention of this sort of item with the amount of practice and study allowed.

10. Hence, testing (perseverative) retention much sooner--e.g., testing afternoon retention after morning discovery--might be in order. Changing the design to exploit anticipative retention, if possible, would also be interesting.

11. How do creative mathematical interests relate to discovery performance? The correlation coefficients and regression analyses within grades (except for grade 5) indicate some relation but invite further investigation.

12. Some pupils reflect lengthily before responding; others respond impulsively. What effect does this difference have on performance and attitude with respect to discovery tasks? Kagan and others (1964) have observed these reflection/impulsivity categories, and Kagan suggests that discovery tasks place the impulsive answerer at a disadvantage since he is wrong more frequently and thus experiences more negative reinforcement (Shulman and Keislar, p. 161).

Improvements

From this vantage point, there are several weaknesses of the present study which might be remedied to make further studies of this sort more informative:

a. The number of subjects could be increased to give greater certainty to the results. Sampling from a larger population would add to the generalizability.

b. A greater number and variety of tasks could be used to see whether the results of the present study are generalizable to other tasks.

c. Additional independent variables--for example, a creative ability test score or scores for various other cognitive factors--could be added to the regression equations to attempt to account for more of the variance. The type of task--numerical, geometrical, verbal--might suggest appropriate independent variables.

d. A greater amount of time and emphasis could be given to the verbalizing treatments to attempt to enhance retention.

e. Additional verbalizing treatments such as having the pupil write his generalization or read a statement of the generalization might give interesting results.

f. A smaller time interval between the discovery test and the follow-up test should also increase retention performance. An appropriate time interval should be determined by pilot testing.

g. Even with an interval of one week, altering the form of the follow-up test might give more meaningful data. For example, two or three completed instances of an item could be given before an incomplete test instance.

In conclusion, this study has given rough guidelines for expected performance by pupils in grades 4-7 on some numerical generalizing tasks. The interesting question of the influence of verbalizing mode on retention was unanswered and invites further study.

Appendix I

PUPIL DATA

Pupil Data

Codes: ID--PQRS: P = Grade
Q = IQ level: High, Middle, or Low
R = Boy or Girl
S = No-verbalizing, Subject-verbalizing, or eXperimenter-verbalizing treatment

Mathematical Interests Questionnaire scores
C = creative
R = routine
N = non-mathematical

COMPAQ: Computation achievement quotient

CONCAQ: Concepts achievement quotient

APPAQ: Applications achievement quotient

ID	AGE	IQ	C	R	N	COMPAQ	CONCAQ	APPAQ
4HBN	10.5	121	13.5	22.0	18.5	0.67	0.49	
4HBS	9.2	146	19.0	22.0	13.0	1.18	1.82	
4HBX	9.3	123	22.4	14.3	22.4	0.95	1.26	
4HGN	9.5	120	18.0	17.0	19.0	1.13	1.21	
4HGS	10.0	122	20.0	16.0	18.0	1.03	1.44	
4HGX	10.1	138	20.0	16.0	18.0	1.15	1.62	
4MBN	10.2	113	17.0	18.0	19.0	1.00	1.10	
4MBS	9.8	110	15.0	23.0	16.0	0.87	0.80	
4MBX	9.8	111	20.0	22.0	12.0	1.03	0.80	
4MGN	9.2	113	21.0	18.0	15.0	1.15	0.97	
4MGS	9.6	111	21.0	21.0	12.0	0.95	0.87	
4MGX	9.8	110	19.5	17.0	17.5	1.00	1.08	
4LBN	9.5	97	20.0	20.0	14.0	1.10	1.15	
4LBS	9.3	98	23.0	17.0	14.0	0.74	0.82	
4LBX	9.9	101	14.0	18.0	22.0	1.05	1.31	
4LGN	9.7	99	18.0	19.0	17.0	0.97	0.92	
4LGS	9.4	100	18.0	22.0	14.0	0.97	0.97	
4LGX	9.8	100	17.0	23.0	14.0	0.77	0.77	
5HBN	10.4	138	20.0	13.0	21.0	1.37	1.55	1.33
5HBS	11.0	125	21.0	15.0	18.0	1.27	1.25	1.47
5HBX	10.3	129	17.0	19.0	18.0	1.27	1.25	1.33
5HGN	10.7	117	22.0	15.0	17.0	1.00	0.73	0.78
5HGS	10.6	123	18.0	21.0	15.0	1.14	1.25	1.33
5HGX	11.0	117	21.0	15.0	18.0	1.18	1.29	1.04

ID	AGE	IQ	C	R	N	COMPAQ	CONCAQ	APPAQ
5MBN	10.5	107	15.0	20.0	19.0	0.90	0.84	0.90
5MBS	10.8	113	17.0	20.0	17.0	1.57	1.45	1.41
5MBX	10.3	108	17.0	21.0	16.0	0.90	1.20	1.00
5MGN	11.1	106	22.0	21.0	11.0	1.22	0.67	0.86
5MGS	11.0	110	23.0	15.0	16.0	1.18	1.06	1.12
5MGX	10.3	113	12.0	22.0	20.0	1.02	1.29	1.08
5LBN	10.5	101	17.0	20.0	17.0	0.90	1.12	1.18
5LBS	10.6	94	22.0	19.0	13.0	1.08	1.02	0.94
5LBX	11.8	98	19.0	20.0	15.0	1.20	1.20	0.94
5LGN	10.8	95	22.0	15.0	17.0	1.06	0.73	0.73
5LGS	10.9	98	18.0	20.0	16.0	0.76	0.80	0.82
5LGX	11.1	86	17.0	17.0	20.0	0.67	0.55	0.69
6HBN	11.3	116	15.0	17.0	22.0	1.32	1.32	1.41
6HBS	11.0	123	24.0	12.0	18.0	1.00	0.95	0.75
6HBX	11.1	127	16.0	17.0	21.0	0.98	1.29	1.54
6HGN	11.6	127	21.0	20.0	13.0	1.10	1.44	1.63
6HGS	11.7	134	22.0	14.0	18.0	0.93	1.10	1.36
6HGX	11.6	120	20.0	16.0	18.0	1.00	1.36	1.36
6MBN	12.1	110	21.0	18.0	15.0	1.20	1.19	1.46
6MBS	11.3	112	20.5	18.5	15.0	0.93	1.00	1.15
6MBX	11.5	113	11.0	23.0	20.0	1.12	1.59	1.88
6MGN	11.8	111	13.0	22.0	19.0	1.12	1.29	1.31
6MGS	12.1	114	19.0	14.0	21.0	0.98	1.12	1.00
6MGX	11.9	107	21.0	20.0	13.0	1.07	1.03	1.12
6LBN	12.5	96	18.0	18.0	18.0	1.05	0.83	0.64
6LBS	12.4	85	15.0	19.0	20.0	0.69	0.88	0.78
6LBX	11.3	96	22.0	16.0	16.0	1.15	1.24	1.10
6LGN	11.3	92	16.0	22.0	16.0	1.25	1.29	1.20
6LGS	12.1	93	13.0	21.0	20.0	1.32	1.03	1.20
6LGX	12.2	77	18.0	18.5	17.5	0.61	0.73	0.92
7HBN	12.5	126	17.0	25.0	12.0	1.07	1.23	1.16
7HBS	12.2	126	17.0	18.0	19.0	1.22	1.16	1.46
7HBX	12.3	121	24.0	14.0	16.0	1.25	1.36	1.39
7HGN	13.0	118	24.0	15.0	15.0	0.99	1.10	0.96
7HGS	12.9	118	20.0	20.0	14.0	0.86	0.75	0.88
7HGX	12.2	130	25.0	11.0	18.0	0.96	1.23	0.96
7MBN	12.6	111	18.0	16.0	20.0	1.07	1.16	1.32
7MBS	12.7	110	21.0	20.0	13.0	0.91	1.23	1.25
7MBX	12.3	112	25.0	15.0	14.0	0.87	0.71	0.88
7MGN	12.4	111	27.0	13.0	14.0	1.42	1.16	1.03
7MGS	12.2	105	19.0	17.0	18.0	0.87	1.10	1.03
7MGX	13.2	110	22.0	15.0	17.0	0.96	0.94	0.83
7LBN	13.0	100	21.0	21.0	12.0	0.78	0.78	0.74
7LBS	13.5	85	22.0	18.0	14.0	0.67	0.67	0.61
7LBX	12.9	96	15.0	23.0	14.0	0.99	0.88	0.91
7LGN	13.0	83	24.0	17.0	13.0	0.90	0.88	0.67
7LGS	13.0	97	19.5	16.0	18.5	0.78	0.75	0.39
7LGX	13.5	97	17.0	22.0	15.0	0.86	0.86	0.74

Appendix II

THE MATHEMATICAL INTERESTS QUESTIONNAIRE

The Instructions

The Questionnaire

Scoring

ADMINISTRATION OF THE MATHEMATICAL INTERESTS QUESTIONNAIRE

(1) Distribute the questionnaires, telling the students to wait for instructions.

(2) Have the students carefully tear off the cover page, exposing the Sample page. Have each student write his first name, then his last name, in the blank on the Sample page.

(3) Be sure to read all the material in quotes to your class. Add any informal remarks which you feel are natural.

(4) Read to the class: "The Sample page is here so we can learn how to fill in the circles and boxes. Read the first three sentences only, those in Sample Group 1 . . ."

(Wait. Feel free to read the sentences aloud or to help students with the reading at any time during the questionnaire.)

"Decide which of these three things you would like to do most, which you like second best, and which you like third best."

(Wait.)

"Now we are going to fill in the circles. Put an 'A' in the circle by the activity you like best in these three sentences. The boxes will be used later. . . .For the same three sentences, put a 'B' in the circle by the activity you like second best. . . .Now put a 'C' in the circle by your third choice in these three sentences."

(5) Read to the class: "Now look at Sample Group 2. Read these three sentences and decide which activity you like best, second best, third best."

(Wait and/or help with the reading.)

"Now put an 'A' in the circle by the activity you like best in these three sentences; a 'B' by your second choice; and a 'C' by your third choice."

(Allow time.)

Permission was granted by SMSG to use the items in this test. The format was developed at The Wisconsin Research and Development Center for Cognitive Learning, pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare under the provisions of the Cooperative Research Program. Center No. C-03/ Contract OE 5-10-154

(6) Read to the class: "Now you try Sample Group 3. Read and decide on your choices; then put 'A' in the circle by your first choice, 'B' by your second choice, and 'C' by your third choice."

(Feel free to help the students with reading or marking problems at any time.)

(7) Read to the class: "Now we are ready to fill in the boxes on the Sample page. Read again the three sentences you put A's by...just the ones you put A's by. Decide which of these three you like best, second best, third best.

(Wait.)

Put a '1' in the box by your first choice, a '2' in the box by your second choice, and a '3' in the box by your third choice. Choose just from the A-sentences on this page right now."

(Allow time, giving help as needed.)

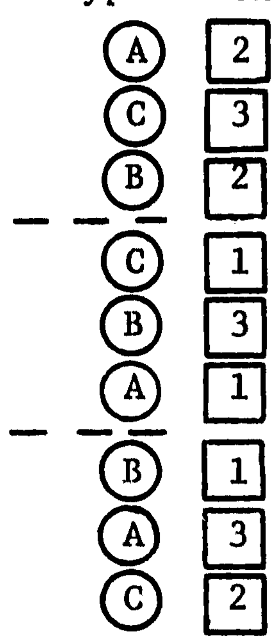
"Now look at your B-sentences, your three B-sentences. Decide which of these three you like best, second best, third best. Then put a '1' in the box by your first choice, a '2' in the box by your second choice, and a '3' in the box by your third choice. Choose just from the B-sentences right now."

(Allow time, giving help as needed.)

"Now look at your three C-sentences and fill the boxes for them."

(Again allow time, helping as needed.)

At this stage, a typical sheet might look like this:



(8) Read to the class: "Are there any questions?" (Answer if so.) "There are 3 more pages in this questionnaire;

we'll do one page at a time. If you need help in reading or filling out, please raise your hand. Notice the STOP at the end of each page."

(Again, feel free at any time during the questionnaire to help in reading or marking. You may read the sentences in each group aloud if you wish.)

(9) Read to the class: "Now turn to page 1, right after the Sample page. Fill in the circles in Group 1 with A, B, and C as you did on the Sample page."

(Wait.)

"Do the same with Group 2."

(Wait.)

"Do the same with Group 3."

(Wait.)

"Now fill in the boxes as we did on the Sample page. Look at just your 3 A-sentences first and rank them 1, 2, 3." (Explain if necessary.)

(Do the same with the B-sentences and then the C-sentences on page 1. Have your students wait until everyone has finished the page.)

(10) Then do page 2 in the same way as in (9).

(11) Finally, do page 3 in the same way.

(12) Have the students check to see that the sheets in the questionnaire are securely fastened together before they hand them in.

MATHEMATICAL

INTERESTS

QUESTIONNAIRE

Permission was granted by SMSG to use items in this test. The test format was developed at The Wisconsin Research and Development Center for Cognitive Learning pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program. Center No. C-03/Contract OE 5-10-154

Sample page

Name (first and last)______________________________

Sample Group 1

○ ☐ Shop at the grocery store.

○ ☐ Shop at the clothing store.

○ ☐ Shop at the drugstore.

Sample Group 2

○ ☐ Ride a bicycle.

○ ☐ Go on a hike.

○ ☐ Run in a race.

Sample Group 3

○ ☐ Read about cowboys.

○ ☐ Go to a cowboy movie.

○ ☐ Watch a cowboy program on TV.

STOP

Page 1

(The type of activity in each group is indicated in parentheses.)

Group 1 (non-math)	◯	☐	Study units of measurement used in Europe.
	◯	☐	Find out what jobs require arithmetic.
	◯	☐	Learn when the number zero was first used.
Group 2 (routine)	◯	☐	Use a map to figure the distances between cities.
	◯	☐	Multiply 9863 by 7215.
	◯	☐	Write the Roman numerals from one to a hundred.
Group 3 (creative)	◯	☐	Learn about numbers less than zero.
	◯	☐	Try to find a fraction with a value between $\frac{8}{13}$ and $\frac{9}{14}$.
	◯	☐	Study a new way to do long division.

STOP

Page 2

Group 4 (routine)	◯	☐	Count the spokes on a bicycle wheel.
	◯	☐	Divide 58236 by 34.
	◯	☐	Mark off a sheet of paper into squares.
Group 5 (creative)	◯	☐	Solve puzzles about numbers.
	◯	☐	Find a way to put 8 dots in 4 rows with 3 in each row.
	◯	☐	Make a machine to average test grades.
Group 6 (non-math)	◯	☐	Choose the best of several designs made with circles.
	◯	☐	Watch a man using an adding machine.
	◯	☐	Learn how the Chinese write numerals.

STOP

Group		
Group 7 (creative)	○ □	Try to cut a triangle into pieces which can form a square.
	○ □	Decide what it means to "forget to forget to forget."
	○ □	Find a short way to add all whole numbers from 1 to 100.
Group 8 (non-math)	○ □	Learn the name of 1,000,000,000,000,000,000.
	○ □	Collect pictures of geometric designs.
	○ □	Study the dates on coins.
Group 9 (routine)	○ □	Find out the number of pennies in a cupful.
	○ □	Count by fives to a thousand.
	○ □	Correct some arithmetic tests.

Scoring

Creative: Sum of ranks in Groups 3, 5, and 7.

Routine: Sum of ranks in Groups 2, 4, and 9.

Non-mathematical: Sum of ranks in Groups 1, 6, and 8.

Appendix III

THE DISCOVERY TEST

Instructions for Administering

A Sample Discovery Test

ADMINISTERING THE DISCOVERY TEST

General remarks. The test consists of two warm-up items and eight items for data. Each item consists of 11 instances of a particular generalization. Detailed instructions for administering the items are contained in the attached flow-chart (see page 30). The sequence is as follows: present instance, wait, record correct answer (whether given by the subject or not), present next instance, etc., until the subject gives two consecutive correct responses. Instances already looked at are left exposed.

The items might be classified into two types: short-cut items or secret-rule items. The secret-rule items involve discovering a simple function and are identified by a symbol such as Δ. The items are presented in a random order, with secret-rule items and short-cut items interlaced but with each test indicating the secret-rule items by the same sequence of symbols--viz., ◇, Φ, Δ, Σ.

Giving the test. (Equipment: stop-watch, paper, two pencils, cardboard mask)

1. Check name and sex of subject. If birthdate is missing, record it on the data form. Move as quickly as possible into the test, but if the subject seems ill-at-ease, try to overcome that by chatting about TV, sports, space, etc.

2. In giving the warm-ups,

a. Explain that the stop-watch is used to "Watch the time to keep from spending too much time on one problem."

b. Record wrong answers during warm-up (after you've written the correct answer) to get the subject accustomed to same. Explain that this is done so that later we can see what kind of wrong answers come up.

3. In general,

a. Respond to a correct answer by saying, "Right." Respond to an incorrect answer by saying, "The correct answer is . . . " and write the answer while saying it.

b. On exposing a new instance, say "How about this?" After writing the correct answer, say "study this (these)," until S does so automatically. Late in the interview you may need to start saying it again.

c. You may need to remind "Try to find a short-cut," if S tries (or seems to be trying) to do the calculation involved in an instance. If S is doing a calculation involved in a possible short-cut, do not interrupt unless it clearly leads to nothing productive.

d. Keep separate the items on which the pupil is successful (two consecutive correct responses) to make it easy to find these items for the verbalizing treatment.

4. Giving the test:

Prefatory remarks "You will be given some sets of problems. This is not a test and it won't count on your grade; we just want to see how -graders do on problems like these. Some of the problems are harder than others. Hardly anyone gets them all correct.

"Each set of problems has parts that are alike in some way. Try to figure out a way to answer the problems. You may not be able to do this at first so the first answer in each set is already given. After each problem you will have a short time to study your work. Use the paper if you want to. Do not be afraid of making mistakes. Don't be afraid to guess."

<u>Warm-ups</u>

"There are two kinds of problems. Here is a sample of one kind. It might be called a short-cut problem. See if you can give the answers without calculating. <u>Do not be afraid of making mistakes</u>."

(Give first warm-up.)

"Here is a sample of the other kind of problem. It might be called a secret-rule problem. <u>Do not be afraid of making mistakes</u>. The sign for the secret rule (point) is so that we can remember it better."

(Give second warm-up.)

<u>Repeat</u>

"Remember this is not a test. Do not be afraid of making mistakes. Hardly anyone figures them all out."

<u>The test proper</u>

(Avoid directive statements or hints. Repeat underlining used in first instance, if any. Emphasize the secret-rule sign: "This is so we can remember it better." After instance 5 of each item if S is unsuccessful, say, "Most people don't get this one right away." If S fails, say, "This is a hard one for a lot of people."--But don't say this eight times!)

<u>Additional cueing</u>

Sum of first n odds: Underline 2 in "First 2," etc. Read and emphasize--e.g., "What is the sum of the first nine numbers from the pattern?" "The sum of the first nine numbers is 81."

If S forms generalization (two consecutive correct responses), say, "Right. You've figured this one out."

NoV

"Let's go back and look at the items you were getting the answers to. We can study them so you can memorize them."

Which instances — Expose only those instances S saw earlier (so can have idea of how much practice).

Time — Devote 40-50 seconds to each page, as S's interest warrants. If S's interest lags before 40 seconds, say, "Do you have the short-cut/the secret-rule and its sign memorized? Study it some more."

Secret rules — "Look over this secret-rule. Notice the sign for this secret rule. (Point.) Memorize the secret-rule and its sign."

Short-cuts — "Look over this short-cut. Memorize it."

If S has forgotten, say, "Try to figure it out again," and when/if he does, say, "Memorize the short-cut/the secret-rule and its sign."

When S is to be dismissed, ask him not to tell what is on the test since you will be talking to other -graders; say he can tell them "It's some arithmetic problems."

SubV

"Let's go back and look at the items you were getting the answers to. We can study them so you can memorize them."

Which instances	Expose only those instances S saw earlier (so can have idea of how much practice).
Time	Including S's verbalization, devote 40-50 seconds to each page as S's interest warrants. If S's interest lags before 40 seconds, say, "Do you have the short-cut/the secret rule and its sign memorized? Study it some more."
Lead-in	"What short-cut did you use?" OR "What secret rule did you use for this one?" (Point to the secret rule sign in the item instructions.)

After/if S has explained, say, "Now study the short-cut and memorize it." OR "Study the secret rule and its sign and memorize them." If S has forgotten, say, "Try to figure it out again," and when/if he does, have him explain his method. Accept solution demonstrated through example. Note that some items can be done reasonably in more than one way--e.g., (n x n) + n instead of n (n + 1).

When S is to be dismissed, ask him not to tell what is on the test since you will be talking to other -graders. Say he can tell them "It's some arithmetic problems."

ExV

"Let's go back and look at the items you were getting the answers to. We can study them so you can memorize them."

Which instances	Expose only those instances S saw earlier (so can have idea of how much practice).
Time	Including your verbalization, devote 40-50 seconds to each page as S's interest warrants. If S's interest lags before 40 seconds, say, "Do you have the short-cut/the secret-rule and its sign memorized? Study it some more."
Verbalizations	"One way to get the answers here is to . . . " (Demonstrate on first instance.)

1. (1 + 2 . . . + 2 + 1) "Find the middle number (point) and take it times itself."

2. (_5 x _5) "Take this number (point) times one more than it, write that, and then write 25."

3. (sum odds) "Find out how many numbers there are and take that times itself."

4. (49 x 51) "Take the first number (point) times the sum of the other two (point), write that, and then write 99."

5. (2 ---> 6) "Add 4 to the number (point) . . . Memorize the sign for this secret rule."

6. (1, 4, 6 ---> 4/7) "Take the first number times the second (point), put that on top; add first and last number, put that on bottom . . . Memorize the sign for this secret rule."

7. (7 ---> 56) "Take the number (point) times one more than it. Memorize the sign for this secret rule."

8. (2, 3, 4, 5 ---> 6/8) "Add the first number and the third number (point), write that on top; add the second and last numbers, write that in the bottom . . . Memorize the sign for this secret rule."

Then, "Study the short-cut/and memorize it." or "Study the secret rule and its sign and memorize them."

Repetitions Repeat above only if asked.

When S is to be dismissed, ask him not to tell what is on the test since you will be talking to other -graders. Say he can tell them "It's some arithmetic problems."

DISCOVERY

TEST

The test contained herein was developed at The Wisconsin Research and Development Center for Cognitive Learning pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program. Center No. C-03/Contract OE 5-10-154

__Warm-up__ See if you can find a short way to get the answers.

11 x 11 = 121

111 x 111 =

1111 x 1111 =

111,111 x 111,111 =

11,111 x 11,111 =

1,111,111 x 1,111,111 =

111,111,111 x 111,111,111 =

11,111,111 x 11,111,111 =

Warm-up Here a number is matched to another number by a secret rule, *. See if you can figure out the secret rule and give the answers.

1 $\overset{*}{\longrightarrow}$ 2

3 $\overset{*}{\longrightarrow}$

4 $\overset{*}{\longrightarrow}$

2 $\overset{*}{\longrightarrow}$

10 $\overset{*}{\longrightarrow}$

8 $\overset{*}{\longrightarrow}$

6 $\overset{*}{\longrightarrow}$

9 $\overset{*}{\longrightarrow}$

7 $\overset{*}{\longrightarrow}$

5 $\overset{*}{\longrightarrow}$

11 $\overset{*}{\longrightarrow}$

Look at the pattern of numbers on the card. Can you tell what the next row is? Here we are interested in the rows of the pattern. See if you can find a short-cut to get the sum when we add the numbers in a row.

(These numbers were on a 3 x 5 card:

1
1 2 1
1 2 3 2 1
1 2 3 4 3 2 1)

1 + 2 + 1 = 4

1 + 2 + 3 + 4 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 7 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 11 + 10 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 10 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 =

1 + 2 + 3 + on up to 20, and then down + 6 + 5 + 4 + 3 + 2 + 1 =

Here a number is matched to another number by a secret rule, $\diamond$.
See if you can figure out the secret rule and give the answers.

$2 \overset{\diamond}{\dashrightarrow} 6$

$5 \overset{\diamond}{\dashrightarrow}$

$7 \overset{\diamond}{\dashrightarrow}$

$12 \overset{\diamond}{\dashrightarrow}$

$25 \overset{\diamond}{\dashrightarrow}$

$33 \overset{\diamond}{\dashrightarrow}$

$50 \overset{\diamond}{\dashrightarrow}$

$62 \overset{\diamond}{\dashrightarrow}$

$18 \overset{\diamond}{\dashrightarrow}$

$43 \overset{\diamond}{\dashrightarrow}$

$81 \overset{\diamond}{\dashrightarrow}$

These problems involve multiplying a number ending in 5 by itself. See if you can find a short-cut to get the answers.

65 x 65 = 4225

25 x 25 =

105 x 105 =

55 x 55 =

45 x 45 =

85 x 85 =

15 x 15 =

75 x 75 =

35 x 35 =

95 x 95 =

205 x 205 =

Here a number is matched to another number by a secret rule, Φ.
See if you can figure out the secret rule and give the answers.

7 $\xrightarrow{\Phi}$ 56

1 $\xrightarrow{\Phi}$

8 $\xrightarrow{\Phi}$

2 $\xrightarrow{\Phi}$

6 $\xrightarrow{\Phi}$

3 $\xrightarrow{\Phi}$

9 $\xrightarrow{\Phi}$

5 $\xrightarrow{\Phi}$

4 $\xrightarrow{\Phi}$

10 $\xrightarrow{\Phi}$

20 $\xrightarrow{\Phi}$

1, 3, 5, 7, 9, 11, , , , , , , , and so on.

See if you can find a short way of getting the answer when we add numbers from the pattern.

First 2: 1 + 3 = 4

First 9: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 =

First 3: 1 + 3 + 5 =

First 8: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 =

First 6: 1 + 3 + 5 + 7 + 9 + 11 =

First 10: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 =

First 7: 1 + 3 + 5 + 7 + 9 + 11 + 13 =

First 11: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 =

First 15: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 + 23 + 25 + 27 + 29 =

First 20: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 + 23 + 25 + 27 + 29 + 31 + 33 + 35 + 37 + 39 =

First 12: 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 + 23 =

Here three numbers are matched with a fraction by a secret rule, Δ. See if you can figure out the secret rule and give the answers.

1,4,6 $\xrightarrow{\Delta}$ $\frac{4}{7}$

3,5,2 $\xrightarrow{\Delta}$

9,2,1 $\xrightarrow{\Delta}$

3,6,3 $\xrightarrow{\Delta}$

2,7,4 $\xrightarrow{\Delta}$

9,8,6 $\xrightarrow{\Delta}$

3,5,4 $\xrightarrow{\Delta}$

8,7,4 $\xrightarrow{\Delta}$

3,5,7 $\xrightarrow{\Delta}$

6,3,5 $\xrightarrow{\Delta}$

3,9,4 $\xrightarrow{\Delta}$

These problems involve multiplying. See if you can find a short-cut to get the answers.

49 x 51 = 2499

79 x 81 =

109 x 111 =

39 x 41 =

99 x 101 =

29 x 31 =

89 x 91 =

69 x 71 =

19 x 21 =

209 x 211 =

59 x 61 =

Here four numbers are matched with a fraction by a secret rule, Σ.
See if you can figure out the secret rule and give the answers.

2,3,4,5 $\xrightarrow{\Sigma}$ $\frac{6}{8}$

1,3,5,6 $\xrightarrow{\Sigma}$

2,4,5,9 $\xrightarrow{\Sigma}$

3,4,5,2 $\xrightarrow{\Sigma}$

6,4,2,3 $\xrightarrow{\Sigma}$

8,9,2,1 $\xrightarrow{\Sigma}$

7,4,5,8 $\xrightarrow{\Sigma}$

3,6,9,2 $\xrightarrow{\Sigma}$

2,5,8,3 $\xrightarrow{\Sigma}$

4,6,7,5 $\xrightarrow{\Sigma}$

6,8,9,2 $\xrightarrow{\Sigma}$

Appendix IV

SCHEDULE OF DISCOVERY TEST ADMINISTRATIONS

Schedule of Discovery Test Administrations

PQR Key: P--High, Middle, or Low IQ level

Q--Boy or Girl

R--No-verbalizing, Subject-verbalizing, or eXperimenter-verbalizing treatment

	Interviewer 1	Interviewer 2	Interviewer 3
		Grade 4	
AM	HBN	MGX	LBS
	MGS	LBN	HBX
	LBX	HBS	MGN
PM	MBN	LGX	HGS
	LGS	HGN	MBX
	HGX	MBS	LGN
		Grade 5	
AM	MBX	LGS	HGN
	LGN	HGX	MBS
	HGS	MBN	LGX
PM	LBX	HBS	MGN
	HBN	MGX	LBS
	MGS	LBN	HBX
		Grade 6	
AM	LBS	HBN	MGX
	HBX	MGS	LBN
	MGN	LBX	HBS
PM	HGS	MBN	LGX
	MBX	LGS	HGN
	LGN	HGX	MBS
		Grade 7	
AM	MBN	HGS	LGX
	LGS	MBX	HGN
	HGX	LGN	MBS
PM	LBN	MGS	HBX
	HBS	LBX	MGN
	MGX	HBN	LBS

Appendix V

THE FOLLOW-UP TEST

Instructions for Administering

A Sample Follow-Up Test

ADMINISTRATION OF THE FOLLOW-UP TEST

General remarks. DO NOT TELL THE STUDENTS THAT THIS TEST IS RELATED TO THE RESEARCH PROJECT. Although the test is given to all the students, the key people to the research are the ones listed below (in point 7), who were involved in the individual interviews. We do not want them to be told that this test is related to the project. The test is a speed test so the students should try to use whatever short-cuts they may know. However, you will determine the time-limits in terms of those students crucial to the experiment, as indicated below.

There are two forms of the test. Each form consists of three parts.

Giving the test

1. Have the students clear their desks except for two pencils. In particular, they should not have access to any scratch paper.

2. The student's names are on the tests. Distribute the tests, telling the students to wait for further instructions before starting the test.

3. Read to the class: "This test is to see how fast you can get answers. Do not worry if you do not finish a page; the test is not going to be used for grading, but just to see how fast -graders can get answers to these kinds of problems. If you can work a problem in your head, do so. Do not

---

This test was developed at the Wisconsin Research and Development Center for Cognitive Learning, pursuant to a contract with the United States Office of Education, Dept, of Health, Education, and Welfare under the provisions of the Cooperative Research Program. Center No. C-03/Contract OE 5-10-154

use other scratch paper. There are three parts to the test. If you finish a part before I say to stop, wait before you go to the next part."

4. Read the instructions on the cover page aloud with your class.

5. Ask if there are any questions. If students ask about how much they should write down, tell them to write down just as much as they need to get the answers. They should write, however, only on the paper provided. If they ask how much time will be allowed, tell them you haven't decided but they should work fast.

6. Checking to see that all pupils are ready, read to the class:

"Turn the page to Part 1 and start. Get the answers as fast as you can. Use any short-cuts you know."

Time started, Part 1: ____________

7. During Part 1 of the test, watch, if you can do so inconspicuously, the following students:

These students were involved in the individual interviews. If any of them seems to erase any calculations, note this in this space:

Who — Which problem (if you can tell)

8. When all the students listed above have finished numbers 1 through 6 of Part 1, read to the class:

"Stop. Do not worry if you did not finish this part . . . Now tear off the cover sheet and Part 1 and hand them in together."

Time stopped, Part 1: ____________ (Collect these papers)

9. Checking to see that all are ready, <u>read to the class</u>:

> "When I say to start, turn to Part 2. Work as fast as you can. If you don't know how to do a problem, skip it. Do not erase, except for mistakes. Don't worry if you do not finish Part 2 . . . Turn the page to Part 2 and start."

Time started, Part 2: ______________

10. Again, watch the students listed above. Note here if they seem to erase any calculations:

<u>Who</u> <u>Which problem</u> (if you can tell)

11. Any time <u>after</u> all the students listed above have finished numbers 1 through 8 of Part 2, <u>read to the class</u>:

> "Stop. Do not worry if you did not finish this part . . . Tear off the blue sheet and the Part 2 sheet and hand them in together."

Time stopped, Part 2: ______________

(Collect these two pages. The students do not need to put their names on these sheets--unless they want to--since the punched hole in the test can be used to identify them.)

(Part 3 was not relevant to the study.)

12. Checking to see that all are ready, <u>read to the class</u>:

> "When I say to start, turn to Part 3. Work as fast as you can. Do not erase, except for mistakes . . . Hold up your hand when you finish . . . Turn to Part 3 and start."

Time started, Part 3: ______________

13. Again, watch the students listed above. Note here if they seem to erase any calculations:

<u>Who</u> <u>Which problem</u> (if you can tell)

14. Collect the remaining two sheets from each student when he holds up his hand. When about half the class has finished, <u>read to the class</u>: "Stop. Do not worry if you are not finished. Hand in these last two sheets."

Time stopped, Part 3 ____________ (Collect these sheets.)

15. Please make any remarks you feel appropriate about the test:

DO NOT START UNTIL YOU ARE TOLD TO.

Work as fast as you can.

Do as much as you can in your head.

Use any short-cuts you know.

To save time--

If you do not know how to do a problem, skip it and go on to the next one.

Do not erase, except for mistakes.

Do not guess.

Part I

1. 100 + 3 + 20 = __________

2. 50 x 10 = __________

3. 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 = __________

4. 95 x 95 = __________

5. 1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 + 23 + 25 + 27 + 29 = __________

6. 69 x 71 = __________

7. 10 + 20 + 30 + 40 + 50 + 60 + 70 + 80 + 90 = __________

8. 550 x 650 = __________

9. 4 + 12 + 4 + 12 + 4 + 12 + 4 + 12 + 4 + 12 + 4 + 12 + 4 + 12 + 4 + 12 = __________

10. 512 + 514 + 516 + 518 + 520 + 522 + 524 = __________

<u>STOP. WAIT.</u>

Part 2

1. 18 --◇-> ____
2. 6,3,5 --Φ-> ____
3. 4 --Δ-> ____
4. 2,5,8,3 --Σ-> ____
5. 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 11 + 10 + 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 = ____
6. 75 x 75 = ____
7. Find the sum of the first 20 numbers from the pattern 1, 3, 5, 7, 9, 11, 13, and so on.
   1 + 3 + 5 + 7 + 9 + 11 + 13 + 15 + 17 + 19 + 21 + 23 + 25 + 27 + 29 + 31 + 33 + 35 + 37 + 39 = ____
8. 59 x 61 = ____
9. Find the sum of the first 15 numbers from the pattern 2, 4, 6, 8, 10, 12, 14, and so on.
   2 + 4 + 6 + 8 + 10 + 12 + 14 + 16 + 18 + 20 + 22 + 24 + 26 + 28 + 30 = ____
10. 11,111 x 11,111 = ____
11. 70 + 73 + 76 + 79 + 82 + 85 + 88 + 91 + 94 + 97 + 100 = ____
12. 50 + 45 + 40 + 35 + 30 + 25 + 20 + 15 + 10 + 5 = ____

STOP. WAIT.

Appendix VI

DISCOVERY TEST PERFORMANCE

Table VI-1

Mean Number of Instances Required for Item and Test

	1	2	3	4	5	6	7	8	Row sum
4	9.0	9.4	8.8	10.0	5.6	8.2	9.5	6.9	67.4
5	6.2	9.0	6.9	9.1	4.6	5.9	8.5	6.2	56.4
6	6.2	8.4	5.8	8.3	4.2	4.8	7.1	3.7	48.6
7	7.1	9.2	5.8	8.4	3.4	4.8	6.4	4.3	49.6
H	6.2	7.8	6.0	8.6	3.8	4.5	6.9	4.5	48.3
M	7.0	9.3	7.8	9.0	5.2	5.8	8.4	4.6	57.1
L	8.3	9.9	6.7	9.3	4.3	7.5	8.3	6.7	61.0
B	6.5	8.9	6.7	9.0	5.0	6.2	7.9	5.7	55.8
G	7.8	9.2	6.9	8.9	3.8	5.7	7.9	4.8	55.1
4B	9.3	9.4	8.9	10.0	6.8	8.3	9.1	7.2	69.1
4G	8.7	9.4	8.7	10.0	4.3	8.1	9.9	6.6	65.7
5B	4.2	8.3	5.8	8.8	4.8	6.4	7.9	7.1	53.3
5G	8.3	9.7	8.0	9.3	4.3	5.4	9.1	5.3	59.6
6B	6.8	8.6	6.4	9.7	4.3	5.0	8.3	4.1	53.2
6G	5.8	8.3	5.1	6.9	4.0	4.7	5.9	3.2	43.9
7B	5.7	9.1	5.8	7.4	4.1	5.0	6.1	4.4	47.7
7G	8.6	9.2	5.9	9.4	2.8	4.6	6.8	4.2	51.4
HB	5.6	7.3	6.3	8.6	4.1	4.9	6.6	5.5	48.8
HG	6.8	8.4	5.6	8.6	3.5	4.1	7.3	3.6	47.8
MB	6.7	9.6	8.0	8.8	5.8	6.1	8.4	4.6	57.9
MG	7.3	9.1	7.6	9.3	4.6	5.6	8.4	4.6	56.3
LB	7.3	9.8	5.8	9.6	5.1	7.6	8.6	7.1	60.8
LG	9.4	10.0	7.6	8.9	3.5	7.4	8.1	6.3	61.3
4H	8.3	8.3	8.0	10.0	5.3	6.5	9.2	6.7	62.3
4M	8.7	10.0	9.3	10.0	7.5	8.2	10.0	7.2	70.8
4L	10.0	10.0	9.0	10.0	3.8	10.0	9.3	6.8	69.0
5H	6.2	7.8	5.0	9.5	4.8	4.0	9.0	5.5	51.8
5M	5.5	9.2	8.3	8.2	3.5	5.7	8.5	4.8	53.7
5L	7.2	10.0	7.3	9.5	5.3	8.2	8.0	8.3	63.8
6H	6.2	7.2	6.0	7.2	2.5	4.0	4.7	3.0	40.7
6M	5.0	8.7	6.2	9.3	6.2	4.8	8.3	2.8	51.3
6L	7.7	9.5	5.2	8.3	3.8	5.7	8.3	5.2	53.7
7H	4.2	8.0	4.8	7.7	2.5	3.5	4.8	3.0	38.5
7M	8.7	9.5	7.3	8.5	3.7	4.7	6.8	3.5	52.7
7L	8.5	10.0	5.3	9.2	4.2	6.2	7.7	6.5	57.5

Table VI-1 (continued)

	1	2	3	4	5	6	7	8	Row sum
4HB	8.7	8.3	10.0	10.0	7.7	7.0	8.3	7.7	67.7
4HG	8.0	8.3	6.0	10.0	3.0	6.0	10.0	5.7	57.0
4MB	9.3	10.0	8.7	10.0	7.3	8.0	10.0	5.3	68.3
4MG	8.0	10.0	10.0	10.0	7.7	8.3	10.0	9.0	73.0
4LB	10.0	10.0	8.0	10.0	5.3	10.0	9.0	8.7	71.0
4LG	10.0	10.0	10.0	10.0	2.3	10.0	9.7	5.0	67.0
5HB	3.7	6.7	4.0	10.0	4.3	4.7	9.0	6.7	49.0
5HG	8.7	9.0	6.0	9.0	5.3	3.3	9.0	4.3	54.7
5MB	4.7	8.3	8.7	6.3	3.7	6.0	7.0	5.7	50.3
5MG	6.3	10.0	8.0	10.0	3.3	5.3	10.0	4.0	57.0
5LB	4.3	10.0	4.7	10.0	6.3	8.7	7.7	9.0	60.7
5LG	10.0	10.0	10.0	9.0	4.3	7.7	8.3	7.7	67.0
6HB	7.7	6.7	7.3	9.0	2.3	4.7	6.0	4.0	47.7
6HG	4.7	7.7	4.7	5.3	2.7	3.3	3.3	2.0	33.7
6MB	5.3	10.0	6.7	10.0	7.3	4.7	9.0	2.6	55.7
6MG	4.7	7.3	5.7	8.7	5.0	5.0	7.7	3.0	47.0
6LB	7.3	9.0	5.3	10.0	3.3	5.7	10.0	5.7	56.3
6LG	8.0	10.0	5.0	6.7	4.3	5.7	6.7	4.7	51.0
7HB	2.3	7.3	4.0	5.3	2.0	3.3	3.0	3.7	31.0
7HG	6.0	8.7	5.7	10.0	3.0	3.7	6.7	2.3	46.0
7MB	7.3	10.0	8.0	8.7	5.0	5.7	7.7	4.7	57.0
7MG	10.0	9.0	6.7	8.3	2.3	3.7	6.0	2.3	48.3
7LB	7.3	10.0	5.3	8.3	5.3	6.0	7.7	5.0	55.0
7LG	9.7	10.0	5.3	10.0	3.0	6.3	7.7	8.0	60.0

Table VI-2

Number of Pupils Generalizing Each Item

	1	2	3	4	5	6	7	8	Row sum
(18 pupils for each grade)									
4	5	2	4	0	14	8	3	10	46
5	11	5	11	4	16	13	6	12	78
6	11	7	12	7	16	16	9	15	93
7	9	5	12	6	17	16	10	16	91
Col. mean	9	4.8	9.8	4.3	15.8	13.3	7.0	13.3	77
(24 pupils for each IQ level)									
H	16	13	16	8	22	23	13	20	131
M	13	5	11	5	19	20	7	20	100
L	7	1	12	4	22	10	8	13	77
Col. mean	12.0	6.3	13.0	5.7	21.0	17.7	9.3	17.7	102.7
(36 pupils for each sex)									
B	20	10	20	8	30	25	14	25	152
G	16	9	19	9	33	28	14	28	156
Col. mean	18	9.5	19.5	8.5	31.5	26.5	14.0	26.5	154
(9 pupils for each group)									
4B	2	1	2	0	6	3	2	4	20
4G	3	1	2	0	8	5	1	6	26
5B	8	4	7	2	8	6	4	5	44
5G	3	1	4	2	8	7	2	7	34
6B	4	3	5	1	8	8	3	7	39
6G	7	4	7	6	8	8	6	8	54
7B	6	2	6	5	8	8	5	9	49
7G	3	3	6	1	9	8	5	7	42
(6 pupils for each group)									
4H	3	2	2	0	5	5	1	4	22
4M	2	0	1	0	4	3	0	3	13
4L	0	0	1	0	5	0	2	3	11
5H	4	3	5	1	5	6	2	5	31
5M	4	2	3	2	6	5	2	5	29
5L	3	0	3	1	5	2	2	2	18
6H	4	4	4	4	6	6	5	5	38
6M	5	2	4	1	4	6	2	6	30
6L	2	1	4	2	6	4	2	4	25
7H	5	4	5	3	6	6	5	6	40
7M	2	1	3	2	5	6	3	6	28
7L	2	0	4	1	6	4	2	4	23

Table VI-2 (continued)

	1	2	3	4	5	6	7	8	Row sum
(12 pupils for each group)									
HB	8	7	7	4	11	11	7	8	63
HG	8	6	9	4	11	12	6	12	68
MB	7	2	5	3	9	9	4	10	49
MG	6	3	6	2	10	11	3	10	51
LB	5	1	8	1	10	5	3	7	40
LG	2	0	4	3	12	5	5	6	37
(3 pupils for each group)									
4HB	1	1	0	0	2	2	1	1	8
4HG	2	1	2	0	3	3	0	3	14
4MB	1	0	1	0	2	1	0	2	7
4MG	1	0	0	0	2	2	0	1	6
4LB	0	0	1	0	2	0	1	1	5
4LG	0	0	0	0	3	0	1	2	6
5HB	3	2	3	0	3	3	1	2	17
5HG	1	1	2	1	2	3	1	3	14
5MB	2	2	1	2	3	2	2	2	16
5MG	2	0	2	0	3	3	0	3	13
5LB	3	0	3	0	2	1	1	1	11
5LG	0	0	0	1	3	1	1	1	7
6HB	1	2	1	1	3	3	2	2	15
6HG	3	2	3	3	3	3	3	3	23
6MB	2	0	2	0	2	3	1	3	13
6MG	3	2	2	1	2	3	1	3	17
6LB	1	1	2	0	3	2	0	2	11
6LG	1	0	2	2	3	2	2	2	14
7HB	3	2	3	3	3	3	3	3	23
7HG	2	2	2	0	3	3	2	3	17
7MB	2	0	1	1	2	3	1	3	13
7MG	0	1	2	1	3	3	2	3	15
7LB	1	0	2	1	3	2	1	3	13
7LG	1	0	2	0	3	2	1	1	10

Table VI-3

Pupil Performance on the Discovery Test

	1	2	3	4	5	6	7	8	Total-instances	Number of gen.	G/I
4HBS	10	10	10	10	4	10	10	10	74	1	.135
4HBX	6	10	10	10	10	6	5	3	60	4	.667
4HBN	10	5	10	10	9	5	10	10	69	3	.435
4HGS	10	10	10	10	2	7	10	5	64	3	.469
4HGX	8	10	4	10	3	4	10	6	55	5	.909
4HGN	6	5	4	10	4	7	10	6	52	6	1.154
4MBS	10	10	10	10	3	10	10	4	67	2	.298
4MBX	10	10	10	10	10	10	10	10	80	0	.000
4MBN	8	10	6	10	9	4	10	2	59	5	.847
4MGS	4	10	10	10	10	7	10	10	71	2	.282
4MGX	10	10	10	10	5	10	10	7	72	2	.278
4MGN	10	10	10	10	8	8	10	10	76	2	.263
4LBS	10	10	4	10	2	10	7	10	63	3	.476
4LBX	10	10	10	10	4	10	10	6	70	2	.286
4LBN	10	10	10	10	10	10	10	10	80	0	.000
4LGS	10	10	10	10	2	10	10	1	63	2	.317
4LGX	10	10	10	10	3	10	9	10	72	2	.278
4LGN	10	10	10	10	2	10	10	4	66	2	.303
5HBS	6	6	5	10	6	3	7	10	53	6	1.132
5HBX	3	10	4	10	3	3	10	2	45	5	1.111
5HBN	2	4	3	10	4	8	10	8	49	6	1.224
5HGS	10	7	10	10	2	3	10	3	55	4	.727
5HGX	10	10	5	7	10	3	10	6	61	4	.656
5HGN	6	10	3	10	4	4	7	4	48	6	1.250
5MBS	10	10	10	6	5	10	10	10	71	2	.282
5MBX	2	9	10	3	3	2	4	1	34	7	2.059
5MBN	2	6	6	10	3	6	7	6	46	7	1.522
5MGS	6	10	6	10	2	4	10	5	53	5	.944
5MGX	10	10	10	10	6	8	10	3	67	3	.448
5MGN	3	10	8	10	2	4	10	4	51	5	.980
5LBS	6	10	6	10	5	10	3	7	57	5	.877
5LBX	4	10	4	10	4	6	10	10	58	4	.690
5LBN	3	10	4	10	10	10	10	10	67	2	.298
5LGS	10	10	10	7	2	3	5	3	50	5	1.000
5LGX	10	10	10	10	4	10	10	10	74	1	.135
5LGN	10	10	10	10	7	10	10	10	77	1	.130

Table VI-3 (continued)

	1	2	3	4	5	6	7	8	Total-instances	Number of gen.	G/I
6HBS	10	10	10	10	2	3	6	1	52	4	.769
6HBX	3	2	2	10	2	4	2	1	26	7	2.692
6HBN	10	8	10	7	3	7	10	10	65	4	.615
6HGS	3	7	2	5	2	3	3	1	26	8	3.077
6HGX	7	10	3	6	2	3	2	4	37	7	1.892
6HGN	4	6	9	5	4	4	5	1	38	8	2.105
6MBS	4	10	10	10	6	4	10	1	55	4	.727
6MBX	2	10	3	10	10	4	10	4	53	4	.755
6MBN	10	10	7	10	6	6	7	3	59	5	.848
6MGS	4	10	10	10	10	6	10	3	63	3	.476
6MGX	4	4	3	10	2	7	10	3	43	6	1.396
6MGN	6	8	4	6	3	2	3	3	35	8	2.286
6LBS	2	10	4	10	2	2	10	4	44	5	1.137
6LBX	10	10	10	10	5	10	10	10	75	1	.133
6LBN	10	7	2	10	3	5	10	3	50	5	1.000
6LGS	10	10	2	6	2	3	4	3	40	6	1.500
6LGX	4	10	3	4	2	4	6	1	34	8	2.059
6LGN	10	10	10	10	9	10	10	10	79	1	.127
7HBS	3	10	4	3	2	3	3	2	30	7	2.333
7HGX	2	8	2	6	2	3	3	1	27	8	2.963
7HBN	2	4	6	7	2	4	3	8	36	8	2.222
7HGS	4	10	4	10	2	6	5	2	43	6	1.395
7HGX	10	8	3	10	5	2	5	3	46	6	1.304
7HGN	4	8	10	10	2	3	10	2	49	5	1.020
7MBS	6	10	10	6	3	7	3	2	47	6	1.277
7MBX	6	10	4	10	2	5	10	6	53	5	.944
7MBN	10	10	10	10	10	5	10	6	71	2	.282
7MGS	10	7	4	5	2	4	10	1	43	6	1.395
7MGX	10	10	10	10	2	3	5	5	55	4	.727
7MGN	10	10	6	10	3	4	3	1	47	5	1.064
7LBS	10	10	10	10	7	10	10	8	75	2	.267
7LBX	10	10	4	10	5	5	10	1	55	4	.727
7LBN	2	10	2	5	4	3	3	6	35	7	2.000
7LGS	10	10	10	10	4	10	10	10	74	1	.135
7LGX	10	10	2	10	3	7	10	10	62	3	.484
7LGN	9	10	4	10	2	2	3	4	44	6	1.364

Table VI-4

Order in which Items Were Given to Each Pupil
(See Table 3, ch. 2, for description of items.)

4HBS	18362547	6HBS	62837451
4HBX	47152836	6HBX	36482517
4HBN	61735284	6HBN	17382645
4HGS	47162538	6HGS	16374528
4HGX	64528371	6HGX	84637152
4HGN	51637482	6HGN	74638152
4MBS	52816374	6MBS	15473628
4MBX	52847163	6MBX	52618473
4MBN	48263715	6MBN	64825173
4MGS	36251847	6MGS	46371825
4MGX	28451736	6MGX	45261738
4MGN	73526481	6MGN	84517362
4LBS	47163825	6LBS	73845261
4LBX	64738251	6LBX	25374618
4LBN	38251647	6LBN	48173625
4LGS	37152648	6LGS	16284735
4LGX	73816452	6LGX	62847153
4LGN	54736281	6LGN	82637154
5HBS	83547162	7HBS	62547381
5HBX	45283716	7HBX	74815263
5HBN	53718264	7HBN	37481625
5HGS	26153748	7HGS	26453718
5HGX	15284637	7HGX	81745263
5HGN	71528463	7HGN	45382617
5MBS	51628374	7MBS	63825471
5MBX	52617384	7MBX	15473826
5MBN	36174528	7MBN	17452638
5MGS	26453817	7MGS	83625174
5MGX	63748152	7MGX	71625384
5MGN	82536174	7MGN	37154826
5LBS	81745362	7LBS	37182546
5LBX	15283746	7LBX	73648251
5LBN	28174536	7LBN	81746253
5LGS	28154637	7LGS	51738462
5LGX	14625183	7LGX	28364517
5LGN	35264817	7LGN	28371645

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, Frank B. "Teaching for Generalization in Geometry," The Mathematics Teacher, 43:245-251, October, 1950.

Ausubel, David P. "Learning by Discovery: Rationale and Mystique," Bulletin of the National Association of Secondary School Principals, 45:18-58, December, 1961.

The Psychology of Meaningful Verbal Learning, New York, Grune and Stratton, 1963.

Baker, Frank B. "An Intersection of Test Score Interpretation and Item Analysis," Journal of Educational Measurement, 2:167-180, 1965.

Baker, Frank B., and T. J. Martin. "FORTAP: A Fortran Test Analysis Package," Madison, Wisconsin Research and Development Center for Cognitive Learning, 1968.

Bartlett, F. Thinking, New York, Basic Books, 1958.

Bloom, Benjamin. Stability and Change in Human Characteristics, New York, Wiley, 1964.

Bruner, Jerome S. "The Act of Discovery," Harvard Educational Review, 31:21-32, Winter, 1961.

Bruner, J., J. Goodnow, and G. Austin. A Study of Thinking, New York, Science Editions, 1956.

Buros, O. K. The Sixth Mental Measurements Yearbook, Highland Park, N. J., Gryphon Press, 1965.

Cattell, Raymond B. Editor Handbook of Multivariate Experimental Psychology, Chicago, Rand McNally, 1966.

Davis, Robert B. "The Madison Project's Approach to a Theory of Instruction," Journal of Research in Science Teaching, 2:146-162, 1964.

"The Range of Rhetorics, Scale and Other Variables," Journal of Research and Development in Education, 1:51-74, Fall, 1967.

Ebert, Reuben S. "Generalization Abilities in Mathematics," Journal of Educational Research, 39:671-681, May, 1946.

Finn, Jeremy D. Multivariance: Fortran Program for Univariate and Multivariate Analysis of Variance and Covariance, Buffalo, State University of New York, 1967.

Flavell, John H. The Developmental Psychology of Jean Piaget, Princeton, Van Nostrand, 1963.

Friedlander, Bernard. "A Psychologist's Second Thoughts on Concepts, Curiosity, and Discovery in Teaching and Learning," Harvard Educational Review, 35:18-38, Winter, 1965.

Gagné, R. M., and L. T. Brown. "Some Factors in the Programming of Conceptual Learning," Journal of Experimental Psychology, 62:313-321, October, 1961.

Gagné, Robert, and E. Smith, Jr. "A Study of the Effects of Verbalization on Problem Solving," Journal of Experimental Psychology, 63:12-18, No. 1, 1962.

Hanson, Lawrence E. "Inductive Discovery Learning, Reception Learning, and Formal Verbalization of Mathematical Concepts," unpublished doctoral dissertation, Florida State University, 1967.

Harris, Margaret L. "Some Methodological Suggestions for Construction of an Objective Measurement Instrument," Madison, Wisconsin Research and Development Center for Cognitive Learning Technical Memo, 1968.

Haslerud, George, and Shirley Meyers. "The Transfer Value of Given and Individually Derived Principles," The Journal of Educational Psychology, 49:293-298, December, 1958.

Heath, Robert W. New Curricula, New York, Harper and Row, 1964.

Heidbreder, E. "A Study of the Evolution of Concepts," Psychological Bulletin, 31:673, 1934.

"Language and Concepts," Psychological Bulletin, 33:724, 1936.

Hendrix, Gertrude. "A New Clue to Transfer of Training," Elementary School Journal, 48:197-208, December, 1947.

"Prerequisite to Meaning," The Mathematics Teacher, 43:334-339, 1950.

"Learning by Discovery," The Mathematics Teacher, 54:290-299, May, 1961.

Houston, Tom. "N-Way Analysis of Variance--ANOVAH," Madison, Laboratory of Experimental Design of the Wisconsin Research and Development Center for Cognitive Learning, 1967.

Humphrey, George. Thinking: An Introduction to its Experimental Psychology, New York, Wiley, 1963.

Husén, Torsten, and others. International Study of Achievement in Mathematics: A Comparison of Twelve Countries, Vol. 2, Wiley, 1967.

Judd, C. H. The Psychology of Secondary Education, Boston, Ginn, 1927.

Kagan, Jerome, B. Rosman, D. Day, J. Albert, and W. Phillips, "Information Processing in the Child: Significance of Analytic and Reflective Attitudes," Psychological Monographs, 78, No. 1 (Whole No. 578), 1964.

Katona, George. Organizing and Memorizing, New York, Columbia University Press, 1940.

Kelley, T. L., R. Madden, E. F. Gardner, and H. C. Rudman. Stanford Achievement Test, New York, Harcourt, Brace and World, 1963, 1964, 1965.

Kersh, Bert Y. "The Adequacy of 'Meaning' as an Explanation for Superiority of Learning by Independent Discovery," Journal of Educational Psychology, 49:282-292, 1958.

"The Motivating Effect of Learning by Directed Discovery," Journal of Educational Psychology, 53:65-71, 1962.

Kittell, Jack E. "The Effect of External Direction During Learning on Transfer and Retention of Principles," Journal of Educational Psychology, 48:391-405, 1957.

Kuhlmann-Anderson Test, Booklets CD and EF, Princeton, Personnel Press, 1963, 1960, respectively.

Lord, Frederic M., and Melvin R. Novick. Statistical Theories of Mental Test Scores, Reading, Massachusetts, Addison-Wesley, 1968.

McSwain, E. T., K. E. Brown, B. H. Gundlach, and R. J. Cooke. Arithmetic 4 (1965a), Arithmetic 5 (1965b), Arithmetic 6 (1965c), Mathematics 7 (1963), River Forest, Ill., Laidlaw.

National Committee on Mathematical Requirements, The Reorganization of Mathematics in Secondary Education, The Mathematical Association of America, Inc., 1923.

Osler, S. F., and M. W. Fivel. "Concept Attainment: I. The Role of Age and Intelligence in Concept Attainment by Induction," Journal of Experimental Psychology, 62:1-8, July, 1961.

Romberg, Thomas A. "Guidelines for Status Studies," Madison, Wisconsin Research and Development Center for Cognitive Learning Technical Memo, 1968.

Romberg, Thomas, and M. Vere DeVault. "Mathematics Curriculum: Needed Research," Journal of Research and Development in Education, 1:95-112, Fall, 1967.

Scheffe, Henry. The Analysis of Variance, New York, Wiley, 1959.

School Mathematics Study Group, Goldmine, Stanford, dittoed copy, 1965.

Schwartz, Bert D. "Certain Relations between Verbalization and Concept Formation," unpublished doctoral dissertation, Princeton, 1948.

Shulman, L. S., and E. R. Keislar, Editors Learning by Discovery: A Critical Appraisal, Chicago, Rand McNally, 1966.

Siegel, Sidney. Nonparametric Statistics, New York, McGraw-Hill, 1956.

Smoke, Kenneth L. "An Objective Study of Concept Formation," Psychological Monographs, 32, No. 4 (Whole No. 191), 1932.

Swenson, Esther J., G. Lester Anderson, and Chalmers L. Stacey. Learning Theory in School Situations, University of Minnesota Studies, Number 2, University of Minnesota Press, Minneapolis, 1949.

Thompson, Jane. "The Ability of Children of Different Grade Levels to Generalize on Sorting Tests," Journal of Psychology, 11:119-126, 1941.

Wetterstrand, W. H. "Correlation Analysis--DSTATI," Madison, University of Wisconsin Computing Center, 1966.

Wills, Herbert III. "Transfer of Problem Solving Ability Gained Through Learning by Discovery," unpublished doctoral dissertation, University of Illinois, 1967.