DOCUMENT RESUME

ED 039 611 24 CG 005 634

AUTHOR Gagne, Robert M.
TITLE Basic Studies of Learning Hierarchies in School Subjects. Final Report.
INSTITUTION California Univ., Berkeley.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-6-2949
PUB DATE Apr 70
CONTRACT OEC-4-062949-3006
NOTE 102p.

EDRS PRICE EDRS Price MF-$0.50 HC-$5.20
DESCRIPTORS Courses, *Elementary School Students, Instructional Materials, *Learning, Problem Solving, *Research, *Secondary School Students, Skills, Task Performance

ABSTRACT

This study explores three areas of learning hierarchies: (1) the condition for learning in accordance with learning hierarchies; (2) the use of hierarchies in the diagnosis and learning of prerequisite intellectual skills; and (3) the conditions of learning and retention of principles. There were seven studies undertaken using elementary and secondary school students as subjects. Results indicated: (1) overlearning was not shown to facilitate either achievement or transfer; (2) learning of subordinate skills produces marked positive transfer in learning of complex problem solving tasks; (3) comparing and contrasting tasks are affected by the attainment of subordinate skills; (4) measuring performance on subordinate tasks gives direction for obtaining tryouts and revisions of materials for instruction; (5) retention of concrete rules is 20 per cent after three days; (6) interference operates in determining retention; and (7) the superordinate context cue has a major effect in retrieval. (Author/EK)

BR 6-2949
PA 24

FINAL REPORT
Project No. 6-2949
Contract No. OEC-4-062949-3066

BASIC STUDIES OF LEARNING HIERARCHIES IN SCHOOL SUBJECTS

Robert M. Gagné
University of California, Berkeley
Berkeley, California 94720

April 1970

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT
Project No. 6-2949
Contract No. OEC-4-062949-3066

BASIC STUDIES OF LEARNING HIERARCHIES IN SCHOOL SUBJECTS

Robert M. Gagné
University of California, Berkeley

Berkeley, California
April, 1970

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinion stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This report contains descriptions and conclusions of studies related to the general topic of learning hierarchies. The work of the project was designed to explore, by means of experimental studies and related theory, three problem areas of learning hierarchies. The first was one of specifying conditions for learning in accordance with learning hierarchies; this is dealt with in Chapter 2, Studies 1 and 2. A second problem was to investigate the use of hierarchies in the diagnosis and learning of prerequisite intellectual skills, the subject of Chapter 3, Studies 3 and 4. A third problem concerned the identification of conditions of learning and retention of principles, the components of learning hierarchies. This is the subject of Chapter 4, Studies 5, 6, and 7.

Study 1. Transfer effects of practice variety in principle learning, by Jeanne R. Gibson, University of California, Berkeley.

This study investigates the effect of certain events following initial learning of a principle on the retention and transfer of the principle. The major independent variables were amount of practice (many or few examples) and variety of examples (broad or narrow). Ninety children of the 3rd and 4th grades, of three IQ levels, were the subjects. All children were instructed in subordinate skills related to decimal numbers: reading decimals, writing decimals, and identifying position of a decimal on a number line. Subjects in different groups worked few (30) or many (75) practice problems of a broad or narrow variety, while a control group had no practice examples. Achievement and transfer tests were given following the practice period and again after an interval of several weeks. Results indicated a significant difference favoring narrow variety on the immediate achievement test only. Overlearning with either broad or narrow variety was not shown to facilitate either achievement or transfer.

Study 2. A study of subordinate skills in science problem solving, by Virginia K. Wiegand, University of California, Berkeley.

This investigation concerned transfer of learning from hierarchically arranged subordinate skills to a final task requiring students to generate a general expression relating variables in

an inclined plane, of the sort height x weight = mass x distance pushed (of a standard block). A learning hierarchy was constructed indicating hypothesized prerequisite capabilities for this task. Sixth grade children were tested for their ability to perform the final task, and 30 (out of 31) were found who could not. Tests were then administered, working downwards in the hierarchy, to determine the presence or absence of prerequisite skills. Following this, in one group of 10 subjects, retesting on alternate forms of items, working upwards, was accompanied by demonstration of each task; in a second group of 10, retesting alone was done; and in a third group of 10, no retesting was done. Retesting of subordinate skills was carried up to the point of the final task. Subjects in all groups were then tested on the final task, and also on a transfer task derived from one used by Piaget, involving equilibrium on an inclined plane. Eighteen subjects (of 20) in both retested groups performed the final task successfully, and 19 (of 20) were also able to do the transfer task. Only three (of 10) control subjects were able to do these tasks correctly; and it is notable that these three were subjects who had no missing prerequisite skills when first tested. The experiment confirms the hypothesis that learning of initially missing subordinate skills produces marked positive transfer in the learning of a complex problem-solving task in science.

Study 3. Transfer of learning in a social studies task of comparing-contrasting, by Lois T. Coleman, Nebraska Wesleyan University, and Robert M. Gagné, Florida State University.

An analysis was made of a final comparing task requiring comparison of the exports of two countries, to identify a hierarchy of 21 subordinate skills hypothesized as prerequisite to the successful solving of the final task. Twenty sixth-grade girls, of moderately high IQ levels, were randomly assigned to control and experimental groups, 10 in each. None of these students was able to accomplish the final comparing task, given in its pretest form. The experimental group were first tested in a diagnostic sense to determine the presence or absence of prerequisite skills, and then were given instruction on missing skills. Students in both groups were then tested again on the final comparing task in its post-test form. Success on the final task was 100% for the experimental subjects, 0% for the controls. Scores on a near- and far-transfer comparing task were also markedly and significantly higher for the experimental

group than for the control group, indicating the transferability of the learned skills. Evidence was sought of the sequential ordering of prerequisite skills in the hierarchy. Some slight evidence of ordering appeared in the patterns of success in skills in the class "generating conceptual categories", but none in the classes "quantitative comparisons" and "table interpretation". The high level of performance on subordinate skills in the group studied, however, was seen as limiting the possibility of obtaining such evidence.

Study 4. Revision of a science topic using evidence of performance on subordinate skills, by James R. Okey, University of California, Santa Barbara, and Robert M. Gagné, Florida State University.

An initial instructional program on solving solubility product problems was studied by a group of 49 chemistry students. Following instruction, performance of these students was measured on a criterion test and on 15 skills identified as subordinate to the final task. Performance on these subordinate skills was used to locate specific skills failed by a substantial number of students. Gagné's cumulative learning model served as the basis for identifying the subordinate skills and for predicting instruction needed to overcome deficiencies. Twenty frames were added to the original program in accordance with the learning hierarchy, and with results of student performance, making up a revised program of 70 frames. A second group of 57 students then studied this revised program. The number of subordinate skills attained following this learning period increased significantly, and with few exceptions. Analysis of covariance also showed significant differences favoring the group using the revised program on a post-test performance measure. The results indicate that by measuring student performance on subordinate tasks related to the learning of science principles, direction may be obtained for cycles of tryout and revision of materials for instruction.

Study 5. Some factors in children's learning and retention of rules, by Robert M. Gagné and Virginia K. Wiegand, University of California, Berkeley.

An experiment investigated effects of (a) number of rules, (b) immediate vs. 3-day recall, (c) verbal vs. verbal plus pictorial cueing, and (d) IQ, on the learning and retention of concrete rules. Each rule was composed of a highly pronounceable

CVC as the name of a thing (a drawn shape), and an action (such as, "underline it"). 96 4th-grade children were the experimental Ss. Following sessions providing prelearning on thing concepts and review of action concepts, different groups of children learned 3, 5, 7, or 9 rules from printed booklets, reading and recording each rule once. Virtually complete retention was obtained for 3 and 5 rules when measured immediately, significantly less for 7 and 9. Significant effects were not found for IQ or cueing method. After 3 days, the amount of retention was about 20% under all conditions.

Study 6. Context, isolation, and interference effects on the retention of facts, by Robert M. Gagné, University of California, Berkeley.

In the setting of a science lesson, children of Grades 4 and 5 were given five facts to learn and remember. Several different kinds of presentation were employed, including massed and spaced, accompanying contexts containing (a) a related superordinate fact in the form of a topic sentence, (b) coordinate facts, (c) unrelated facts, and (d) an "isolated" presentation condition containing no context facts. Retention of the facts was measured primarily by recall scores obtained by having the children complete blanks in paraphrased sentences. Results indicated significant effects of the variables of school grade and type of context, but no significant effects of massed versus spaced presentation. Most favorable condition for recall was one with no context during learning. Superordinate context was superior to coordinate, and this in turn superior to unrelated context. The results are interpreted as indicating the operation of interference, and also of organizing factors, in the determination of retention.

Study 7. Effects of a superordinate context on learning and retention of facts, by Robert M. Gagné, Florida State University, and Virginia K. Wiegand, California State College at Hayward.

Children of the fourth grade were given five facts to learn and remember about howling monkeys, each being presented in a context with four other facts. Half of the total group of 44 learned the facts in a context containing a superordinate statement (topic sentence), the other half with a coordinate (related) statement. Retention was measured in half of each group by

presenting the topic sentence immediately before the retention test, and in half without. Significantly greater remembering, measured by recall and recognition scores, was found in groups having the topic sentence presented just prior to the retention test. The results indicate the major effect of the superordinate context cue to occur in retrieval.

CHAPTER 1 - INTRODUCTION

A number of studies of school-relevant learning have investigated the use of systematic analyses of content to be learned in terms of learning hierarchies (Gagné, 1962; Gagné and Paradise, 1961; Gagné, Mayor, Garstens and Paradise, 1962; Gagné and Bassler, 1963; Gagné et al, 1965). Since these studies originally appeared, a good deal of interest has been generated in their potential usefulness for understanding the nature of instruction as exhibited in school-learning activities. The theory of learning hierarchies and its application to the design of instruction has been described by Gagné (1969).

The basic method has been used by other investigators to test the appropriateness of curriculum sequences, ranging over a variety of content. Hively (1963), for example, used the method to test several forms of curricula on finite sets, including the notions of "universe," "complement," "intersection," and "union." Schutz, Baker, and Gerlach (1964) carried out a comprehensive set of experiments using the method to reveal gaps, inversions, and unnecessary units in curricula on capitalization, fractions, and punctuation. Newton and Hickey (1965) used the method to test the feasibility and effectiveness of two different curriculum sequences concerned with economic principles of the Gross National Product. Smith and Moore (1965) explored the sequencing of curriculum units in programmed materials on Chinese History, Programmed Instruction, Logic Sentences, Federal Relations in Education, the Russian Alphabet, and Set Theory. Their findings support the notion that tests can be designed to reveal appropriate learning hierarchies in each of these content areas. Graham and Cox (1966) found that the testing of young children on adding two-digit numerals revealed both correct and incorrect sequences within a curriculum, and made possible the design of a learning hierarchy having a coherent structure as shown by a high coefficient of reproducibility (Guttman, 1944).

Other authors have seen additional implications in the idea of learning hierarchies for an understanding of human functioning. For example, Underwood (1964a, p. 144) sees the possibility of analyzing learning tasks into constituent learnings as a possible new approach to the problem of learning wholes vs. parts. Ferguson (1965) points out the potentiality such analyses have for an improved understanding of human abilities in relation to transfer of learning. Considerable evidence pertaining to the transfer of learning from subordinate intellectual skills to higher-level concepts and rules is reviewed by Gagné and Rohwer (1969).

Problems Investigated

The studies cited have served to indicate the promise of learning hierarchies as a method of analyzing and specifying content to be learned, and also as a model for the acquisition of knowledge by the human learner. They have also indicated many questions still to be answered, before one is able to apply the method with confidence and precision to the design of instructional sequences useful in the schools. The problem in its most general form is this: Can additional evidence be obtained to indicate the conditions under which knowledge of learning hierarchies can be used to design instruction for school-relevant subjects? Particularly, the problems investigated in the present project were concerned with finding evidences of transfer of learning in a "vertical" sense, that is, from simpler intellectual skills to the learning of more complex ones; evidence concerning the breadth of "lateral" transfer to novel problems within the same class as those which learners had mastered; and evidence about the factors affecting learning of the "units" of a learning hierarchy, namely, principles.

Purposes of the Project

The work of the project was designed to explore, by means of a series of experimental studies and developing theory, three problem areas of learning hierarchies that appeared to hold promise of applicability to the design of instruction

A first purpose was to try to specify the conditions of acquisition of knowledge which can be rationally analyzed into hierarchies of principles (or "rules"). Previous work has shown that "minimal requisite" sequences of rules can be identified (in the sense that a rule such as factoring the parts of a fraction must be known before addition of fractions can be learned), and also that hierarchies expressing such sequences can be verified. Unanswered questions are: What are the effects on student achievement when subordinate learning hierarchies are deliberately altered in accordance with certain pedagogical principles? What are the effects of these alterations on the transferability of knowledge? What are the effects of individual differences (other than achievement) on the rate of acquiring knowledge? Answers to such questions as these are expected to improve teachers' understanding of the instructional process, and to increase their facility in instruction.

A second purpose was to investigate the applicability of hierarchies to the problem of diagnosing specific deficiences (gaps) in the knowledges of individual students. This is an application of task-analysis methods to the design of instruction for the individual student, particularly the "slow" student, the student who "doesn't quite get it." The purpose was not to develop new techniques of what is usually called remedial instruction, but rather techniques which might be used by the teacher to deal with the individual, and specific, learning difficulties she encounters every day.

A third purpose was to identify with as much precision as possible the conditions of learning (cf. Gagne, 1965) of the individual components of a learning hierarchy, namely, the learning of principles. Given a particular learning background, how does a student learn and remember principles (some of which are called "facts") of the sort "Water vaporizes at 100°C"? What are the conditions in the student and in the surrounding environment, which affect the learning of principles, and how may these be optimized? It should be emphasized that such learning is to be measured by a criterion of ideational reproduction, not verbatim reproduction as is usual in studies of rote learning (Underwood, 1964b). The identification of optimal learning conditions for principles would be of use in the design of instruction, by the teacher, the textbook writer, or other agent.

Methods, Results, and Conclusions

In view of the three-fold purpose of the project, the methods for experimental study varied with the specific problem investigated. Accordingly, these methods, as well as the results and conclusions to which the studies have led, are reported as part of the description of each single investigation

CHAPTER 2 - STUDIES OF LEARNING SEQUENCES

(Study 1) Transfer Effects of Practice Variety in Principle Learning*
by
Jeanne R. Gibson

In the learning of concepts, a number of studies (Hull, 1920; Callantine and Warren, 1955; Morrisett and Hovland, 1959; Dukes and Bevan, 1967) have shown transfer to new instances of the concept class to be enhanced by variety of instances employed in the learning situation.

Principles may be conceived as combinations of two or more concepts; that is, they represent relationships among concepts. One might expect, therefore, that variety of context in which principles are presented, or variety of problem-solving situations in which principles are acquired, might have a similar effect on the transfer of principles to new application situations of the same general class. There are few studies in this field, however, and their results are not entirely consistent.

Variety in task examples involving the application of principles was studied by Gagné and others (1965) with sixth grade students learning non-metric geometry. Three different degrees of variety were employed in the examples which followed the presentation of each principle to be learned. In terms of scores on a final test covering these principles, significant effects of variety of examples were not found. It may be noted, however, that the level of mastery of subordinate skills in this learning hierarchy was not high, and this fact may have affected the results. Traub (1966) investigated variety by varying the context of examples used in a learning program for addition of integers, using sixth-grade students. The portion of the number line employed, the size of numbers, and the signs of numbers were all varied. The group instructed with heterogeneous problems attained higher scores in a final test of addition of integers. Retention and transfer were not studied.

---

*This study was performed as part of a dissertation for the degree of Ph.D. in Education, University of California, Berkeley, 1969. (See Gibson, 1969).

Related to the factor of variety in a methodological sense is the variable of number of repetitions of examples. If one intends to study different amounts of variety, he must make a decision whether to vary number of problems presented simultaneously, or independently. Thus the variable of amount of practice, or number of examples, must be considered in such investigations.

In the learning of verbal associations, the evidence that increased repetition leads to increased retention comes from many studies, as Underwood (1964) points out. Again, however, the effects of repetition in the sense of number of examples has been little studied. Gagné, Mayor, Garstens, and Paradise (1962) failed to find significant effects of this variable in a study of learning of integer addition, and a similar lack of significance for the effects of this variable was obtained in a study of non-metric geometry (Gagné and Staff, U. of Md. Mathematics Project, 1965).

Purpose of the Study

The purpose of the present study in principle learning is:

To determine the effect of narrow versus broad variety of practice examples (which constitute applications of the principle), following the initial acquisition of the principle, on achievement, retention, and on a transfer task.

To determine the effect of few versus many practice examples (applications of the principle) following initial learning of the principle, on achievement, retention, and transfer.

To ascertain whether or not an interaction exists between these variables and measured I.Q.

The specific hypotheses this study was designed to test are as follows:

1. A greater number of practice examples following initial learning will increase retention more than a lesser number of examples.

2. A broad variety of practice examples will facilitate transfer to a related task more than a narrow variety.

3. A broad variety of practice examples will facilitate retention more than a narrow variety.

4. Many examples of a broad variety will facilitate both retention and transfer more than many examples of narrow variety, few examples of narrow variety, or few examples of broad variety.

Method

The Learning Task and the Experimental Variables

Previous research suggests that hierarchies of learning sets can be derived from a final learning task, and that in such hierarchies the probability of achieving a higher order learning set is increased if the set below it has been learned.

If the topic Decimal Numbers were analyzed into its component learning sets, by defining a final task and then identifying subordinate knowledges that support the task, a section of the hierarchy looks like this:

d. Read and record a number in decimal form from a number line.

c. Identify a decimal number by indicating its location on a number line.

b. Write a number in the decimal form

a. Read a number in the decimal form.

Topics a, b, and c are the learning tasks taught in this study to prepare the subject for topic d which is a transfer task. For each topic, a, b, and c, the sequence of instruction on decimals was: 1) tenths between zero and one; 2) tenths between whole numbers greater than one; 3) hundredths; and 4) thousandths. After the original learning of subskills a, b, and c, and their measurement by subskill proficiency tests, the experimental treatment embodied in booklets of practice examples was administered.

A factorial design, 2 x 2 x 2 x 3, was used in which the variables were variety (broad and narrow), number of examples (many and few), grade level (3rd or 4th), and I.Q. level (high, medium, or low).

In this study the variable *number* of practice examples provided for the repetitions in the applications of the learning sets. There were two levels of number: 75 examples, referred to as the *many* condition, and 30 examples, called the *few* condition. The many condition practice booklets contained 25 examples for each of the three subskill learnings in the hierarchy. The few condition consisted of 10 examples for each subskill. The purpose of the variable number is to assess the effect of overlearning on retention and transfer.

The variable variety is defined in this study as the amount of diversity of context, form, and wording among the examples which involved application of the subordinate principles during the practice period. There were two levels of variety: broad and narrow. Broad variety is characterized by variation of practice examples within a broad range of contexts, forms, and wordings while narrow variety restricts the forms to those which would be used in the posttest. By separating sheer number of practice applications from variety of applications, the effect of each can be isolated in the case of both retention and transfer.

The Subjects

All the third and fourth grade children at Whittier University Elementary School in Berkeley, California participated in the study. The school arrangement was multi-aged, so that each of the five classrooms in the Intermediate Stage contained a combination of third and fourth grade children. Heterogeneous grouping was employed throughout the school. The school drew its students from the almost all-white neighborhood which is adjacent to one side of the University of California. Many parents of the children were employed by the University in clerical or para-professional positions; others worked in or owned small local businesses; a few were graduate students; there was also a sprinkling of hippie types.

From a total of 137 subjects who participated at one time or another, 90 constituted the group on which final statistical analyses were performed. The study took place over a period of several months, so cases were lost because of absences and moving, in addition to the cases which were excluded due to the child's speaking a foreign language or being designated as educationally handicapped by the school psychologist. These children were treated like other children in the various experimental groups, but were not included as a part of the total sample that was used as a source of data for the study. The various treatment groups were formed by random assignment within I.Q. levels. Accordingly, when a case was lost in one treatment, cases had to be deleted from other treatments, randomly at the same I.Q. level.

Materials

The basic guide for the lessons came from the A.A.A.S. Science - A Process Approach, Third Experimental Edition, Part Five. To meet the purpose of this study, adaptations and changes were made in the original materials.

Before any of the materials were constructed, a list of the behavioral objectives expected as an outcome of instruction was made. There was one objective for each learning set in the hierarchy.

Objectives:

a. Read numbers in the decimal form: see 3.4 and say or write "three and four-tenths."

b. Write numbers in the decimal form: see or hear "thirty-four hundredths" and write .34.

c. Identify a number written in decimal form by indicating its correct position on a number line.

Tests

The subskill tests were made up of two questions for each objective at the level being tested, i.e., tenths between zero and one, tenths falling between numbers greater than one, and hundredths from zero to any positive number. There was no subskill test on thousandths: children had difficulty drawing a number line showing hundredths; accordingly, in order to prevent frustration, thousandths were read and written but never drawn by the children. (Part of the difficulty was that they felt compelled to show every hundredth).

Both the immediate posttest and the retention test contained 15 items and were constructed in the same manner as the subskill tests except that they covered all levels of the objectives at once: tenths between zero and one, tenths between numbers greater than one, hundredths between zero and one, hundredths between numbers greater than one, and thousandths. A substitution of a hundredths item for the thousandths item was made for objective c to avoid the difficulty mentioned above.

The transfer tests were designed to test the amount of transfer the subject was able to achieve from objectives a, b, and c to a new situation d, where he was asked to read a decimal from a number line which was an instrument calibrated in decimal units. Never in the instructional periods had the child been asked to read a decimal from a number line, yet in each transfer test, he was presented with seven situations in which this behavior was requested. The instruments the subject was to read on the two transfer tests were several thermometers, barometer, metric ruler, map, graduated cylinder, stopwatch, and a metric spring scale. The first transfer test was given after the posttest measured immediate retention; the second was given after several weeks had elapsed, just before giving of the retention test.

Treatment Booklets

To administer the treatment, four experimental programs and an entirely different type of activity for the control group were composed. By definition, narrow variety in the practice booklets consisted only of those types of forms and wordings that would appear on the posttest. The narrow variety booklets were constructed accordingly by using 10 items of this type for each of the three objectives in the narrow variety-few examples (NF) booklets and 25 items for each objective in the narrow variety-many (NM) booklet.

Broad variety is defined by the examples being presented in as many forms, contexts, and wordings as possible. To do this, it was necessary to list other behaviors related to decimals, but not listed as objectives, and make practice items for these. The enlarged list of objectives contained behaviors such as writing a decimal as a fraction, writing a certain fraction in decimal form, labeling columns in a decimal number (tenths, hundredths, thousandths), telling which is larger 1.07 or 1.70, which is smaller?, changing x parts in ten (or in a hundred) to the appropriate fraction or decimal, changing dollars and cents from words to decimal form and listing coins needed to attain a given total of money. All of these additional objectives were other ways of talking about, writing, or representing decimals. To make the broad and narrow variety booklets comparable in difficulty, the same actual numerals were used whenever possible. The booklets for the broad variety-few example group (BF) contained 30 examples, 12 of which were the same as those in booklet NF except for wording, context, and format. Broad variety-many example (BM) booklets contained 75 examples, 24 of which were the same as those in booklet NM except for a different presentation.

The practice booklets, made with purple ditto, had several examples on each page. After the first page, answers to the items on the preceding page were inserted between the examples. These answers were printed in green ditto and the subject was directed to check the answer to each example immediately after he had completed it and to correct it if necessary before going on to the next item.

The practice booklets for the control group had the same outward appearance as those for the experimental groups, but consisted of word and number games totally unrelated to the topic of decimals. The word games were <u>Jumbles</u>, scrambled letters which form words when rearranged. The number game was finding groups of three adjacent digits that added to 15 in rows of single digit numerals.

Similar Jumbles were used as the "pleasant task" the children who had passed the subskills test were permitted to engage in.

Procedure

Before the study began, all the children were given the Henmon-Nelson Group I.Q. test (Form A, Grades 3-6).

From the beginning, the cooperation of the children was obtained by explaining frankly the reason for all the procedures they questioned: the pretest, subskills tests, some children doing jumbles while others were being taught, and the control group. This cooperation on the part of the children and their teachers made it possible to have the children come to another classroom for a lesson they missed because of absence, thus cutting the loss of subjects due to absence.

Pretest. A ten-item group test on decimals was given at the beginning of the instruction solely for the purpose of establishing that the subjects were not familiar with decimals; they were not. The test was not related to the tests given during the study.

Instruction. In previous studies of transfer from the subordinate learning sets to learning sets of the next higher level, the instruction in the subsets has often been from programmed material. Because each child progressed through the same program, what he could learn was held constant, but the degree of learning was not controlled. By continuing each child's learning until he passed a criterion test on each subset, the initial learning can be somewhat equated and the effects of practice variables determined more accurately. Therefore, in this study, procedures similar to those employed in computer-assisted instruction were used. The instruction was done by a teacher who presented each topic in the usual classroom manner. After each lesson the group was given a subskill test of six examples, two for each subskill, to prevent overlearning of the material by some students and incomplete learning by others. Those who made errors on the test were reinstructed while the children with perfect scores worked at an unrelated pleasant task. Both individual and group instruction was given to those who did not meet the criterion of 6/6 on the subskill test. After the second period of instruction another form of the subskill test was given only to those who had failed it the first time. The idea that each child must pass the subskill test on a topic before the whole group progressed to the next topic was adhered to within the reasonable limits imposed by the classroom situation.

The type of instruction done by the teacher was very conventional for an elementary school classroom. The experimenter went into each of the five rooms one day a week for an hour. All five classes received the same seven hours of instruction. The lessons were taught by means of guided discovery, basing each learning on what was known already. The children asked and answered questions orally. Both the teacher and the students used the blackboard and, at times, the children drew diagrams or wrote answers on a sheet of paper so they could check their answers against those of a person who gave his answer aloud, and thereby feel that they were getting a "turn" every time. Unpainted one-inch cubes, 1" x 1" x 10" rods, a 10" x 10" x 1" solid, and a large 10" cube were used as visual aids by the teacher to demonstrate fractional parts.

Treatment. The pupils were ranked according to grade and I.Q. scores, then as each group of five children was taken in order from the ranked list, the subjects were assigned randomly, one child to each experimental group. Materials used by the groups were dittoed booklets of practice problems for the five conditions.

a. Broad variety, many examples (BM)

b. Broad variety, few examples (BF)

c. Narrow variety,many examples (NM)

d. Narrow variety, few examples (NF)

e. Control group, no practice examples (C)

The treatment varied in length of time taken to complete it from one hour to four hours. After the children had spent the first class period working on the treatment booklets, the investigator noted that there were answer errors in the booklets that had not been corrected by the children as the instructions had prescribed. In order to assure each child's really having the desired experiences, it was necessary for the investigator to mark items wrong and have each child correct his booklet until he had all the answers correct. Even though the subjects had the correct answers available, knew they were to check each item as they completed it, and correct any errors, they were only eight and nine year old children and as such could not be expected to meet adult standards of performance; consequently the above adaptation was made to insure all subjects getting the appropriate treatment for their group.

Posttest and First Transfer Task. The 15-item posttest tested the listed objectives just as the subskills tests had. There were five items, each at a different level of difficulty, for each objective. In effect, the posttest was a test of achievement based on everything that was taught by the teacher; in the subsequent discussion it will be referred to as the test of immediate achievement. A seven-item transfer test was incorporated in the test booklet following the posttest. The transfer task asked the child to read decimals from a number line as is done in reading instruments such as thermometers. Pictorial representations of the instruments were used, one to each page.

Second Transfer Task and Retention Test. Several weeks later, with no intervening instruction or practice, a second test of transfer and a retention test were administered. These tests were identical in length and conceptually equivalent to the first transfer task and the posttest. The retention test will be referred to in future discussions as the delayed achievement test.

The purpose of reversing the order of the tests in the delayed assessment was to permit an evaluation of transfer which was unassisted by cues from the delayed achievement test. The subsequent retention test told how much of the original learning had been forgotten./

Results

Achievement of Subordinate Skills. A comparison of the mean scores of the subjects in the various groups on the tests of subordinate skills is shown in Table 1-1. It indicates clearly that a criterion of learning was reached and that the groups were comparable in initial learning. With a maximum score of 6 possible on these tests, one does not hesitate to state that learning of the subordinate skills was achieved when the means for all the groups fall within one-tenth of 5.5. A one-way analysis of variance performed on these scores indicated that they did not differ significantly ($F = .04$, df 4, 85, $p > .05$).

Table 1-1

Scores Analysis of Variance on Subordinate Skills Test

Experimental Group	Mean Score	df	F	P
Broad Many	5.48	4, 85	.04	N.S.
Broad Few	5.49			
Narrow Many	5.44			
Narrow Few	5.52			
Control	5.42			

Tests of Learning Outcomes, Comparison of Mean Scores. Mean scores and F values for differences among the five treatment groups are shown for each of the four tests in Table II. None of the differences between means is significant.

On the test of achievement immediately after the treatment, the narrow variety groups, which had practiced only examples similar to those on the test, had the highest means, but differences were not significant. This outcome is in opposition to the third hypothesis based on Traub's (1966) research, but it is logical to expect better performance from those who practiced similar tasks. Another prediction was that practice on many examples would result in higher achievement than practice on just a few. This was not found to be the case.

Understandably, mean scores were generally higher on the immediate achievement test than they were on the conceptually equivalent retention test administered after an elapsed time interval. The narrow variety groups have higher means on the posttest, although not significantly so. These tendencies fail to show up as differences in the retention test.

On the tests of transfer, the slight advantage of the broad variety groups is not a significant one, according to this analysis. Subjects tended to do better on the second test of transfer than on the transfer test given immediately after the initial practice. The general tendency of transfer scores to increase while retention scores decrease suggests that different mechanisms may be operating (Dukes and Bevan, 1967). The retention test required the recall of principles, whereas the transfer test assessed the application of principles to new problem settings. It may be that these processes are not governed by the same rules.

Table 1-2

Mean Scores for the Five Experimental Groups on the Various Tests

	Experimental Groups				
Tests	Broad Many	Broad Few	Narrow Many	Narrow Few	Control
Posttest	11.83	10.94	12.67	12.50	10.44
Transfer test 1	3.17	2.39	2.67	2.00	3.11
Transfer test 2	3.33	3.22	3.11	3.17	3.50
Retention test	11.61	10.72	11.33	11.89	11.00

Analysis of Variance. The major findings of an analysis of variance indicated non-significant effects of the variables of number and variety. The absence of these effects was found on the post-test, on transfer tests 1 and 2, and on the retention test.

Some few significant effects were indicated for the variable of IQ on the transfer tests and on the retention test. However these effects were not unexpected, and are not of major importance insofar as the main purposes of the study are concerned.

Summary and Conclusions

This study had the purpose of investigating the effects of certain events following initial learning of a principle on the retention and transfer of the principle. The events were: practice with different numbers of examples, and practice with examples of "broad" and "narrow" variety. The major variables in this study were amount of practice (Many or Few examples) and variety of examples (Broad or Narrow). Ninety children (third or fourth grade) of three I.Q. levels were the subjects. All the children were instructed in subordinate skills related to decimal numbers: reading decimal numerals, writing decimals, and indicating with an arrow the position of a given decimal on a number line. Subskill tests were administered following instruction. Children who met the learning criterion on the subskill tests were given an unrelated pleasant task to prevent rehearsal, while others were reinstructed. Subjects were randomly assigned to one of five

treatment groups (BM, BF, NM, NF, or C) by I.Q. level. Treatment consisted of the subjects working few (30) or many (75) practice examples of a broad or narrow variety. These decimal practice examples were administered in booklets of a programmed instruction type. The control group did no additional examples. Achievement and transfer tests were given after the practice period and again after an elapsed time interval of several weeks. The transfer task was the reading and recording of a decimal from a pictured representation of an instrument such as a thermometer, graduated cylinder, or barometer.

No significant effects were found in retention and transfer for the variables of number of examples or variety of examples. Overlearning with either broad or narrow variety was not shown to facilitate either achievement or transfer. Variety was also not a significant variable on the tests of transfer, or on the delayed achievement tests. As expected, I.Q. was a contributing variable, particularly on the transfer tests and the delayed achievement test.

An important difference between the design of this and previous studies of effects of variety on principle learning lies in the placement of the practice examples. This study placed the variety of examples <u>after</u> the original learning; others have placed it during the learning period. It may be that variety of examples is not a variable which can exert its effects once learning has initially occurred. Further investigation is needed to affirm this. Such a study could be conducted as a variant of an earlier experiment with the variety placed after learning has been completed. Alternatively, the tasks of this study could be used in incorporating practice variety in the original learning program, as a contrast to the present procedure.

The absence of any significant effects on either retention or transfer attributable to the amount of practice raises a question concerning the appropriateness of generalizing the findings of one type of learning study to another without careful consideration of the differences in learning tasks. Although overlearning facilitates retention and transfer in verbal learning situations, this study has not found an effect of additional practice, once a principle is learned. These results suggest that verbal learning and principle learning are, in at least this respect, not governed by the same rules.

(Study 2) A Study of Subordinate Skills
in Science Problem Solving*
by
Virginia K. Wiegand

The role of learning in the development of logical thinking and problem-solving remains an unresolved issue in contemporary psychology. Two contrasting views are currently predominant. One stresses the importance of innate incipient cognitive structures within the child, while the other emphasizes the effects of learning. The foremost proponent of the maturation of cognitive structures, Jean Piaget, believes that the order of cognitive development is invariant. Environmental exigencies influence the rate of development, however, through the operation of the functional invariants, accommodation and assimilation. New experiences are assimilated into the existing cognitive structures of the child and, in turn, enable the child to accommodate to the demands made by the environment.

The development of complex intellectual processes is considered by other psychologists to be an outcome of learning. Behavior is viewed as a response to stimulus changes in the environment, and learning is inferred as the learner exhibits altered patterns of responding following experience in a particular stimulus situation. A cumulative effect of learning within varied stimulus situations is the acquisition of skills, strategies, and/or learning sets that permit the individual to deal more effectively with environmental contingencies. Learning leads to more efficient responding both in similar environmental situations of equal complexity and in similar, but more complex, situations. Gagné (1969) considers problem-solving capabilities to be the result of the cumulative learning of specific intellectual skills which are relevant to a particular class of problems. The transfer of learning, in this view, plays a critical role in the development of cognitive abilities, particularly those concerned with complex thinking.

---

*This study was performed as part of a dissertation for the degree of Ph.D. in Education, University of California, Berkeley, 1969. (See Wiegand, 1969).

One way to study the relation of learning, transfer, and logical thinking is to choose a task used by Piaget in his studies of logical thinking, or one very much like it; select subjects who are unable to perform this task in a manner that demonstrates logical thinking, according to Piaget's criteria; train these subjects by means of a learning sequence to perform simpler tasks considered to be needed for the Piaget task or its equivalent; then present the complex task a second time to see whether the effects of learning transfer so that the learners can now perform the complex task. The learning sequence can be devised, using a method developed by Gagné (1962), to ascertain the subordinate skills needed to perform the complex task. Since these subordinate skills are expected to generate the potential for positive transfer to a whole class of problems, a learning hierarchy based on a problem of the same complexity as the Piaget, and involving the same operations and variables, should also be effective in generating the potential for positive transfer to the Piaget task. A careful assessment of the initial capabilities of each student, in combination with a deliberate and careful attempt to have each student learn any needed subordinate skills before presenting the complex task, can reveal information about the role of learning and positive transfer in a problem-solving situation. Transfer from one problem to another, e.g., from a comparable alternate problem to a Piaget task, can provide information about lateral transfer or generalization to another problem of the same class. Such information is of interest not only to the psychologist but also to the educator and curriculum planner faced with the problems of what to teach and how to teach it.

The major emphasis of the model proposed by Gagné (1968) is upon the cumulative effects of learning. The intellectual development of the child proceeds, not because of the increasing numbers of associations formed between stimuli and responses, but because of the learning of ordered sets of capabilities that build upon each other in progressive fashion by means of the processes of differentiation, recall, and transfer of learning. It is proposed that the entities that are learned in a learning sequence are relatively specific intellectual skills. "They are specific enough so that one must specify them by naming the class of properties of external objects or events to which they will apply. At the same time, they possess great potential for generalization, through combination with other learned entities by means of a little understood, but nevertheless dependable, mechanism of learning transfer" (ibid., p. 189).

Purpose of the Study

The success of instruction based on learning hierarchies in enabling students to perform conservation and mathematical tasks makes it seem likely that the same general procedure would be equally effective in the learning of a science task requiring, in Piaget's terms, logical thinking. Further, it is believed that the deficiencies in the ability of students to solve this particular type of science problem can best be described as absences of specific capabilities (subordinate skills), each of which can be overcome by specific learning. The individual assessment of each student's initial capabilities in order to identify the skills which he lacks, followed by a learning sequence designed to provide him with the specific learning of needed subordinate skills, by maximizing the potential for positive transfer from the attainment of a subordinate capability to a related capability higher in the learning hierarchy, should lead to the successful performance of the science problem and to others of the same general class.

The specific research questions to be investigated are the following:

Within a learning hierarchy describing specific intellectual skills subordinate to a final science problem, children's achievements will be found to differ with respect to the level of subordinate skills achieved with each branch of the learning hierarchy.

Learning of the originally non-achieved skills in the hierarchy will enable each student to perform successfully on the final task, in contrast to the performance of control subjects who have not achieved needed skills.

The effects of the initial testing, characterized by thorough explanation of test problems, will lead to the later achievement of specific subordinate skills, but will not affect achievement of the final task (except in those instances in which all subordinate skills are achieved).

Positive transfer of training will be shown on the performance of a second problem by those students who have learned the subordinate skills when compared with those who have not.

Method

Measures

The transfer task (Criterion Task 2). Selection of a complex science problem from among those used by Piaget (Inhelder and Piaget, 1958) to study logical thinking was based primarily on three factors: 1) is the problem concerned with a construct basic to many areas of science? 2) is the formulation of this construct or general relation one that requires mathematical operations (division or multiplication), so that the child cannot directly formulate the relation on the basis of immediate untransformed perceptual information? 3) is it possible to devise a similar problem that illustrates the same form of general relationship but involving different perceptual information, while still making use of some of the same independent variables? The third question has special importance in the context of the present experiment because performance measures on similar, but somewhat different, problems can shed light on the nature of transfer. The problem selected is discussed in the chapter entitled "Hauling Weight on an Inclined Plane" (Inhelder and Piaget, 1958, pp. 182-198). In this problem a toy wagon, suspended by a cable, is pulled up an inclined plane by counterweights at the other end of the cable. The cable passes over a pulley at the end of the plane. The inclination of the plane can be varied from 0 degrees to 90 degrees. The child's task is to "predict the movements or equilibrium position of the wagon as a function of three variables--the weight it carries, the counterweight suspended by a cable fastened to the wagon, and the inclination of the track" (ibid., p. 182).

This problem deals with work relationships, certainly of central concern in physics and mechanics. It was adapted for the present study first, by having the child make his observations in a specified order; second, by having him tabulate the results of his observations; third, by using vertical heights of 4, 5, and 6 inches for the inclination of the plane rather than the inclinations of 30, 45, 60, and 90 degrees commonly used by Piaget's subjects.

A plane whose inclination could be varied, a small car used alone or with one or two 10-gram weights added, a pulley fastened to the end of the plane, and six counterweights were used in the present version of the Piaget problem. The student was asked to find a rule relating 1) the height of the plane at the position of equilibrium and the weight of the car and 2) the counterweight and the length of the plane from the horizontal to the point of equilibrium. Three vertical heights of the incline and three weights

of the car were used in all possible combinations. The same position of equilibrium was used throughout. All observations were made by the student in accordance with a predetermined random order. The subject was instructed to tabulate the results of all his observations systematically and then asked to derive from them the general relation H x W = T x C, where H is the vertical height of the incline, W is the weight of the car, T is the length of the track along the incline from the horizontal to the position of equilibrium, and C is the weight of the counterweight. This relation was to be expressed either in words or symbols.

The final task. As an immediate measure of the effects of the experimental treatments, a final task was devised which used the same inclined plane. The small car weighing 40 grams, alone or with one or two 10-gram weights added, was placed at a designated place on the incline. When the car was released, it traveled down the inclined plane and struck a 20-gram block placed 2 inches from the bottom of the incline along the horizontal. The student was to derive from his measures and observations the relation between 1) the vertical height of the incline at the starting position of the car and the weight of the car and 2) the weight of the block at the bottom of the incline and the distance the block was moved after being struck by the car. All combinations of three heights and three weights were used, and the same block was employed throughout. The position of the block and the distance from the bottom of the inclined plane to the vertical support from which the height of the incline was measured were held constant over all trials. The order in which the student varied the combination of height and weight was the same for all subjects and was randomly selected. The distance moved by the block was measured and recorded by the subject on every trial. From his tabled and recorded values, the student was asked to derive and explain the general rule, H x W = M x B, i.e., the weight of the car multiplied by the height of the incline at the point from which the car was released is equal to the weight of the block multiplied by the distance that the block was moved along the horizontal. This relation could be expressed in words or symbols. Subjects who performed successfully were then asked to predict the distance the same 20-gram block would be moved if the car weighed 30 grams and was released from a position 4 inches above the horizontal.

This task was the final task of a learning hierarchy, that is, it was the basis for derivation of a set of subordinate capabilities which were independently tested.

The learning hierarchy. (See Figure 2-1). Using a method proposed by Gagné (1962) to identify subordinate skills needed to solve the final task, a learning hierarchy was constructed. This was done by asking, for each subskill revealed in the analysis, the following question: What would an individual have to be able to do to perform this task, given only instructions? Beginning with the final task, this question was asked for all subskills identified as part of the learning hierarchy. It was expected that students who could perform successfully on one subskill of the hierarchy could also perform successfully any and all subordinate skills related to this subskill that were lower in the learning hierarchy.

Once the learning hierarchy had been constructed, it was possible to devise problems to test the subject's capabilities in the subskills represented. Because it was considered desirable to assess the subskills in as general a way as possible, it was decided to vary the problems with respect to the variables of interest, the mathematical operations employed, and the organization of the data presented in the problems themselves. Only variables with which it seemed safe to assume that the child had had some prior acquaintance were used in problems to be done by the child himself. It was desired that the variety of variables, operations, and sequencing of variables would help focus attention on the operation and/or other requirements of the problems.

The subskills identified for the final task, together with behavioral objectives indicative of the attainment of a particular subskill are described in the following paragraphs. The order in which the subskills are discussed is the same as the order in which the test was given.

I. The final task. This task has already been described. Here the student is asked to demonstrate and derive the physical relationship H x W = M x B (height of plane times weight of car equals weight of block times distance moved).

II. Given a set of values relating two physical variables, the student is to place them in order in a table, derive from these values the general equation, then demonstrate the derivation of one variable given the value of the others. The student had to get two out of two problems correct to pass the test of this subskill.

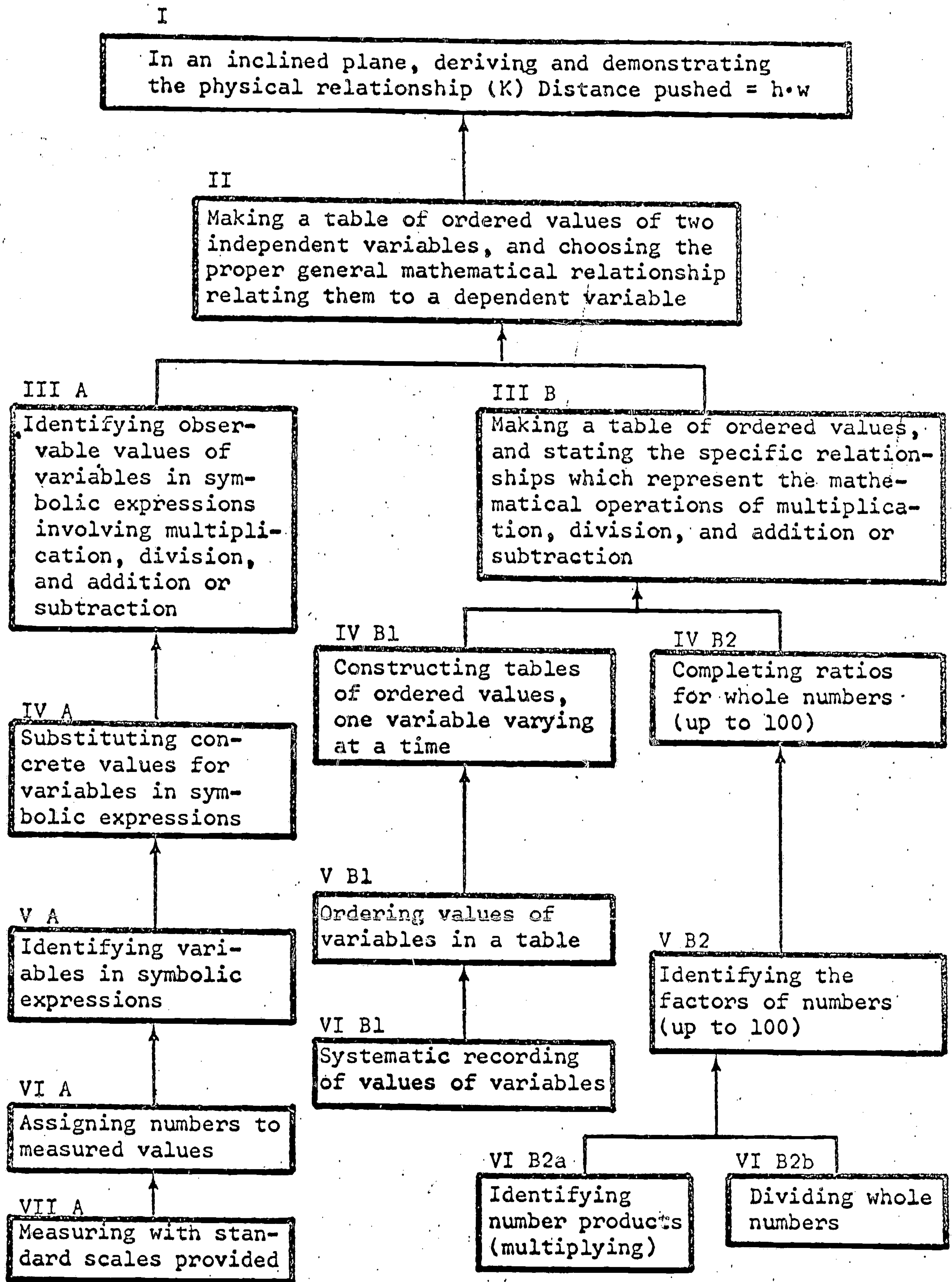

Figure 2-1. Learning Hierarchy for the Science Task.

IIIA. Given an equation stated in words, the student was to show how to find the value of the dependent variable using the equation and values for the independent variables shown in an accompanying illustration. To pass this test, the student had to have two out of three problems correct.

IVA. Given an equation and an assortment of measurements of various types of variables, the student is to substitute the proper numerical values, selected from those given in the assortment of measurements, into the equation. To demonstrate this subskill satisfactorily, the student had to get two of three problems correct.

VA. Given an equation and an illustration, the student was to label the illustration with the proper letters symbolizing designated variables. The criterion for this subskill was two of three problems correct.

VIA. Given illustrations of objects of different weights and of arrows pointing to different values on a number line, the student was to assign correct numerical values to the weights of the objects and the position of the arrows. 2 of 2 multiple-part problems had to be done correctly for the student to pass the test of the subskill.

VIIA. Given numbered scales of length, the student was to draw arrows pointing to particular values on a number line. The criterion was two of three sets of values correct.

IIIB. Given a set of values relating two physical variables, the student was to place them in a table in order and derive from these values the general equation illustrated by these values. The student had to get two of three problems correct to demonstrate this subskill.

IVB1. Given a set of values relating two physical variables, but varying values of only one variable, the student was to place them in order in a table and derive from these values the general equation. To pass this test the student had to get two of three problems correct.

VB1. Given a set of values relating two physical variables, the student was to place them in order in a table for which columns and headings were provided. To pass the test, the student had to be correct for two of three problems.

VIB1. Given a set of values relating two variables, the student was to record the values of the variables in a table

for which columns and headings were provided. Two of two problems correct were required for this test.

IVB2. Given 3 of 4 values of a ratio, the student was to complete the ratio. To pass the test, four of the five problems had to be completed correctly.

VB2. Given a number, the student was to supply the required number of factors. The student had to get two of three problems correct.

VIB2b. Given a dividend and a divisor, the student was to find the quotient. The criterion was two of three problems correct.

VIB2a. Given two numbers, the student was to find their product. To pass the test, two of three problems had to be done correctly.

All scores for subskills contained in the learning hierarchy were pass (+) or fail (-). The final task and the transfer task were scored in two ways, plus (+) or minus (-) and in points. For these tasks, 1 point was given for the correct recording and tabling of values, 1 point for the correct ordering of the values, and 1 point for the correct derivation of the general rule; a total score of 3 points could be attained on either task. Only Ss given 3 points were asked the question involving an additional application of the rule derived in the final task, Task 1. Answers to this question were also scored plus (+) or minus (-).

Subjects

A total of 30 students, 14 boys and 16 girls, enrolled in the Summer Program of the Mount Diablo Unified School District, Concord, California, took part in this experiment. The group was heterogeneous with respect to ability, with an intelligence test range from approximately 90 to 125. Two-thirds of the subjects regularly attended public schools, and one-third attended parochial school. All had recently completed the sixth grade. All were initially unable to pass the criterion tasks (the transfer task and the final task). One student was excluded from the study because she was able to perform one of the two criterion tasks on its first presentation. There were 10 subjects in each of the three experimental conditions, 1) Demonstration Test-Retest (DTR), 2) Test-Retest (TR), and 3) Test (T). All subjects were tested individually.

Materials and Apparatus

Apparatus. The same apparatus was used for both of the criterion tasks, which were the final task of the learning hierarchy and the transfer task. Both tasks required the use of an inclined plane.

Materials. A copy of every problem used in the assessment of subskills included in the learning hierarchy was prepared for each student. In addition, subjects in the DTR condition were given a copy of every problem used in the demonstrations as a part of the instructional procedure. Paper and pencils were also supplied. Problems were presented one at a time in a pre-arranged sequence until the subject reached a criterion of two consecutive related subordinate capabilities correct or until all problems had been presented. The test was designed to be given in a standard order, beginning with those problems which tested the highest-level subskill and descending, one subskill and one branch of the learning hierarchy at a time, through all subskills or until the criterion was reached. The order of testing for all the subordinate capabilities was II, III A, IV A, V A, VI A, III B, IV B1, V B1, VI B1, IV B2, V B2, VI B2b, VI B2a.

Procedure

All students were first presented with the two criterion tasks, the final task and the transfer task. Only those subjects who were unable to perform these tasks were tested further. One subject was rejected at this point in the experiment. After failures on these tasks, each student was assigned at random to one of the three experimental conditions, 1) Demonstration Test Retest (DTR), @) Test Retest (TR), 3) Test (T) (see Table 2-1).

Table 2-1

Sequence of Experimental Events

Experimental Condition	Criterion Tasks Pretesting		Testing Sequence	Posttesting
DTR	Final task	Transfer task	Demonstration test retest	Final task Transfer task
TR	Final task	Transfer task	Test, retest	Final task Transfer task
T	Final task	Transfer task	Test	Final task Transfer task

Group DTR. Each subject was given a test of subordinate skills (see Figure 3), beginning with the highest-order capability in the learning hierarchy and proceeding downward to the point at which the student was successful on tests of two contiguous related subskills or until all problems had been given. Instructions which were given included a demonstration by the investigator of an example with an accompanying explanation of each step in the example. The demonstration was followed by a test problem to be done by the subject. All test problems were read aloud by the investigator before the subject attempted to do them. A second problem at the same level of complexity was then given, and then a third, if there was one. Instructions and problems for the next subordinate capability were then introduced. This procedure was followed until the criterion was met or until all levels of the learning hierarchy had been tested.

At the conclusion of the initial assessment, each student was retested on every subordinate capability which he had been unable to demonstrate during the initial testing. Instructions were repeated, but no demonstration was given. The order of testing was reversed during the retest so that lower-order skills were tested prior to higher-order ones. Students were asked to perform the same tasks that they were unable to perform in the initial test. Students were told when they performed correctly, but nothing was said in the cases where they did not.

All subjects were asked to perform the final problem at the conclusion of the retest. The instructions for this task were explicit, i.e., they delimited and defined the problem, but they did not include a demonstration. Those students who performed the final task successfully were then asked to supply a missing value for a variable in the general rule they had just derived.

Following the final task, all subjects were asked to perform the transfer task. Instructions for the transfer task were similar to those for the final final task.

Group TR. The sequence of the test, retest, final task, and transfer task was the same as for Group DTR. The instructions and procedures were identical, except that no demonstration was given.

Group T. The procedure for Group T was the same as that for Group RT, except that subjects in this group were not given a retest.

The number and length of the testing sessions varied according to the subject's performance and the experimental condition to which he had been assigned. Testing sessions were from 30 to 45 minutes in lenth. From three to eight sessions were needed for each subject.

Results

The Ordering of Subordinate Skills

The learning hierarchy developed for the complex science task investigated in this study was based upon an *a priori* psychological analysis of subordinate intellectual skills believed to be relevant to its performance. Evidence of its effectiveness is provided by a comparison of the before-and-after final task performance of all subjects. On the first presentation of the final task, prior to any experimental intervention, none of the students were able to perform the task. Following the assessment of subordinate capabilities, 22 subjects were successful in performing the final task, a significant change yielding a X^2 value of 16.2 (df = 1; $p < .001$), according to the McNemar test for the significance of changes.

Evidence concerning the ordered dependence of subordinate skills provided by an examination of patterns of scores on subskills and a comparison of these patterns with success on the final task, as shown in Tables 2-2 and 2-3. If the subskills have been identified correctly as belonging in the learning hierarchy, and if the order of their inclusion is correct, attainment of a skill higher in the learning hierarchy should indicate the prior attainment of any and all simpler related skills. In addition, the attainment of a skill is expected to generate the potential for positive transfer to a more complex related skill. Pass-fail scores on contiguous related subskills provide an empirical indication of such positive transfer.

There are four possible pass-fail combinations for any two related subskills: 1) Lower +, Higher +; 2) Lower +, Higher -; 3) Lower -, Higher -; and 4) Lower -, Higher +. Of these, only the fourth, Lower -, Higher +, is contrary to the expectation that transfer to a higher-order skill occurs only after the attainment of lower-order skills. Pass-fail scores were computed for the subordinate capabilities in two ways; First, the pattern of responding on the initial test were examined for all three

experimental groups, DTR, TR, and T (see Table 2-2). Second, patterns of responding on the retest of subordinate skills were scrutinized for Groups TR and T (see Table 2-3). Because of the large number of zero entries in the various cells of these tables no statistical tests of significance could be performed. However, the data are overwhelmingly consistent with expectation: only 6 of the 286 pass-fail combinations show successful achievement of a higher-level skill in the absence of a contiguous lower-level skill.

Table 2-2

Number of Subjects in Each Group Displaying Specified Patterns of Success on Related Adjacent Higher- and Lower-Level Subordinate Skills, on the Initial Test

Lower to higher skill	Group DTR				Group TR				Group T			
	++	+-	--	-+	++	+-	--	-+	++	+-	--	-+
III A to II	5	4	1	0	3	7	0	0	3	6	1	0
IV A to III A	4	1	0	0	4	0	0	3	6	0	1	0
V A to IV A	1	0	0	0	0	3	0	0	0	1	0	0
VI A to V A	0	0	0	0	3	0	0	0	1	0	0	0
III B to II	5	2	3	0	3	1	6	0	3	1	6	0
IV B to III B	2	1	2	0	1	5	1	0	1	5	1	0
V B1 to IV B1	1	2	0	0	5	1	0	0	5	1	0	0
VI B1 to V B1	2	0	0	0	1	0	0	0	1	0	0	0
IV B2 to III B	2	0	3	0	1	4	2	0	1	4	2	0
V B2 to IV B2	0	2	1	0	4	2	0	0	4	0	0	0
VI B2b to V B2	2	1	0	0	2	0	0	0	2	0	0	0
VI B2a to V B2	2	1	0	0	2	0	0	0	2	0	0	0
Total	26	14	10	0	29	23	9	3	29	18	11	0
Percent in Predicted Direction	100				95				100			

Table 2-3

Number of Subjects in Each Group Displaying Specified Patterns of Success on Related Adjacent Higher- and Lower-Level Subordinate Skills, on the Retest

Lower to higher skill	Group DTR ++	+-	--	-+	Group TR ++	+-	--	-+
III A to II	10	0	0	0	10	0	0	0
IV A to III A	5	0	0	0	7	0	0	0
V A to IV A	1	0	0	0	3	0	0	0
VI A to V A	0	0	0	0	3	0	0	0
III to II	10	0	0	0	10	0	0	0
IV B1 to III B	5	0	0	0	7	0	0	0
V B1 to IV B1	3	0	0	0	6	0	0	0
VI B1 to V B1	2	0	0	0	1	0	0	0
IV B2 to III B	4	0	0	1	6	0	0	1
V B2 to IV B2	2	1	0	0	5	1	0	0
VI B2b to V B2	3	0	0	0	2	0	0	0
VI B2a to V B2	3	0	0	0	2	0	0	0
Total	48	1	0	1	62	1	0	1
Percent in Predicted Direction		98				98		

The effects of retesting, whether with or without a demonstration (as in Groups DTR and TR) was to change failure on subordinate skills to success, in all but a few instances. A comparison of scores for number of failed subskills with scores on these same subskills on the retest is shown in Table 2-4. There were 13 failures on subordinate capabilities made by subjects in Group DTR on the initial test, but only one failure on the retest. Only one failure was made by subjects in Group TR in the retest, in contrast to 19 made in the initial test. Chi-Square one-sample tests of the direction of change were significant for both Group DTR ($X^2 = 13.07$, df=1; $p < .001$) and Group TR ($X^2 = 17.05$, df=1; $p < .001$).

The retest had a significant effect on performance on the test of subordinate capabilities. Did it also affect performance on the final task when the latter was presented after the retest?

Table 2-4

Scores on Initial Test, Retest, Final Task and Transfer Task for Subjects in Groups DTR, TR and T

Subjects in Group	No. of Subordinate Skills Scored -		Score on Final Task	Score on Transfer Task
DTR	Initial Test	Retest		
1.	0	0	3	3
2.	0	0	3	3
3.	0	0	2	3
4.	0	0	3	2
5.	0	0	3	3
6.	1	0	3	3
7.	1	0	3	3
8.	3	0	3	3
9.	3	0	3	3
10.	5	1	3	3
M	1.3	0.1	2.9	2.9
TR				
1.	0	0	3	3
2.	0	0	3	3
3.	0	0	3	3
4.	1	0	3	3
5.	2	0	3	3
6.	2	0	2	1
7.	2	0	3	3
8.	4	0	3	3
9.	4	1	3	3
10.	4	0	3	3
M.	1.9	0.1	2.9	2.8
T				
1.	0		3	3
2.	0		3	3
3.	0		3	3
4.	1		1	0
5.	2		1	0
6.	2		1	1
7.	2		2	2
8.	2		2	1
9.	4		2	2
10.	5		1	0
M	1.8		1.9	1.5

Here a comparison of scores of performance on the final task of subjects in Groups DTR and RT with those of subjects in Group T should be significantly different, if the retest had an effect. In Groups DTR and TR, 19 of 20 subjects were able to perform the final task following the retest of subskills. In Group T, in which there had been no retest, only 3 of 10 subjects successfully performed the final task. The Chi-Square value for this comparison is 11.9 (df = 1; $p < .001$): it can be stated that the retest led to improved performance of the final task.

As Table 2-4 indicates, the three subjects of Group T who did perform successfully on the final task were precisely those who had no missing subskills. A retest on the final task was all that was needed to enable them to perform correctly. As control subjects, they must be counted along with those subjects in Groups TR and DTR who had no missing subskills. However, it is noteworthy that the change in their performance is in a strict sense in accord with the hypothesis under investigation.

Performance on the Final Task and the Transfer Task

Besides being scored on a pass-fail basis, final task performance was also scored in points, and the results are indicated in Table 2-4. One point was awarded for correct observations and tabling, one for the correct ordering of data in a second table, and one for the correct derivation of the general rule. All subjects were able to make correct observations and table the data correctly, so that differences in performance between Group T and Groups DTR and TR must be due to differences in ordering values in a table and/or in deriving the general rule. A significant difference in ordering was found ($X^2 = 6.88$, df = 1; $p < .01$); subjects in Group T were correct less often than subjects in the other groups. There was also a significant difference in the derivation of the correct relation between Group T and Groups DTR and TR ($X^2 = 14164$, df = 1; $p < .001$). Both the ability to order data in a table and to derive the correct relation from the table contributed to the obtained differences in performance on the final task.

Like the final task, the transfer task was also scored in points, on the same basis. A Chi-Square test of the difference between Group T and Groups TR and DTR showed a significant difference in respect to the correct ordering of data ($X^2 = 6.88$, df = 1; $p < .01$). The difference between the total score on the

transfer task for Group T and Groups DTR and TR was also significant (X^2 = 25.43, df = 1; $p < .001$). The capability to order data and the ability to derive a rule correctly from such data both contributed to the obtained differences in performance on the transfer task.

Summary and Conclusions

The present study was designed to test whether the performance of a science problem similar to one used in Piaget's studies could be accounted for on the basis of a cumulative learning model. If such problem solving is an outgrowth of the cumulative effects of the learning ordered sets of capabilities that build upon each other in progressive fashion, then it should be possible 1) to identify the subordinate capabilities involved in performing a complex task, 2) to demonstrate that children unable to perform this task can perform it once they have acquired needed subskills, and 3) to show that the capability to perform the complex task transfers to another similar complex task.

Two tasks requiring the derivation of a general rule describing a multiplicative physical relationship illustrated in a concrete situation were selected. In the Piaget task, (the transfer task) the child was to derive the relationship between 1) the height and weight of a car on an inclined plane and 2) the distance from the bottom of the incline to the equilibrium position of the car and the counterweight attached to the car by means of a cord. In the complex science (the final task) the relationship between 1) the height and weight of a car on an inclined plane and 2) the weight of a block at the bottom of the incline and the distance moved by the block when the car was released and hit it was to be derived.

A learning hierarchy of subordinate capabilities was constructed for the complex science task, together with problems used to assess the subskills.

There were three phases in the experiment: pretesting on the final task and the transfer task, the assessment of subordinate capabilities, and posttesting on the final and transfer tasks. Thirty subjects approximately 12 years of age who failed both problems were assigned to one of the three experimental conditions: Demonstration Test Retest (DTR), Test Retest (TR), and Test (T).

In the DTR condition, a demonstration of an example with an explanation of steps in its solution was included in the instructions during the initial test of subordinate capabilities. The initial test was followed by a retest, without demonstrations, of those capabilities failed in the initial test. In the TR condition, an initial test and a retest were given; there were no demonstrations. In the T condition, only the initial test was given.

All subjects were then given the final task and the transfer task. Scores on the initial test and the retest show that higher-order skills indicate the prior attainment of simpler related skills.

The demonstration had no significant effect on performance on the initial test of subordinate capabilities. The initial test alone did not enable subjects to perform the final task or the transfer task, except when they had already attained the needed subskills, as revealed by their performance on the test. The retest of subordinate capabilities failed in the initial test enabled subjects to acquire needed subskills. The attainment of needed subskills enabled subjects to perform the final task successfully. Subjects successful in performing the final task were, in all cases, successful in performing the transfer task.

Children who could perform neither a Piaget task (the transfer task) nor a very similar complex science problem (the final task) learned, in a relatively short period of time, to perform both tasks, providing they learned any needed subordinate capabilities in the interval between the first and second presentations of the final task and transfer task. The results indicate that intellectual development has been brought about by the cumulative effects of the learning of concretely referenced intellectual skills, rather than by the adaptation of structures of intellectual growth.

CHAPTER 3 - HIERARCHIES AS DIAGNOSTIC TOOLS

(Study 3) Transfer of Learning in a Social Studies Task of Comparing-Contrasting*

by

Lois T. Coleman
Nebraska Wesleyan University

and

Robert M. Gagné
Florida State University

Considerable interest has centered in recent years upon the possibility of a "process" emphasis in the teaching of social studies, as opposed to a "content" emphasis. This trend of thought is represented in Fenton's (1967) book, as well as in a number of efforts of curriculum planners and designers, most of which do not yet appear in formal publications (cf. California Statewide Social Sciences Study Committee, 1968). It is generally agreed that a "process" strand is probably only one of several kinds of emphasis to be sought by curricula in the social sciences, but that it is nevertheless of considerable importance. The model employed in certain elementary science curricula, such as Science-A Process Approach (A A A S Commision on Science Education, 1967), has appealed to many as worthy of exploration and tryout, when suitably adapted to the requirements of instruction in the social sciences.

An emphasis on process in social studies can be taken to mean, in brief, an attempt to have students learn a variety of relevant intellectual skills (as opposed to information) which make it possible for them to undertake problem solving activities in the social realm (Gagne, 1970). Some writers would prefer the term inquiry skills, which we consider to be a synonymous phrase, so far as present purposes are concerned.

---

*This study was performed as part of a dissertation for the degree of Ed.D., University of California, Berkeley, 1969. (See Coleman, 1969).

What are the intellectual skills relevant to inquiry in the social sciences? It would appear that these skills can be identified by analyzing the kinds of inquiry tasks that represent instructional objectives in modern social studies curricula, with the aim of determining for each type of task the subordinate intellectual skills which support it. Specifically, this would mean developing a set of hypotheses concerning those learned capabilities which mediate positive transfer among themselves, and to the final inquiry task, so as to form a learning hierarchy (Gagné, 1969); and then testing these hypotheses in some systematic way. Confirmation of such hypotheses would be expected to lead to identification of the capabilities (intellectual skills) that are necessary to assure the competent performance of problem-solving activities by students in various areas of social studies.

The present paper reports an effort to apply this method to one particular type of problem-solving task, <u>comparing</u>. Such tasks appear with relatively high frequency in modern social studies curricula. Typically, students first learn basic <u>information</u> concerning some area of knowledge of social phenomena, say, for example, trade among countries. Following this, they learn to use <u>defined concepts</u> (cf. Gagné, 1970) relevant to this area, such as manufactured products, imports, exports, balance of trade, etc. Often, such instruction is then followed by learning of <u>rules</u> relating these concepts, such as the rule represented by the statement, "A favorable balance of trade results from an excess of exports over imports." What such instruction leads to, it is hoped, is the successful performance by students of <u>problem solving</u> tasks, of several varieties, in the same general area. Such tasks are typically introduced in the form of questions posed by the teacher, which become the basis for class discussion of student projects. One kind of problem, with which we are here concerned, might be introduced by the question "How would you compare and contrast the exports of Country X and Country Y?"

We first undertook to make an analysis of the comparing task by identifying the subordinate intellectual skills which contribute to its performance. In accordance with the description given by Gagné (1970), we asked the question "What would the student have to know how to do in order to learn this capability most readily?" The question was applied successively to each capability identified, so that a hierarchy of intellectual skills could be described. Of particular importance in

the conduct of this analysis was the rigorous avoidance of the tendency to answer the question by identifying information (what the student knows) as opposed to intellectual skill (what the student can do). This tendency is much less easy to shun in a subject like social studies than it is in a subject like mathematics, which is virtually information-free. It should be noted that we were interested in making the assumption, not that the student of social studies needed no information, but rather that he would have the required information readily available either from recall or in task directions. The point is, if one sets out to identify process skills, he must carefully avoid becoming distracted by content.

Having derived the learning hierarchy, we set out to test it, in a group of sixth graders. First, we identified twenty sixth-grade girls who were unable to perform a comparing-contrasting task pertaining to the exports of two countries. Half of these students were then assigned randomly to an experimental group, half to a control group which was to be given no instruction. The subordinate skills of the hierarchy were then tested individually in the children of the experimental group, and they were subsequently given instruction on those intellectual skills that were "missing." They were then presented with an alternate form of the comparing problem, followed by two additional comparing tasks involving different content, one a "near-transfer" problem, the other a "far-transfer" problem. The same tasks, including the transfer tasks, were also given to the children who constituted the controls. Thus, we were interested, first, in obtaining internal evidence concerning the adequacy of the hypothesized hierarchy of skills, to be supplied by evidence of dependence of higher-order skills upon lower-order ones. Second, we wanted to determine whether the learning of the subordinate skills we had identified would lead to successful learning and performance of the final comparing task. And third, we wished to determine the extent to which such learning would be transferable to other problems of the comparing sort.

Method

Analysis of the Comparing Task

The task chosen as the focus of interest, for which prerequisite skills were to be identified, was as follows: "Given a detailed list of exported items and their monetary value for

two countries, formulate a summary comparison of the exports of the two countries in terms of a limited number of categories which are exhaustive and exclusive." By asking the question about subordinate skills described previously, a set of 21 such skills were defined. The structure of the learning hierarchy derived in this manner is illustrated in Figure 3-1. It appeared to have three major "branches", whose components may be described as follows:

A. Quantitative Comparisons

IA. Translate as 'same' or 'greater' the quantitative values of categories, from summed values for two category-fields.

IIA 1. Use 'greater' when one number is twice or more as large as another number, and 'essentially the same' when neither number is twice or more as large as the other.

IIA 2. Add columns of as many as five 5-digit numbers.

IIIA 1. Distinguish pairs of numbers in which one number is twice the other.

IVA.1. Compute two times a whole number.

VA 1. Distinguish the smaller from the larger number in a pair.

B. Generating Conceptual Categories

IB. Perform a check to test whether constructed categories are exhaustive (self-editing).

IIB. Reformulate or revise categories across two category-fields to achieve optimal distinctiveness while maintaining limited numbers of categories.

IIIB. Formulate categories which are mutually exclusive and limited in number.

IVB. Generate three-level superordinate concept categories.

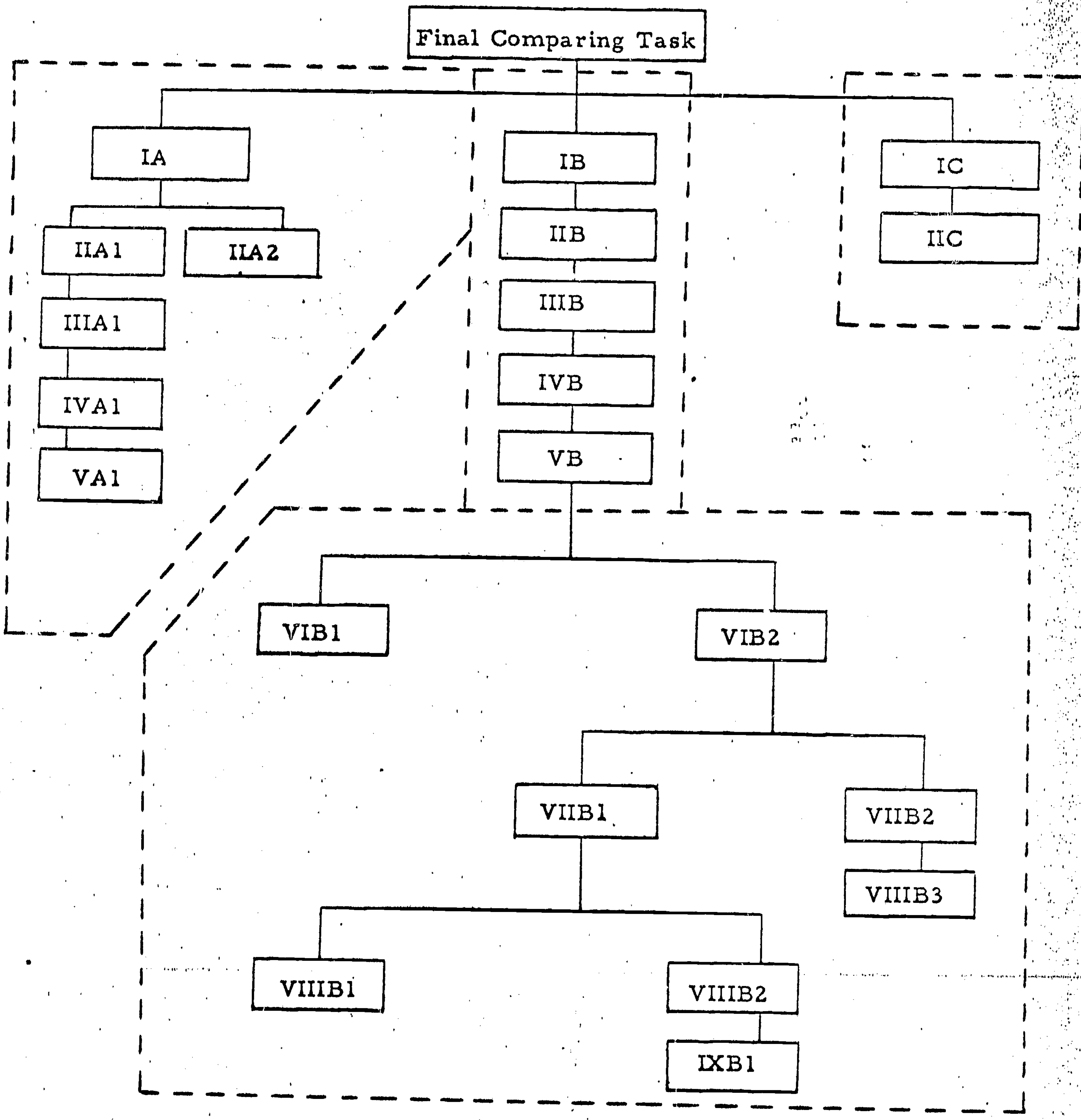

Figure 3-1. Outline of Hierarchy Chart for the Final Comparing Task and Subordinate Intellectual Skills.

V B. Use three-level superordinate category labels.

VI B1. Use rule of thumb about convenient number of categories.

VI B2. Test the mutual exclusiveness of categories by comparing operations which define them.

VII B. Construct operational definitions of superordinate categories.

VIII B1. Generate superordinate categories.

VIII B2. Given objects defined by their class, transform the definitions of categories to include statements of their relations.

IX B1. Distinguish thing-words from relation-words.

VII B2. Distinguish, in terms of their relations (verbs, prepositions) things which are sometimes alike.

VIII B3. Use relational terms (prepositions, verbs).

C. Table Interpretation

I C. Summarize information in a two-way category table.

II C. Read direct data in a two-way category table.

Each of these subordinate skills was capable of being independently represented by one or more tasks, each with its own set of verbal directions. It should be emphasized that tasks constructed in this way may pertain to any area of knowledge, and thus can be entirely unrelated in content among themselves.

Materials

The Final Comparing Task contained two problems requiring the comparison of the exports of two countries. All the problems were designed to be conceptually equivalent. The task was presented in test booklets, each of which contained an identical introduction giving an explanation and definition of the process

of comparing, as well as a definition of "exports." Each booklet contained an identical Sample Problem of the same class as the test problems, which was explained and worked through to solution. Following this were the two problems for the student to solve.

Diagnostic Tests were constructed to cover all of the 21 prerequisite skills. These were preceded by an introductory section identical for both pre- and post-testing forms. For each skill a sample problem was given, worked, and the answer recorded; this was followed by the two problems for the student to do. The problems had a variety of content, and in all cases were unrelated to exports. An instruction booklet was used to accompany the administration of problems for each prerequisite skill. These contained material pertaining to a variety of concrete physical objects, for example, the moon, bracelets, electric mixers, etc., which differed from the content of the tests themselves. A few actual household items were used in some instances to supplement this material.

For the purpose of measuring transfer of learning, Near Transfer and Far Transfer Tasks were also constructed. The Near Transfer Task contained an introduction and two comparing problems, one deriving its content from ecology, the other from medical research. The Far Transfer Task, following an introduction, presented two complex comparing problems from the area of human behavior. These problems required the student to accomplish comparisons within a category and also across many categories. It was considered, therefore, that these problems were not simply different in content, but were in all likelihood "higher-level" problems which required some prerequisite skills additional to those represented in the learning hierarchy.

Instruction

Instruction was given to each experimental subject on each prerequisite skill in which she had shown herself to be deficient. No instruction was given on the final comparing task. Instruction had the specific aim of making it possible for the subject to acquire the missing skill, recall it, and use it in producing a solution to the problem presented.

Specifically, instruction provided: (1) a problem which could be solved in a variety of specific ways; (2) sufficient verbal directions so that the subject was aware of the variety of choices and their relationships to solution; (3) a wide variety of experiences to develop the student's ability to distinguish between relevant and irrelevant material; (4) sufficient training so that the skill was readily recallable; and (5) opportunities for verbalization.

Instruction on concept categorizing skills was designed to teach the student to use, generate, and understand rule-consequence relationships. Stress was laid on the student's awareness that she had a repertoire of bases for categorizing, and on the hierarchical nature of concept categories. Concrete household items were used first as illustrations, followed by the student's written responses of categorizing. The instructor's frequent request of the student was, "Tell me what happens when you do that." Students were considered to have mastered the intellectual skill when they were able to pass two test problems consecutively.

Subjects

The subjects were sixth grade girls from a single pod in the Livermore, California Unified School District. Twenty subjects were randomly selected from a larger group of girls who scored above the 75th percentile on both the verbal and non-verbal parts of the Lorge-Thorndike Intelligence Test. The range of IQ's was from 111 to 138. The twenty students were assigned randomly to either the experimental or the control group, ten in each. A student so assigned was replaced by another, randomly chosen from the larger group, when it was discovered that she was able to perform the Final Comparing Task (one out of two problems). Such replacement was made in only one instance.

Procedure

Control subjects were given the pre-test of the Final Comparing Task, the post-test on the Final Comparing Task, the Near Transfer Task, and the Far Transfer Task on four consecutive days.

Experimental group subjects were given the pre-test of the Final Comparing Task, followed by the Diagnostic Tests. On the next following school day, they entered the instruction phase of the experimental treatment. Instruction began with a review of the student's diagnostic test performance. The student was then instructed on the learning hierarchy, so that she would understand the purposes of instruction. The period of instruction varied for each subject, depending on how many prerequisite skills were to be learned.

Upon completion of the instruction, the post-test versions of Diagnostic Tests were administered. On the following day, the post-test version of the Final Comparing Task was given. Before opening the booklets for this task, students were asked to recall what they had learned how to do during instruction,

to state what aspects were important to remember, and where their mistakes were likely to occur. On the two following days, the Near Transfer and Far Transfer Tasks were administered.

Subjects met with the experimenter individually in a portion of the teachers' workroom within their pod area. They worked with the experimenter approximately one hour per day on consecutive school days, until all aspects of the individual's program were completed. With few exceptions, the subjects were given their respective programs alternating between experimental and control.

Results

There are two major portions of the findings to be reported. First to be described are data pertaining to the performance of subjects in the experimental group on the diagnostic tests given before and after instruction. Second, there are results of a comparison between the performance of instructed and uninstructed (control) subjects on the Final Comparing Task and the transfer tests.

Performance on Diagnostic Tests of Prerequisite Skills

The diagnostic tests for the 21 prerequisite skills in the hierarchy contained two items for each skill. A rigorous criterion for passing was adopted--namely, successful completion of both problems. Performances of the ten subjects in the experimental group are exhibited in Table 3-1.

For diagnostic purposes, one would expect to find pass-fail patterns in pretesting which are regular and ordered. The hypotheses underlying the hierarchy lead to the expectation that each individual student, having failed the Final Comparing Task, will exhibit a non-pass performance on a sequence of prerequisit skills extending downward in the hierarchy until some point is reached when he does pass. Thereafter, skills lower in the hierarchy should all be passed.

The evidence for such ordered patterns of skills in the present study is not outstanding, as an examination of Table 3-1 reveals. Primarily, it appears that the group of subjects tested did not exhibit a number of deficiencies great enough to permit a good verification to be made of the ordering of prerequisite skills contained in the proposed hierarchy. The largest number of prerequisite skills scored "not-passed," using the rigorous criterion described, was 10, and the average was only 6.9.

Table 3-1

Prerequisite Skills Passed and Not Passed by Subjects in Experimental Group on Pretest (First Column) and.Post-test (Second Column).

Skills	Subjects 1	2	3	4	5	6	7	8	9	10
I A	++	++	++	++	-+	++	++	++	-+	++
II A1	++	++	++	-+	-+	++	++	-+	-+	++
II A2	++	++	++	--	+-	-+	++	++	++	++
III A1	++	++	++	-+	++	-+	++	++	++	-+
IV A1	++	++	++	++	++	++	++	++	++	++
V A1	++	++	++	++	++	++	+[illegible]	++	++	++
I B	++	++	++	-+	+-	-+	-+	-+	++	-+
II B	-+	-+	-+	-+	--	-+	-+	+-	--	-+
III B	-+	-+	-+	-+	--	-+	--	-+	-+	-+
IV B	-+	+-	++	-+	++	++	-+	++	++	-+
V B	++	-+	++	++	++	++	+	-+	-+	++
VI B1	-+	-+	-+	++	++	++	+	-+	+	++
VI B2	-+	++	++	++	-+	-+	+	++	+	++
VII B1	--	+	+	++	++	++	+	+	+	++
VII B2	++	+	+	-+	++	++	+	+	+	++
VIII B1	-+	+	+	+	++	++	+	+	+	++
VIII B2	-+	+	+	+	-+	-+	+	+	+	-+
VIII B3	++	+	+	+	+	+	+	+	+	+
IX B1	++	+	+	+	+	+	+	+	+	+
I C	-+	-+	-+	++	-+	-+	-+	-+	++	++
II C	-+	--	-+	-+	-+	-+	--	-+	+-	++
Total Not Passed	10;1	6;2	5;0	9;1	8;4	7;0	6;2	7;1	5;2	6;0

Part A of the hierarchy, Quantitative Comparisons, shows the smallest frequency of occurrence of "missing" skills. For the students in the experimental group, these were skills that by and large were already present and available, once the intensive verbal directions and examples of the tasks were presented.

Part B, Generating Conceptual Categories, shows some evidences of ordering of prerequisite skills, as may be seen by examining the patterns of "pass" and "not pass" in the first columns under each subject, in Table 3-1. For example, the sequence II B - III B - IV B - V B shows some evidence of ordering, as does also the sequence VI bl - VI B2 - VII Bl (for definitions of these skills, see previous list). At the same time, it is notable that I B, as tested in this study, does not appear to be in correct order, since it is frequently passed when hypothetically lower-order skills are not passed. I B should be a kind of "capstone" skill, involving self-editing and revision of formulated conceptual categories; one suspects in this instance some inadequacy in the test problems, rather than in the sequence suggested by the hierarchy. A different situation is encountered in skill VI Bl (using rule about a convenient number of categories), which is not passed with some frequency, whereas the hypothetically higher skill V B (using three-level superordinate category names) is frequently passed. In view of the differences in kind of activity of these two skill instances, one is inclined to suspect that VI Bl has not been correctly placed in the hierarchy.

Evidence of ordering of prerequisite skills is absent in the third branch of the hierarchy, Part C, Table Interpretation. The most frequent pattern is that of failure to pass both skills, I C and II C. The existence of these deficiencies is of interest, however, in their bearing on the successful accomplishment of the Final Comparing Task.

Learning of prerequisite skills. The results in Table 3-1 provide evidence that the prerequisite skills measured by the diagnostic tests were to a large extent effectively learned by students in the experimental group. Changes in the individual skills can be noted by comparing the first and second columns listed under each subject. The average number of skills not passed on the pretest was 6.9; the average number on the post-test, 1.3. As can be seen in the table, the distributions of these pre- and post-test scores do not overlap.

Comparison of Experimental and Control Groups

Final Comparing Task. Like the tests for prerequisite skills, the tests for the final comparing task were scored in accordance with a rigorous criterion; two problems correct meant pass, and any problem missed meant not-pass. As it turned out, however, using a different scoring criterion would not have changed the results, since subjects were able either to do both problems or none at all.

The results on the effects of instruction on prerequisite skills are quite clearcut: all of the subjects tested of the instructed (experimental) group passed the Final Comparing Task in its post-test form, whereas none of the control subjects were able to do so. (It should be noted that because of an unplanned contingency, 9 out of 9 experimental subjects were given the post-test, rather than 10 out of 10). In other words, the results pertaining to effects of the experimental treatment are: Experimental Group - 100%; Control Group - 0%.

Transfer tasks. In the case of the near- and far-transfer tasks, an attempt was made to obtain a score for each subject which graded his performance in terms of the number of component activities of the total task which were correctly performed. Points assigned to these activities were such as to provide a total score of 36 for the near-transfer task (both problems), and a total score of 46 for the far-transfer task. Means and standard deviations of the scores on these tasks for experimental and control groups are shown in Table 3-2.

Table 3-2

Means and SDs of Total Scores for Experimental and Control Groups on Near and Far Transfer Tasks (N=10 in Each Group).

Group	Task: Near Transfer M	Near Transfer SD	Far Transfer M	Far Transfer SD
Experimental	34.5	2.12	40.2	3.40
Control	25.2	8.83	14.8	7.50

The transfer scores for the experimental group of subjects are higher than those for the control subjects on both the near- and far-transfer tests. Student's t test, corrected for small sample size and unequal variance, was used to test these differences. For the near-transfer scores, $t = 2.943$, $p < .01$; and for the far-transfer scores, $t = 9.7057$, $p < .01$. Evidence is thus provided that the learning of prerequisite skills has been effective not only for the accomplishing of a given comparing task, but also for transfer of learning to other comparing tasks of considerably different content.

Conclusions

The results of this study have provided evidence that the performance of comparing-contrasting tasks, frequently employed in social studies instruction, is markedly affected by the attainment of prerequisite intellectual skills. Students who learned prerequisite skills of this sort became capable of completing a problem of comparing the exports of two countries, of which they were not capable before they learned the prerequisite skills. An equivalent set of students who did not learn the prerequisite skills continued to be unable to perform the comparing task successfully.

Similar contrasts were obtained between instructed and uninstructed subjects on near- and far-transfer taks presenting problems of comparing in entirely different content areas (ecology, medical research; human behavior). Significantly greater amounts of transfer were obtained on both these tasks for the students who had learned prerequisite skills than for those students who had not.

Evidence was sought that the prerequisite skills identified were arranged hierarchically, so that the results of diagnostic testing for the presence of these skills would reveal ordered patterns of capability in the students. Little evidence on this question was obtained in this study, although ordered skill patterns are suggested by one "branch" of the hypothesized hierarchy, Generating Conceptual Categories. The lack of evidence in this area seems primarily attributable to the infrequency of "missing" prerequisite skills among the subjects of the study.

The demonstration of effectiveness of instruction on prerequisite skills, whether ordered or not, is relevant to the general problem of instructional planning and design for social studies areas of the school curriculum. Comparing-contrasting tasks occur with considerable frequency in modern social studies curricula. It is notable that the intellectual skills which underlie successful solving of problems of the comparing type do not themselves necessarily have a "social studies content." They are the kinds of skills sometimes referred to as process skills, in the sense that they reflect what students are able to do, rather than what they "know." In this investigation, these skills were identified as falling into the three general categories (1) making quantitative comparisons; (2) generating conceptual categories; and (3) interpreting tabular data. The first category appeared to be one in which the sixth-grade girls in this study were already accomplished. Instruction in the remaining skill categories, however, made a considerable difference in their ability to solve successfully comparing problems in social studies.

(Study 4) Revision of a Science Topic Using Evidence of Performance on Subordinate Skills

by

James R. Okey
University of California, Santa Barbara

and

Robert M. Gagné
Florida State University

A cumulative learning model has been described by Gagné (1962,1968) which proposes the structuring of learning materials so as to efficiently promote their achievement. In this model, Gagné proposes that the attainment of subject-matter competence is hierarchically organized; that is, for students to master certain intellectual skills it is on the average more efficient for them to have previously acquired certain relevant subordinate skills.

The analysis procedure used to determine intellectual skills subordinate to some final task is called "deriving a learning hierarchy." To derive a learning hierarchy for some task the question is asked, "What would the student have to know how to do already in order to learn this task with a minimum of instruction?" One or more subordinate tasks may be identified in this way. The same question can then be asked of the subordinate tasks and thus lead to the identification of additional lower-order subordinate tasks. Carrying out such an analysis leads to a learning hierarchy which, by hypothesis, includes the tasks needed to bring about efficient learning of the final task.

A learning hierarchy derived in this manner can serve as a sort of instructional map for curriculum development (Gagné, 1967). After deriving a learning hierarchy one can identify which tasks hypothetically need to be mastered by students, and also their probable sequence. If the instructional materials are not successful in promoting mastery of the final task, a learning hierarchy has several built-in diagnostic features. Following instruction, a test on each task in a learning hierarchy can be used to locate places of difficulty. It may be found, for example, that students who fail a specific task are the same students who have failed certain relevant subordinate tasks. The procedure thus becomes one of finding trouble spots or places of low mastery in subordinate skills (Okey, 1968).

It is also possible to locate sequence irregularities in instructional materials with the use of a learning hierarchy. Suppose, for example, that on a pair of tasks hypothesized to be sequentially dependent, it is found that many students who achieve the higher task have failed the lower-order task. This suggests the possibility that the tasks are not subordinate to one another in the way hypothesized in the original learning hierarchy. A reordering of tasks and the related instructional materials may then be warranted.

These diagnostic procedures make possible an empirical approach to instructional development that is based on the cumulative learning model. If it is assumed that the model accurately describes how intellectual skills are acquired, then any inability to perform a final task should be attributable to failure on skills subordinate to the final task. One can also predict that a way to improve performance on a final task is to provide additional instruction leading to attainment of subordinate skills failed by a substantial number of students.

The purpose of this study was to investigate this latter prediction. The question studied was whether adding instruction on those subordinate skills failed by many students using an initial version of a learning program improved the performance of students using a revised version.

Method

A group of students studied an initial version of a learning program followed by tests on the criterion task and on subordinate skills failed by many students. A second student group studied the revised program and took the same tests. The group performances were compared to see if deficiences found in the initial program had been corrected and whether improved performance occurred on the criterion tasks.

Selection and analysis

An existing instruction program* on solubility product calculations was selected as the initial learning program. After completing this program, students were expected to (a) compute the solubility product constant given the molecular or ionic concentrations, (b) compute the molecular or ionic concentrations given the solubility product constant, and (c) (c) compute the ion product in order to predict whether a precipitate would form.

The analysis procedure previously described was used to identify subordinate skills needed to achieve the final task. Fifteen such subordinate skills were identified and are shown in Figure 4-1.

Learning programs

The materials in the learning programs were stenciled on 8 1/2 x 11 sheets of paper and folded so that a frame was on the front of the sheet and its answer on the back. The folded pages were held together with two-1 1/4 inch diameter rings.

---

*Permission to use a portion of the program entitled "Equilibrium Calculations: Solubility Product" from the book Understanding Chemistry by Gordon M. Barrow et al. was kindly granted by W.A. Benjamin, Inc. and Gordon M. Barrow.

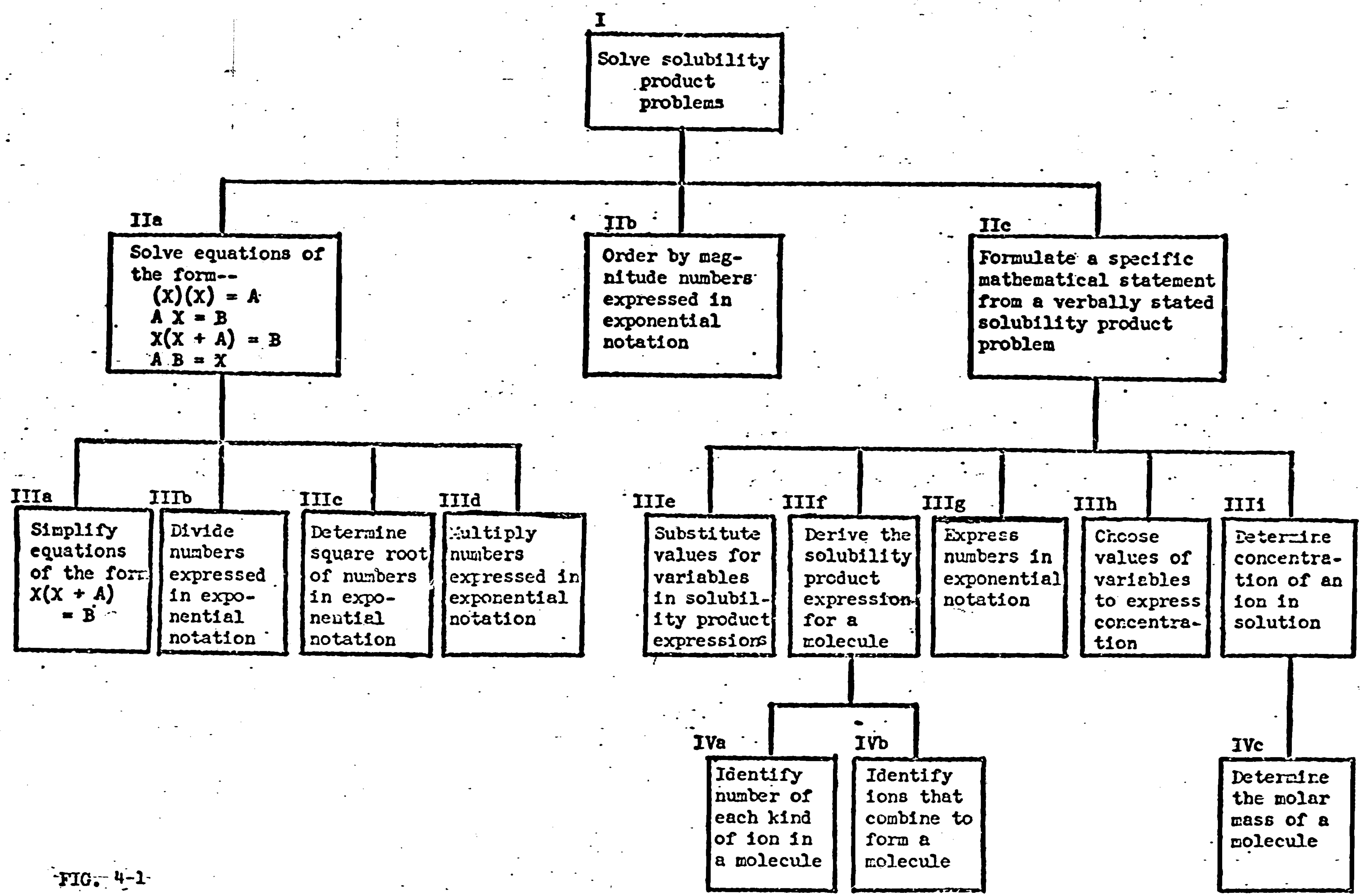

FIG. 4-1

A Learning Hierarchy for a Science Topic

Performance measures

Four different tests were used to measure student performance: a pre- and post-test on the criterion task (C Pre-Test and C Post-test) and a pre- and post-test on the subordinate skills in the learning hierarchy (SS Pre-Test and SS Post-Test).

C Pre-Test and C Post-Test. Each of these equivalent-form tests consisted of seven written problems on solubility product calculations. Six pairs of test items were tested for equivalence by chemistry students at a local high school. A pair of items can be considered equivalent if a student either passes or fails both of them. If the number of students who pass or fail both members of a pair is divided by the total number of students tested, an index of equivalence is obtained. The index of equivalence for the six pairs of items ranged from 0.84 to 1.00 with a mean of 0.90, and was considered satisfactory in each case. The test items from each pair were then assigned at random to the pre- and post-tests. A seventh test item was added later to each test, for which no index of equivalence was obtained.

SS Pre-Test and SS Post-Test. Each of these equivalent-form tests consisted of 32 items covering the 15 subordinate skills. From one to four items were used to test each subordinate skill. The two equivalent forms were produced in the same way as described for the criterion tests. The indices of equivalence for these items ranged from 0.81 to 1.00 with a mean of 0.95.

Subjects

The subjects were 106 tenth, eleventh, and twelfth grade students from five chemistry classes.

Assignment to experimental groups

The names of the students in each class were placed in alphabetical order, and alternate names on each class list were assigned to two treatment groups. In this way, 68 and 67 students, respectively, were assigned to two groups. Attrition resulting from absences and from the unavailability of IQ and mathematics percentile scores resulted finally in 49 and 57 students, respectively, in the two treatment groups.

The means and standard deviation for groups 1 and 2, respectively, on four pretreatment measures were as follows: IQ (M = 121.1, SD = 11.0; M = 118.0, SD = 12.4), mathematics

percentile ($M = 87.7$, $SD = 14.0$; $M = 87.9$, $SD = 14.5$), C Pre-Test ($M = 0.86$, $SD = 0.98$; $M = 0.16$, $SD = 0.45$), and SS Pre-Test ($M = 17.5$, $SD = 5.0$; $M = 15.7$, $SD = 3.9$). The most recently available IQ scores were used for each student. For most students this came from the Lorge-Thorndike Intelligence Test administered in tenth grade. The mathematics percentile scores were taken from the quantitative thinking subscore on the Iowa Test of Educational Development administered in either ninth or eleventh grade.

Administration of learning programs

Approximately seven class periods of 50 minutes each were required for a treatment group to take the tests and complete the learning program. The criterion and subordinate skills pre-tests were given on the first two days followed by three days on the learning program and finally the criterion and subordinate skills post-tests on the last two days.

Each student was given a learning program and told how to write in the answers and check them for correctness. The students were told that they could work at their own pace but that they ought to try and finish the program in three class periods. Because the revised program was longer, approximately one-fourth of the second treatment group arranged to come in during a study period in order to complete the program in three days.

While the first group participated in the experiment, the second group did a series of laboratory investigations on titration and acid-base indicators. Twelve days after the first group finished, the second group followed the same seven day procedure except that they used the revised learning program. There were only three school days in the twelve days between the two treatments because of weekends and Easter vacation. On the three days when school was in attendance and no experiment was in progress, the entire class worked on matters unrelated to solubility product calculations.

Results

Revision of the preliminary program

The initial version of the program proved to be only moderately effective; the mean score was 3.02 ($SD = 1.94$) for the 7-item C Post-Test. Item scores on the SS Post-Test were then examined to identify subordinate skills which had been failed by many students.

Table 4-1 shows the percentage attainment on items used to test 3 of the 15 subordinate skills. These items are used to illustrate how achievement on subordinate skills revealed specific difficulties in the learning program.

The data for subordinate skill IIIc show that most students could determine square roots when the exponent was even but had difficulty when it was odd. The decision was made that more instruction was needed on finding square roots of numbers with odd exponents.

Another specific weakness was revealed by the test items for subordinate skill IIId. A specific kind of multiplication was shown to be a problem. In this case, the decision was made to add instruction on the kinds of multiplication problems in which one factor was squared.

The low percentage attainment on both items in IIIi revealed that this was a major problem area which required attention.

The item data for each of the 15 subordinate skills were analyzed in the manner just described. Following this analysis, 20 additional frames were written and added to the initial 50 frame program. These frames included instruction on 10 of the 15 subordinate skills and were designed to eliminate the specific difficulties identified in the SS Post-Test. The frames were distributed as follows:

a. Thirteen frames on mathematical skills associated with numbers written in exponential notation. The frames were designed to remedy specific weaknesses found on subordinate skills IIa, IIb, IIIa, IIIb, IIIc, IIId, and IIIg.

b. Four frames on translating verbal problems into mathematical statements. Low achievement on two of the items for subordinate skill IIc had shown this to be a problem.

c. Three frames on determining the concentration of ions in solution. The attainment on subordinate skills IIIi and IVa had identified some particular difficulties.

Comparison of treatment groups

Figure 4-2 shows the percentage attainment for the two treatment groups on the seven problems that constituted the criterion task. The mean score on the C Post-Test for students using the initial program was 3.02 (SD = 1.94) and for students using the revised program was 3.84 (SD = 2.08).

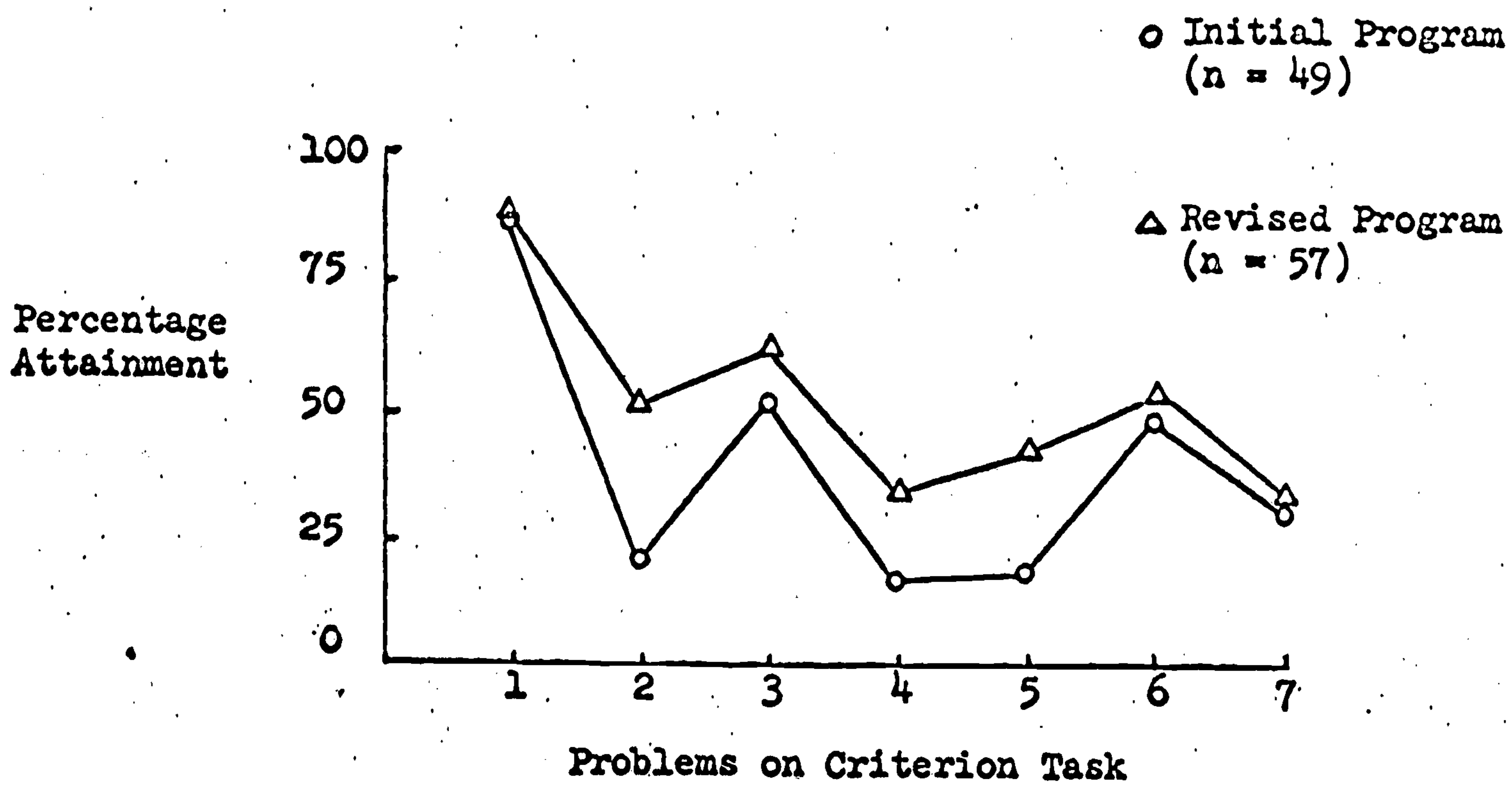

Fig. 4-2. Percentage Attainment on Criterion Task for Students Using the Initial and Revised Programs

A one-way analysis of covariance comparing criterion test gain scores shows the differences in performance to be significant and favor the treatment group using the revised program ($F = 19.45$, $df = 1/100$), $p < .001$). The four pretreatment measures of IQ, mathematics percentile, C Pre-Test, and SS Pre-Test were used as covariates.

The changes in performance on the items used to test attainment of the subordinate skills are shown in Table 4-2. Comparison of the two treatment groups indicates that performance increased on 12 of the 15 subordinate skills and decreased on 3 others. The mean gain score for the initial group on the 32-item SS Post-Test was 7.0 (SD = 3.0) and for the revised group was 10.4 (SD = 3.5). Analysis of covariance of these gain scores on the subordinate skills test showed the differences in performance to be significant and favor the group using the revised program ($F = 15.41$, $df = 1/100$, $p < .001$).

Table 4-2

Percentages of Students Passing All Test Items for a Given Subordinate Skill With the Initial (n = 49) and Revised (n = 57) Programs

Subordinate Skill	Program		
	Initial	Revised	Difference
IIa	39	54	+15
IIb	78	86	+8
IIc	20	38	+18
IIIa	73	95	+22
IIIb	78	86	+8
IIIc	61	77	+16
IIId	65	77	+12
IIIe	69	36	-33
IIIf	86	77	-9
IIIg	94	98	+4
IIIh	39	40	+1
IIIi	12	32	+20
IVa	69	73	+4
IVb	100	98	-2
IVc	80	82	+2

Summary and Conclusions

An initial learning program on solving solubility product problems was studied by a group of 49 chemistry students. Following instruction, performance of these students was obtained on a criterion test and 15 skills which had been identified as subordinate to the final task. The performance of students on these subordinate skills was used to locate specific skills failed by a substantial number of students. Gagné's cumulative learning model served as the basis for identifying the subordinate skills and for predicting that a way to improve performance on a criterion task is to improve performance on the skills subordinate to it.

Twenty frames were added to the original program in accordance with the above analysis making up a revised program of 70 frames. A second group of 57 students then studied this revised program. Both an increase in the number of subordinate skills attained and in criterion test performance were found for this group thereby supporting the major prediction of the study.

The results of this experimental study show that adding instruction leading to improved performance on subordinate skills in a science learning program was markedly successful in improving performance on the criterion task. This result had been predicted from the cumulative learning model. Although there was room for additional improvement, substantial progress was made with one revision of the learning program. Additional cycles of tryout and revision could continue, guided by performance on the tasks in the learning hierarchy.

The practical implications for instructional development suggested by the findings in this study are quite straightforward. In order to bring about attainment of objectives one needs to determine specific difficulties students have in preliminary versions of an instructional program. It is possible to achieve this diagnostic capability by analyzing performance objectives to determine probable subordinate skills. Following instruction, student performance on subordinate skills can be used to pinpoint specific difficulties in the learning program. Finally, empirically based inferences can be made concerning revisions that should lead to improved student performance.

CHAPTER 4 - STUDIES OF RULE LEARNING

(Study 5) Some Factors in Children's Learning and Retention of Rules*

by

Robert M. Gagné and Virginia K. Wiegand
University of California, Berkeley

The learning and retention of ideas have usually been studied with the use of prose materials (e.g., Briggs & Reed, 1943; Cofer, 1941; Hall, 1955; King & Russell, 1966). In such investigations, the learner typically first listens to or reads a prose passage for a specified number of trials, and later is asked to recall or recognize certain selected ideas embodied in the passage. Such an approach generally assumes that what is learned and retained will be influenced by the meaningful context within which the particular ideas to be recalled are imbedded. The context may be presumed to provide cues, as appears to be true with paired associates (McGovern, 1964); it may be a source of mediators (Davidson, 1964); it may provide "organizers" of the sort studied by Ausubel (1962); or it may have still other effects, as yet unidentified.

In contrast to this context-imbedded approach, the present authors were interested in taking an analytic view of the learning of ideas, one which at the outset attempts to separate the effects of context variables from others. This investigation was begun with a specification of the concept of "idea" that was as definite as possible. In pursuit of the latter goal, research of previous writers and theorists provided scant information. Accordingly, in order to proceed with this account, the authors had to propose some of their own definitions. These are tentative, subject to further development, and are given here for the purpose of aiding communication.

---

*This study is reported in the Journal of Educational Psychology (see Gagné and Wiegand, 1968).

For the concept of "idea" itself, the authors refer to Gagne (1965), who considers an idea to be a principle, or rule. The latter is a capability inferred as having been learned, which enables the individual to demonstrate a sequence of concepts such as "If A, then B." The concepts which make up a rule are (simpler) capabilities, evidenced in overt behavior by the identifying of classes of objects or events. Included as concepts are classes of the individual's own actions. A rule may, of course, include a rather large number of concepts in its sequence. However, the simplest rule in a formal sense is one that contains only two concepts. A common variety is a rule combining a "thing concept" with an "action concept," as in "birds fly," "water wets," and "If green, go."

A rule may be concrete or abstract, depending upon whether its component concepts are concrete or abstract. If the two components of a rule may be represented by a concrete noun and a verb describing overt action, the rule is concrete. A partially abstract rule combines a concept represented by a concrete noun with a concept represented by an abstract verb, as exemplified by "women suffer" or "snakes live." The reverse combination, abstract noun—concrete verb, may also be called partially abstract. There are also fully abstract rules, like "power corrupts." The authors believe these distinctions are of some importance to the study of rules. However, no more extensive discussion of them will be undertaken here; they are mentioned in order to convey the meaning of "concrete rule."

The present study had the purpose of investigating some factors with potential influence on the learning and retention of concrete rules, of the sort sometimes called facts. The major variable of interest was the number of such rules to be learned in a single set, when presented one after another. Another variable was the manner of cueing used in eliciting the rule during a test of retention. Both immediate and delayed (3-day) retention were measured.

It is of some interest to note that the amount of retention of learned ideas varies over a wide range, as reported in different studies. It is also true that variations in procedure are many, making comparisons among studies difficult or impossible. An early study by Yoakam (1921), for example, found retention of a prose passage read a single time to be from 48% to 80% as complete after 20 days as retention measured immediately, in different groups of children. In the study of Dietze and Jones (1931), a single reading of a 1,000-word article

yielded scores ranging from 44% to 84% for immediate retention, and from 26% to 42% after 30 days. Newman (1939) reports 87% retention after 8 hours of sleep, 86% after 8 hours of waking, of 12 main ideas contained in a 300-word story. Comparable scores were 47% after sleep and 23% after waking for 12 nonessential ideas. In Cofer's (1941) study of the learning of prose passages, no evidence of retention was found after 9 months, as measured by relearning scores, in Ss who learned the material for logical (as opposed to verbatim) reproduction.

Rationale for an Analytic Study

It is assumed that a single concrete rule, whose component concepts are readily available to recall, is learned in a single trial. If one then undertakes to learn two or more such rules in a row, the possibility exists that they will not exhibit perfect immediate retention. How many such rules can be learned (and retained) when presented one after another? What is the effect of the variable of number of concrete rules on the retention of such rules as indicated both by a test immediately following learning and by a test given 3 days later? These are the questions addressed in this study.

Specifically, fourth-grade children who had previously learned thing concepts, and recalled some familiar action concepts, were given concrete rules to learn. Each rule combined a thing (a drawn shape) with an action (such as underlining the shape). In designing the thing and action concepts making up each rule, the factor of familiarity was considered. Since retention of the rules was to be cued by verbal names, it was desired to keep them equally unfamiliar, and thus to avoid the possible differential effects of mediating verbal associations. Accordingly, nonword names were learned for the things. For action concepts, however, since these were to be recalled, no necessity was seen to insure that they were unfamiliar. Instead, they were chosen to be highly familiar and highly overlearned.

The children learned these rules from the pages of booklets, each page of which gave a printed statement of the rule and then asked the learner to draw it. Different groups of children learned different numbers of rules: three, five, seven, or nine. Retention of the rules was then measured immediately; and in separate groups, after 3 days. For the purpose of investigating cues to recall, half of the retention booklets cued the rules by means of only a verbal name of the thing concept, half by the verbal name plus a drawing of the thing concept.

Method

Subjects

The learners in the study were 96 fourth-grade children in an elementary school located in a primarily middle-class suburban community. With the cooperation of their four teachers, the experiment was conducted as an exercise in science instruction, in which the children were informed participants. The study began with 104 children. The rules of random assignment were followed throughout, with the exception that two students were excluded because they did not complete the prelearning. With absences and other contingencies taken care of, there was a total N of 96 for the study.

Materials

Prelearning of thing concepts. Four different study sheets were prepared to accomplish the prelearning of thing concepts. Each sheet showed a thing (shape) with its printed name (nonsense syllable) underneath it in nine boxes having a scattered arrangement on the page. Ten different test sheets were also made up, containing nine boxes running vertically, each with a printed concept name and a space for the shape to be drawn.

The concept names, typed in capital letters, were nine consonant-vowel-consonant (CVC) syllables with relatively high pronounceability values, averaging 2.49, as given by Underwood and Schulz (1960). No consonants were used in common in either the first or third position, and vowels were balanced in frequency as nearly as possible (BOT, FAC, JUM, LAR, NOP, REL, SUD, TIS, ZIN). The figures were chosen to be distinctive and easy to draw; they included such common figures as a square, a triangle, a heart, and a crescent. In successive study sheets, the figures were shown in four different orders from the top to the bottom of the page. In addition, considerable variation was introduced into nonrelevant features such as size and thickness of outline, without changing the essential shape. A similar plan was followed for test sheets.

Review of action concepts. A booklet was used to present nine different actions. Each action was first shown accompanied by printed instructions which were read by E (i.e., "There is a line under the shape."). The child then executed the action with a second figure of the same kind on the same page, and again with another figure on the next page. The figures employed in

this booklet were different from those used in learning. The nine actions represented were: line under, tail on, circle around, question mark after, dot in, check before, hat over, X through, and legs on. The booklet contained a final test for all nine actions, presented four on one page and five on the next, with printed instructions ("Draw a circle around the shape.") for each action.

Rule learning. The booklets for rule learning were designed to be used by each child in individual learning. Each page first identified a thing concept ("This is a NOP."). Then the statement of the rule was made: A NOP has a circle around it." The rule was then illustrated. Finally, the printed instructions said, "Draw the rule for a NOP," and a blank space was provided where this was to be done. To meet design requirements, a fourth of the booklets presented three rules; a fourth, five; a fourth, seven; and a fourth, nine.

Each thing concept, and each thing-action combination, was made to have an approximately equal frequency of occurrence in booklets for each of the treatment groups pertaining to number of rules (three, five, seven, and nine). Within this constraint, there was random assignment of rule booklets among Ss within each of these groups.

Filler activities. Several booklets were prepared containing filler activities, to be used when each learner finished his learning booklet. These were necessary in view of the fact that the number of rules presented in learning was as few as three for some Ss, and as many as nine for others. The filler booklets contained both geometrical and numerical puzzles.

Retention. Booklets used to measure retention contained 1 page devoted to each rule. The four varieties of booklets were 3, 5, 7, and 9 pages long, to correspond with the number of rules learned by different groups of Ss. Each page contained a printed statement of the form, "Draw the rule for a NOP." To permit the investigation of the cueing variable, half of the booklets included a drawn figure of each thing concept just below its name.

Design

The Ss were first divided into six subgroups representing six levels of IQ, as measured by traditional group test scores. In these Ss, this variable had a median of 110, a range of 89-149, and a roughly normal distribution. The Ss within each IQ

level were then assigned randomly to each of the 16 treatment conditions of the experiment. At the beginning of the study, it was possible to assign seven or eight to each condition. After attritional factors had taken their toll, the assignment of Ss to conditions ended up as shown in Table 1, achieved by the final discarding of data from only three Ss, chosen by a random process.

Table 5-1

Number of Subjects Assigned to Treatment Cells Representing the Experimental Conditions

Retention condition	Cueing condition	Number of rules			
		3	5	7	9
Immediate	Verbal	6	6	6	6
	Verbal + Pictorial	6	6	6	6
Delayed	Verbal	6	6	6	6
	Verbal + Pictorial	6	6	6	6

It will be seen that the design is a four factorial one, making possible the determination of the effects of four different variables: (a) number of rules learned; (b) immediate versus 3-day retention; (c) verbal versus verbal plus pictorial cueing of recall; and (d) IQ level.

Procedure

The schedule for the study was as follows: Days 1-4, prelearning of names of thing concepts; Day 5, review of action concepts; Day 6, learning of rules, followed by retention test for half of the total group; Day 7 (preceded by an interval of 72 hours), retention test for the other half of the group. Each day's session occupied approximately a 1/2 hour period.

The prelearning of the names of thing concepts (shapes) began with a session in which E gave a general introduction to the study. Following this, E drew each shape on the blackboard and placed beside it its printed name. When all nine shapes had been introduced in this way, study sheets were passed out to all the children. They were asked to learn the names of all nine shapes. When a child thought he knew them all, he asked for a test sheet,

and turned in his study sheet. When this in turn was turned in, he received another study sheet, followed by a second test sheet. Two study-test trials of this sort were conducted during each session for 4 days. In addition to the introductory trial, each S thus had a total of eight such trials. On the second test given on the fourth day, only two of the children did not identify all of the names correctly; data from these children were not used in the experiment. The criterion of the prelearning of thing concepts was thus a conjunctive one: completing eight study trials and identifying all nine concepts correctly on the final test.

On the day following the prelearning of thing concepts, the review and test session was conducted for the nine action concepts. Following a brief introduction, these trials were administered by means of the booklet previously described. All the children knew the action concepts after the first trial (and most probably before it).

Whereas prelearning and the action concept review were conducted in the four regular classrooms, the children were assembled in one large room for the rule-learning and retention sessions. Their places at the tables in this room were determined before their assembly, and booklets pertaining to the respective subgroups of the experiment were distributed at these places. In the learning session, the children were first instructed to work through the booklets page by page, and when finished, to go on to the second booklet. They were also asked to record the time of starting each booklet. In this manner, once the trials for rule learning were finished, the learners went on to the other activities contained in the filler booklets, regardless of whether they had studied three, five, seven, or nine rules. The session was brought to an end when it was observed that all children having booklets containing nine rules had finished. Each rule the child encountered was first presented in printed form and illustrated, and then drawn by him on the lower part of the same page. The mean time per rule (each rule being composed of these two steps of observing and recording) was found to be .33 minutes, in the three-rule group; .23 minutes in the five-rule group; .22 minutes in the seven-rule group; and .21 minutes in the nine-rule group.

For half of the total group, distributed equally in relation to the other variables of the experiment, the second booklet was a test of immediate retention, whereas for the other half, the

second booklet contained filler activities. Retention booklets began with no preliminary instructions, simply containing a statement like "Draw the rule for a ZIN" on each page, as previously described. Those who finished these booklets before others went on to a third booklet of filler activities.

Delayed retention was measured with the other half of the total group on a Monday, following a weekend, 3 days after the rule-learning session. In its essentials, the same procedure was followed as for immediate retention. It may be noted that these children had not been told on Friday that they would be asked to recall the rules they had learned.

Results

The raw data of the experiment were number of rules recalled, under the various conditions described in Table 1, and for six IQ levels within each of these conditions. An examination was also made of total overt errors, as well as of various categories of overt errors. Insofar as sufficient data existed, trends in errors appeared similar to those of correct responses; therefore, no more detailed analysis was undertaken.

The data first were converted to percentage remembered. An analysis of variance of these data revealed a significant effect of the immediate versus delayed retention variable ($F = 147$, $df = 1/15$, $p < .01$). Nonsignificant differences were found for the variable of IQ; and also for the variable of verbal versus verbal-plus-pictorial cueing. The analysis did not reveal differences for the variable of number of rules learned; for the data as a whole, the percentage of retention was not significantly different whether three, five, seven, or nine rules had been learned ($F = 2.3$, $df = 3/15$, $p > .05$). In addition, no significant interactions were found.

Cueing effects

The extra cue to recall of the correct rule, namely, the provision of a picture of the thing concept rather than only its name, is not shown by these results to have aided recall of the rule ($F < 1$, $df = 1/15$, $p > .05$). Thus, the group of children who were cued to recall only by the name of the thing concept contained in it, had no more difficulty in recalling the rule than did children who were cued by a picture plus the name. These results may also be considered as evidence that the children were able to use the name effectively, and that their performance in recalling rules was not affected by inability to recall the names per se.

Number of rules learned

The mean percentage recalled is shown (for the verbal and verbal-plus-pictorial groups combined) in Figure 1, for three, five, seven, and nine rules, under conditions of immediate and 3-day retention. It appears that three and five concrete rules are learned virtually intact, when recall is measured immediately. This near-perfect performance appears to drop off, however, when seven rules are learned, and to undergo a further drop for nine rules. The special circumstance to be considered here is that, as compared with these mean values, the performance of the group which learned five rules is 100%, with an SD of 0. It would seem legitimate to pose the question as to whether the means for the seven-rule and the nine-rule group are significantly different from a perfect, performance. The usual statistical tests, however, are not designed to deal with measures having zero variance. Accordingly, the following test was applied.

The standard error of the mean for the group which learned seven rules was 7.6, yielding a confidence interval at the 99% level extending to the score 98.9. The mean of 79.8 for this group is thus significantly different from 100, the score attained by the group which learned five rules. Similarly, for the group which learned nine rules, the upper limit of the confidence interval is 77.9, and by similar reasoning it may be said that the mean of this group (63.9) differs significantly from the perfect performance of the five-rule group. A $\underline{t}$ test of the difference between the immediate retention scores of the seven-rule and nine-rule groups indicated nonsignificance at the .05 level.

As for percentage recalled after 3 days, this tends to be around 20%, and shows only slight and insignificant change depending upon the number of rules learned.

Conclusions

There appear to be some useful facts in the results of this experiment. When both thing concepts and action concepts are previously well learned, fourth-grade children can learn up to five concrete rules by reading and recording each once, one after another, when they are studied at an average rate of one every 15 seconds. When retention of seven and nine such rules is measured immediately, however, losses of the order of 20% and 36%, respectively, are found.

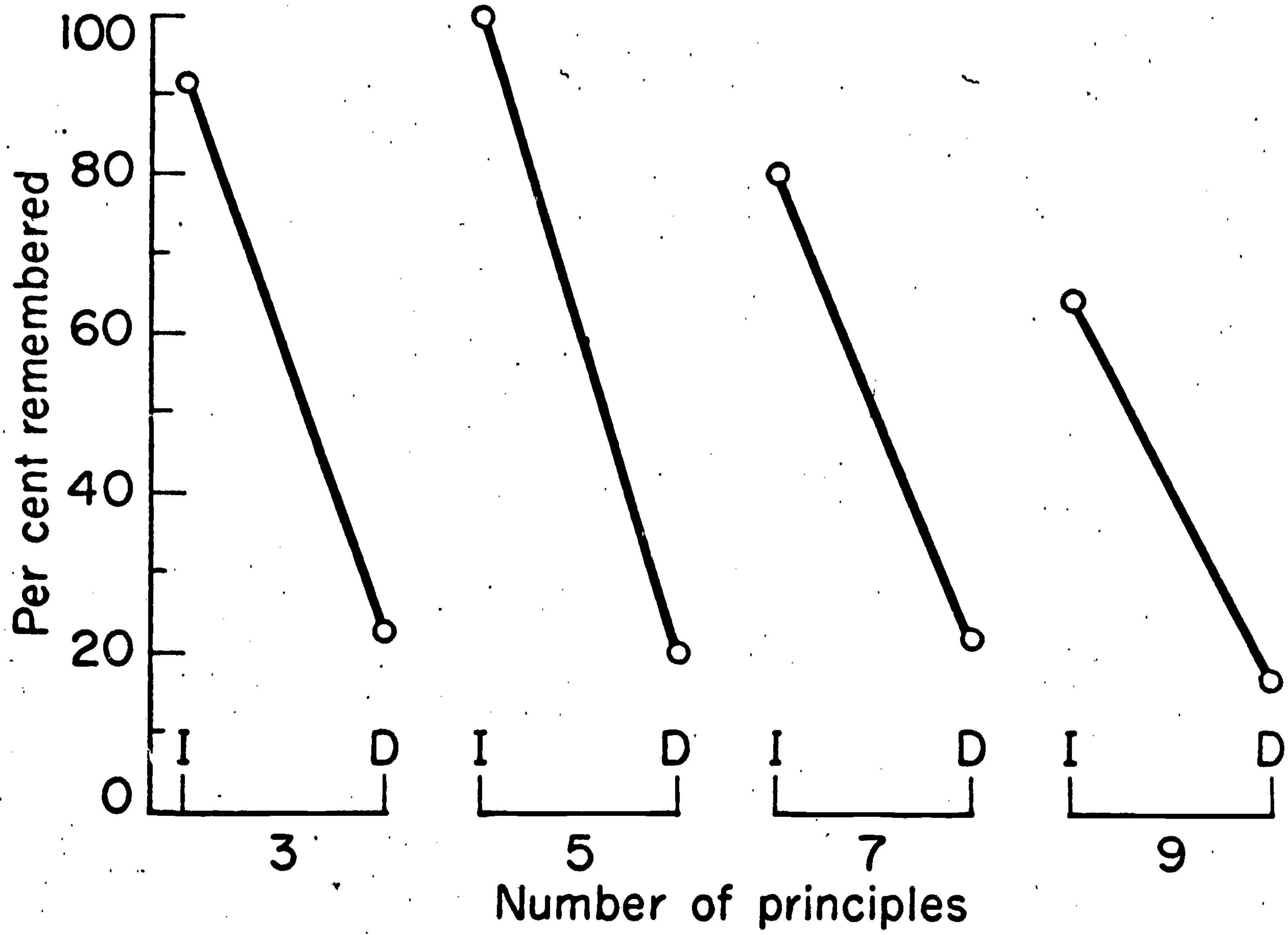

Fig. 5-1. Percentage of principles remembered immediately following learning (I) and after 3 days (D), by groups who were given different numbers of principles (three, five, seven, and nine) to learn.

The dropping off of immediate retention scores under these conditions would be expected to occur as a result of interference (cf. Keppel, 1968), both proactive and retroactive. This experiment, of course, provides no direct confirmation that interference, rather than some other factor, is at work. Such a conclusion would need to be based upon an additional study in which, for example, conditions of learning were arranged in order to contrast the effects of variables favoring interference with variables not favoring it. Differing degrees of similarity among the elements of the rules to be learned might be a variable feasible to investigate with this theoretical purpose in mind. So far as the present results are concerned, interference is merely a likely cause of the observed relation between number of rules presented and percentage retained immediately.

The results also suggest another kind of possible relationship with theory. Postman (1964) points out the relationship of the immediate memory span to the initial events of memory for items to be reproduced serially. Jensen (1964) found a high degree of relationship between measures of immediate memory span and the learning of serial lists of verbal items. Although the rules of the present experiment were not to be learned in serial order, they were presented serially. It may be of some significance that the falling off of immediate retention scores occurs in the interval (five-seven items) which has often been measured as the limit of memory span in children of the age used in this experiment.

As for the amount retained after a 3-day delay, it is most interesting that this was no more than 20%. It should be borne in mind that several obvious factors favor this relatively low degree of retention. First, the children were given no particular incentive to remember, and it therefore seems likely that they undertook little rehersal during the period intervening between learning and recall. Second, the facts they were asked to learn and recall were isolated from each other and from any other meaningful context (except the most general one that they were learning rules containing new names for things). Under these circumstances, the 80% loss in retention appears to be the order of what may be expected. However, it is a high degree of loss when compared to that obtained in studies in which the ideas are imbedded in a context (e.g., Yoakam, 1921; Dietze & Jones, 1931; English, Welborn, & Killiarn, 1934).

The apparent contrast between the remembering of isolated versus context-imbedded ideas would appear to suggest one promising direction for additional study. On the one hand, such research needs to be related to current theoretical formulations of verbatim learning which emphasize the role of contextual stimuli (cf. Melton, 1967). On the other hand, there exists the possibility of further experimental exploration of different kinds of contexts, such as the "subsuming" and "correlative" types emphasized in Ausubel's (1967) theory. It seems likely that continuing to focus attention on the individual "fact" or "rule" will be a useful strategy in such research.

(Study 6) Context, Isolation, and Interference Effects on the Retention of Facts*

by

Robert M. Gagné
University of California, Berkeley

The learning and retention of meaningful materials have been studied in somewhat sporadic fashion over a period of many years (Welborn & English, 1937). One matter of contemporary interest centers on the effects of contextual surroundings on material to be learned and retained. A prominent theory (Ausubel, 1968, pp. 37-48) proposes that meaningful learning occurs when potentially meaningful ideas, in the form of propositions, are encountered by the learner and incorporated into preexisting cognitive structures. Anchoring ideas within such structures may be mobilized by communications to the learner which function as "advance organizers" (Ausubel, 1968, pp. 136-137), and thus act to facilitate both learning and retention (Ausubel & Fitzgerald,1962). Although derived from a different theoretical position, the work of Rothkopf (1968) and his associates also emphasizes the effects of context on meaningful learning and retention, when context, in this case, is provided by text-imbedded questions (Rothkopf & Bisbicos, 1967). These investigators have found marked effects of questions which are relevant to the material being remembered, in addition to "mathemagenic" effects of questions that are not relevant to the material.

If context can facilitate learning and retention of propositional material, an interesting question is raised concerning the matter of interference. When entities to be rotely learned are imbedded in a context of similar materials, it is usually possible to demonstrate that their retention has been markedly affected in an adverse sense by the learning of this contextual material (cf. Keppel, 1968). However, a number of studies have failed to find evidence of interference of materials in a meaningful context (Ausubel & Blake, 1958; Hall, 1955). According to Ausubel (1968, p. 117), interference becomes relatively insignificant in meaningful learning and retention, and is overshadowed by the positive effects of assimilation with relevant anchoring ideas. It is noteworthy, however, that evidence of retroactive interference was found by Entwisle and Huggins (1964) in a study using principles of electrical circuit theory as learning material.

---

*This study appears in the Journal of Educational Psychology (see Gagne, 1969).

In a previous paper, Gagné and Wiegand (1968) undertook to investigate the retention of individual concrete rules (facts) in isolated form. The "things" mentioned in the rules (nonsense names for initially unfamiliar shapes) had been learned to a mastery criterion. The actions were familiar activities performed with a pencil (underlining, circling, etc.). Each rule was a stated relation between a thing and an action, presented and practiced once. It was found that fourth-grade children could learn perfectly as many as five such rules, presented at a rate of one every 15 seconds. Retention after 3 days was about 20%. This previous study provides two items of interest to the present work. First, it has shown the feasibility and reasonableness of investigating retention of meaningful material by using single-trial presentation as opposed to repeated presentations. Second, it raises the question of how retention might be increased from the relatively low value of 20% obtained with unfamiliar "isolated" facts. Meaningfulness, in the sense of contextual relatedness, may turn out to be a variable with powerful effects, when considered and investigated in this way.

The purpose of the present study was to examine the effects of context on the retention of propositional materials, and to contrast these with the effects of "isolation." It was desired to study context effects, under conditions which might reveal the influence of "organizers," on the one hand, and the influence of interference on the other. Although the study planned no direct comparison of the relative power of these two processes, it was expected that some idea could be gained concerning their degree and manner of operation in retention.

The study presented fourth- and fifth-grade pupils with facts about howling monkeys, in the general setting of a science lesson. Five facts, later to be tested for retention, were presented in isolation, contrasted with their presentation in a context of other facts. Because of theoretical interest in the way context operates, three different contexts were employed: one containing a superordinate idea as a topic sentence, one containing coordinate ideas, and one containing unrelated ideas. To test the effects of interference, a condition of massed presentation was contrasted with a condition of spaced presentation (cf. Keppel, 1964). As a further question related to the latter contrast, a condition in which rehearsal was probable was compared with one in which it was improbable.

Method

Subjects

The learners in the experiment were 56 pupils in the fourth grade and 56 in the fifth grade of an elementary school located in a primarily middle-class suburban community. The Ss in each grade were first divided into seven subgroups representing seven levels of IQ, as measured by traditional group test scores. The range of these scores was 85-139, with a median of 119. Within each IQ level, Ss were then assigned randomly to each of the eight treatment conditions of the experiment. Repeating the procedure for each grade meant that seven fourth graders and seven fifth graders were assigned to each individual cell. Absences in the initial group of 118 pupils were handled by discarding cases according to a random procedure, in order to achieve equal numbers for each treatment cell.

Materials

All the children were initially given some introductory material on howling monkeys, presented by means of a brief lecture accompanied by projected slides. This material related information about the habitat and appearance of howling monkeys.

The five facts later tested for retention were presented on slides which were projected for 5 seconds ($\pm$ 1 second) each, and read to the children while on the screen. The facts were as follows: (a) young howlers wrestle and chase each other in play; (b) mother howler monkeys carry their babies with them wherever they go; (c) howlers roar and throw objects at animals they are afraid of; (d) a howler uses his tail as another hand; (e) howler monkeys prefer to live in forests of tall trees.

Sentences representing context ideas were prepared in three different categories for slide projection and reading to the children. For the superordinate context, there were four sentences (in addition to the fact to be remembered) related to each of the topics of (a) howler play; (b) mother-child relationships; (c) actions toward enemies; (d) the howler's tail; and (e) forest living. The initial sentence of the five was a "topic sentence" introducing the general idea; for example, the topic sentence related to the first topic was, "Young howlers have many forms of play." For the coordinate context, there were also four additional related sentences for each topic, but the initial topic

sentence was replaced by another sentence conveying a particular related fact. The unrelated context was provided for each topic by four sentences which did not pertain to howling monkeys. These sentences were chosen to be not too startlingly contrastive with the general subject, however; accordingly, they dealt with the appearance and habits of other animals, trees, terrain features, or climate.

Retention of the facts was measured by two different tests. A recall test was constructed by placing on each page of a booklet a paraphrase of the proposition to be recalled containing a blank for a missing key word (e.g., Howler babies are always _______ by their mothers), along with three other blank-containing propositions which were not contained in the learning material (e.g., Howler babies are always _______ by their sisters). Instructions for this test directed the child first to identify by a check mark the idea he remembered learning, and then to complete it by filling in the blank. This test, therefore, provided two different scores: (a) recognition of the idea, and (b) recall of the substance of the idea. Since the propositions were paraphrased versions of the sentences presented for learning, verbatim retention was presumably not involved.

A second kind of recognition measure was devised for the purpose of testing retention of the context material, as opposed to the five facts which were the major focus of interest. This test contained paraphrased propositions representing each of the context statements (4 for each topic, or 20 in all). These were to be marked "true" or "false."

There were four different context subgroups under both massed and spaced conditions. These were (a) superordinate, in which four facts related to the fact to be remembered were presented, of which the first was a topic sentence; (b) coordinate, in which a context of four related facts was presented without a topic sentence; (c) unrelated, a context containing four other unrelated facts; and (d) isolated, in which each fact to be remembered was presented without any context propositions.

In the spaced condition, a fact to be remembered was separated from the next such fact by the presentation of four other facts (or, in the no-context condition, by an equivalent amount of time). In the massed condition, all five facts to be remembered were presented one after another at the end of the learning session.

There were, then, four different context groups in the massed-presentation condition, and four in the spaced-presentation condition. Half of each subgroup were pupils from Grade 4, and half from Grade 5. Seven levels of IQ were represented in each subgroup for each grade. Sex was not controlled, although it was observed to be roughly evenly divided throughout.

Procedure

The children were assembled in groups of 14-16, fourth and fifth graders together, in a small room separate from their regular classrooms. The learning session was conducted on Tuesday. The children were seated, given the introductory talk with slides about howling monkeys, and then presented the lesson consisting of a series of verbal propositions, projected on the screen and read to them. They were not informed that retention of the facts would be measured.

After each fact was read, the children were directed to complete a blank in a booklet before them, writing in a word that was missing from the sentence that had just been shown. The sentence was read to them, in approximately 3 seconds, including the word "blank" where appropriate. Then they were allowed 6 seconds to complete the blank, before the next sentence was projected on the screen. This procedure was employed in order to insure attention to (and, presumably, learning of) each fact. The 25 facts were presented one after another in this manner for all the context conditions. For the no-context condition, only the five facts to be remembered were shown.

In the spaced condition, the fact to be remembered came between two prior context facts and two subsequent context facts, all belonging to the same topic. Thus the spacing between facts was four context facts, requiring a total of 60 seconds (slide on, 5 seconds; reading sentence with blank, 3 seconds; completing blank, 6 seconds; changing slide, 1 second). When no context was employed, under the spaced condition, an equivalent amount of time (60 seconds) was allowed to elapse before the next fact to be remembered was presented. During this time, a picture of a howling monkey was shown on the screen, and the children were told to wait a bit, and to "think about the fact they had just seen."

In the massed condition, all the context facts (20) were presented first, with the same order of topics as that used for the spaced condition. The five facts to be remembered were presented one after another toward the close of the learning session, following the same procedure as had been used for the remaining facts, including filling in a blank for each. Thus the spacing between facts in this condition was 15 seconds.

Two days later, the same groups were assembled in the same room, and completed the retention tests. Since the measurement of retention was completed in a few hours and during school hours, little opportunity was afforded for discussion of the retention measurement with classmates while it was going on. However, each group was informed about the purpose of the study after testing was completed.

Results

Scores were obtained for three measures of experimental outcome. First, there were scores representing number (out of five) of facts recognized, under the various treatment conditions. Second, and of primary interest, scores were obtained measuring number of facts recalled. Third, scores were available to measure the recognition of contextual facts (of which there were 20) under the various context conditions.

Recall measures

Significant variables. Results of an analysis of variance performed on recall scores are shown in Table 6-1. These results clearly indicate significant effects for the context variable, while none are shown for the massed-spaced presentation contrast. Grade also turns out to be a significant variable, with fifth graders showing greater recall scores, on the average, than fourth graders.

Scheffé contrasts were performed on a number of variable pairs, in order to isolate further the effects of context. Using the .05 level of confidence, the results indicate that each pair of differences is significant. The conditions arrange themselves in the following order, insofar as their facilitation of recall is concerned: no context $>$ superordinate context $>$ coordinate context $>$ unrelated context. Several interesting conclusions are apparent in these results.

Table 6-1

Analysis of Variance on Recall Scores for Five Facts

Source	df	MS	F
Context (C)	3	.3548	10.30*
Spaced/massed presentation (P)	1	.0129	---
Grade (G)	1	.4629	13.44*
IQ (I)	6	.0424	1.23
C X P	3	.0633	1.84
C X G	3	.0181	---
C X I	18	.0148	---
P X G	1	.0057	---
P X I	6	.0345	1.00
G X I	6	.0245	---
C X P X G	3	.0467	1.35
C X P X I	18	.0494	1.44
C X G X I	18	.0442	1.28
P X G X I	6	.0224	---
Residual	18	.0344	

* $p < .01$.

First, the marked difference in presentation represented by the massed and spaced conditions has not shown an effect on recall. If indeed a difference in interference is generated by these two presentation conditions, it has not occasioned differences in amount of recall of meaningful propositions.

Second, the no-context condition is clearly superior to all those in which context was employed. Presumably, this result could occur because interference is reduced under the condition in which facts to be remembered are "isolated."

Third, the use of a context containing a topic sentence, which might be expected to function as an organizer, brings about a significantly greater amount of recall than context conditions which are coordinately related or unrelated. The theoretical position which expects an advantage from organizers may also be said to be supported by the finding that a coordinate context facilitates recall more than does an unrelated context.

Table 6-2

Mean Recall Scores and Percentage of Recall for Pupils in Grades 4 and 5, under Four Context Conditions (Massed and Spaced Presentation Combined)

Condition	Grade	M	Combined M	%
Superordinate	4	2.79	3.09	62
	5	3.38		
Coordinate	4	2.71	2.86	57
	5	3.00		
Unrelated	4	2.14	2.54	51
	5	2.93		
No context	4	3.43	3.86	77
	5	4.29		

Amount of recall. The mean recall scores of the five facts, for pupils in Grades 4 and 5, and for both grades combined, are shown in Table 6-2. Means of the massed and spaced conditions have been combined, in view of the finding of a nonsignificant difference between them. Percentage of recall scores are also shown.

It is of interest to note that the facts presented under isolated (no-context) conditions attained a value of 3.86, or 77%, in recall. (Even for the massed condition the figure is 3.57, or 71%.) This contrasts markedly with the findings of a previous study (Gagné & Weigand, 1968), in which five isolated facts were recalled to only a 20% level after 3 days. The difference of 1 day no doubt accounts for some of this discrepancy, but probably not for all. Opportunity for rehearsal under the massed presentation conditions was, if anything, smaller than that which obtained in the previous study. A difference in degree of familiarity may be the variable of importance.

Recognition of Facts

An analysis of variance was performed on the data representing recognition of the facts to be remembered, obtained from scores made by the children in checking the paraphrased sentences. Nonsignificant differences were found for the massed-spaced presentation variable ($F = 3.10$, $df = 1$, $p > .05$), and also for the context variable ($F = 2.47$, $df = 3$, $p > .05$). In contrast to recall scores, the latter variable does not appear to affect scores for recognition. Significant effects on recognition scores were obtained for grade ($F = 8.19$, $df = 1$, $p < .05$), and for IQ ($F = 9.23$, $df = 6$, $p < .05$). On the whole, these differences exhibit quite a different pattern of influence of the variables than that shown for recall.

As for the amount of retention measured by these recognition scores, it was, not surprisingly, higher than that for recall. The combined means for the context conditions were as follows: no context, 4.29; superordinate context, 3.97; coordinate context, 3.67; unrelated context, 3.86. Thus the range of percentage of recognition scores is 73-86%.

Recognition of Context Facts

For all except the no-context groups, it was possible to obtain measures (from a true-false test) of recognition of the facts included in the context material. When obtained in this way, measures of recognition were relatively high, and comparable in magnitude to recognition scores for the facts to be remembered. The combined means were: superordinate context, 15.86 (79%); coordinate context, 16.57 (83%); unrelated context, 16.47 (82%).

No significant differences were found, in these context recognition scores, associated with any of the major variables. The results were: context ($F = 1.12$, $df = 2$, $p > .05$); massed/spaced presentation ($F = 1.16$, $df = 1$, $p > .05$); grade ($F = 1$, $df = 1$, $p > .05$); IQ ($F = 1.01$, $df = 6$, $p > .05$). At the .05 level, a significant interaction appeared with the variables Massed/Space X Grade. Examination of the means indicates that the spaced condition yielded higher scores for the fifth graders, the massed condition for the fourth graders. While this finding is suggestive, it does not seem possible to provide an interpretation at the moment.

Conclusions

There appear to be two major points which can be made about the findings of the study. First is the conclusion that presentation of facts to be remembered in a way which isolates them from other facts (as in the no-context condition) improves their retention over a 48-hour period. This is true even when the facts isolated in this way are presented for learning one after another, with brief (1-second) pauses between them, and little opportunity for rehearsal. When five facts are presented along with 20 other facts, they are not remembered as well as when they are presented alone.

It seems reasonable to conclude that what is being exhibited in this contrast is the effect of interference. Results for recognition of the context facts imply that these were initially learned as well as the to-be-remembered facts, since the percentages of recognition scores are quite comparable. The remembering of these "extra" facts has the effect of reducing the retention of the facts to be remembered, as indicated by recall scores. Apparently, interference does happen, and its effect is a significant one.

Why is the effect of interference not shown in a contrast between recall following massed presentation versus spaced presentation of the facts to be remembered? Perhaps no definite answer can be given. It is in itself an interesting finding that spaced presentation of facts has no particular advantage over a presentation of considerably greater massing (one fact/60 seconds versus one fact/15 seconds). These results, however, are consistent with the notion that interference exerts its main effects during the storage-retrieval phases of remembering, rather than during the phase of initial learning. Facts to be remembered may be difficult to "sort out" from other context facts when recall is asked for, even though they may not be more difficult to learn when surrounded by other facts.

The facilitating effects of isolating facts cannot be attributed to the factor of rehearsal, according to these results. The failure to find a difference between spaced presentation (in which rehearsal was encouraged) and massed presentation (in which opportunity for rehearsal was slight) indicates the weakness of this variable as a causal agent in recall. It is possible, of course, that effects of rehearsal might be shown with even more extreme differences between massed and spaced presentations. It is also possible that this

variable, which has often been hypothesized to affect recall of verbatim material, simply does not influence the remembering of meaningful material of the scope studied here.

The second major conclusion relates to the effects of different kinds of context on the recall of facts. A context which is superordinate, introducing context facts with a general topic sentence, facilitates recall of facts to a greater extent than does a coordinate or unrelated context. The coordinate context, containing facts related to the fact to be remembered, also leads to greater recall than does an unrelated context. Both these findings are consistent with the theory of Ausubel (1968) to the effect that retention of meaningful facts is improved when efforts are made to mobilize anchoring ideas within a pre-existing cognitive structure. Somehow, even in the face of interference, the organization of facts into topics operates to facilitate retention of particular facts.

A finding of interest, although possibly of less central importance, is that the amount of recall of five facts ranges from 51% under least favorable context conditions to 77% under most favorable (no context) conditions, after 2 days. A previous study dealing with fully learned facts of minimal familiarity found 20% retention after 3 days (Gagné & Wiegand, 1968). This finding suggests the possible importance of the factor of familiarity in retention of facts. In the previous study, the "thing" component of the fact to be learned (a nonsense name) was learned to a 100% criterion; the "action" component (drawing a figure on paper) was presumably highly overlearned. The relation between the two (the rule to be remembered) was learned in one trial. In contrast, the facts used in this study contained highly familiar thing concepts (mother, child, tail, tree, etc.). It appears, therefore, that the familiarity factor was, on the whole, quite different in the two studies. The differential effects of familiarity of "things," "actions," and "relations" on recall of meaningful propositions may well be a worthwhile subject for experimental study.

A word should probably be added regarding the definite finding of superiority in recall of fifth graders versus fourth graders. The greater maturity in intellectual development of the former group is, of course, a possible hypothesis. So also is the idea that the fifth graders have had more time to acquire an efficient set of strategies for remembering facts.

(Study 7) Effects of a Superordinate Context on Learning and Retention of Facts*

by

Robert M. Gagné
Florida State University

and

Virginia K. Wiegand
California State College at Hayward

Research on the learning and retention of meaningful material has been pursued for many years. A number of studies have suggested that the context in which ideas are presented during learning has a significant facilitative effect on retention of these ideas. In the work of Ausubel (1968), the context provided by prior "organizers" (Ausubel, 1960) or by previously learned correlative ideas (Ausubel and Fitzgerald, 1961: Ausubel and Youssef, 1963) has been shown to affect retention. Such findings have been interpreted as supporting a theoretical position which makes meaningful learning and retention dependent on the pre-existence of more general cognitive structures under which the ideas represented in newly presented propositions can be subsumed. In this theory, the provision of context material is conceived as having the effect of mobilizing anchoring ideas of the cognitive structure, thus facilitating the subsumption process; an effect which is reflected in heightened learning and retention.

Another source of evidence regarding effects of context comes from the work of Rothkopf (1966) on the insertion of test-like events in passages of meaningful material. Rothkopf and Bisbicos (1967) found marked increases in retention in groups of learners who read passages of meaningful material containing questions. Part of this increase was attributable to the direct instructive effect of the questions, while a smaller but significant part was shown in this study to be attributable to more general "inspection behaviors".

---

*This study is reported in the Journal of Educational Psychology (see Gagné and Wiegand, 1970).

A variety of context effects on the retention of facts was investigated in a previous study (Gagné, 1969). Results indicated that the retention of facts presented along with a "superordinate" context (introduced by means of a topic sentence) was superior in retention to a context containing "coordinate" facts, which was in turn superior to one containing "unrelated" facts. In that study, evidence of the operation of the factor of interference in retention was also obtained, under circumstances which suggested its locus of operation as the storage and retrieval stages of remembering, as opposed to the learning stage.

A similar question about locus naturally arises in connection with the "organizing" effects of context, particularly those associated with the use of a superordinate context. In the study cited, the effects of the "topic sentence" presentation appeared even when the topic sentences were separated from the facts-to-be-remembered by time and other facts; i.e., even when the latter facts were massed together at the end of the presentation, rather than following the topic sentences by short intervals. Is the influence of the topic sentence exerted during the learning phase, or does it have its organizing effect during the retrieval phase, or both?

To gain additional information on this question, an experiment was designed to test the effects on fact retention of the presentation of a superordinate context during learning, and just preceding recall. The superordinate context, as in the previous study (Gagné, 1969) was provided by a "topic sentence", and many other portions of the procedure were similar. Fourth grade children participated in a science lesson on howling monkeys, the essential components being 25 sentences, divided into five topics. The learning conditions contrasted the presentation of a topic sentence introducing each topic, with the presentation of the same material containing a related (but non-superordinate) sentence in place of the topic sentence. The recall conditions contr sted the completion of paraphrased sentences when a "topic sentence" was read to the children just prior to this action, with a similar condition in which the topic sentence was not read.

Method

Subjects

The learners for the experiment were 62 pupils in the fourth grade of a suburban school in a predominantly middle class community. After allowing for attrition caused by absences, the total N for the study was 44, assigned randomly to four subgroups of equal number. The division of these groups was approximately equal by sex, although the effect of this variable was not studied. The Ss were assigned randomly to the four conditions of the experiment within 11 levels of IQ. The latter measure was obtained from school records based upon traditional group tests, and ranged from 85-145 with a median of 113.

Materials

Materials were the same as those used in the previous study (Gagné, 1969). A brief introductory lecture told the children about howling monkeys, their appearance and habitat. Sentences embodying the context ideas and the to-be-remembered ideas were presented by means of projected slides, one sentence to each. Each child responded in an individual booklet, containing each of the sentences he saw (but with a blank included), typed one per page.

The five facts-to-be-remembered, presented in positions 3, 8, 13, 18, and 23 in sequence with other facts, were (1) young howler monkeys wrestle and chase each other in play; (2) mother howler monkeys carry their babies with them wherever they go; (3) howler monkeys roar and throw objects at animals they are afraid of; (4) a howler monkey uses his tail as another hand; (5) howler monkeys prefer to live in forests of tall trees.

In the superordinate learning condition, a topic sentence was used as the first of five slides in a topic. These sentences, occurring in positions 1, 6, 11, 16, and 21 in the sequence, were as follows: (1) young howlers have many forms of play; (2) howler mothers take very good care of their children; (3) howlers are usually peaceful, but they do try to frighten their enemies; (4) his tail is very important to a howler monkey; (5) groups of howlers live in treetops deep in the forest. In this condition, topic sentences were also accompanied by the brief remark "This sentence tells what the next few are all about". In the coordinate learning condition, a sentence containing a related specific fact was substituted for the topic sentence, and the extra instruction was omitted.

Retention of the five facts-to-be-remembered was measured by means of a five-page booklet constructed as follows. Each page contained four sentences each containing a blank. One of these was a paraphrase representing one of the facts that had been presented for learning, the other three "misleads". Children were first to check the fact they had learned (providing a measure of recognition), and then to complete the blank (for a measure of recall). In addition, recognition of the twenty other context facts was measured by means of a true-false test (administered last) containing items which were paraphrases of the context facts as originally presented.

Design

Within each IQ level, obtained by ranking the entire class, Ss were assigned to four different groups. There were two learning conditions, each in turn subdivided into two different recall conditions. The learning conditions were labeled Superordinate (S), indicating a context using a topic sentence, and Coordinate (C), in which the context had no topic sentence. The recall conditions were with a topic sentence present (P) or absent (A). Thus the four groups were SP, SA, CP, and CA.

Analysis of variance was used to examine the effects of the two major experimental variables, Learning Presentation (S or C), and Recall Presentation (P or A), as well as the effect of IQ.

When absences were noted at the time the children were to be assembled for recall sessions, random substitution at the same IQ level was done where possible between recall conditions; otherwise, cases were randomly dropped from the analysis in the interest of achieving equal numbers in each treatment cell.

Procedure

The children were first assembled in two different groups, half to receive the Superordinate presentation, half the Coordinate, in a room separate from their classrooms. They were given the introductory talk, accompanied by slides, and proceeded through the 25 facts. Each fact was projected on the screen and read to the children (5 secs., ± 1). The children were then directed to fill in a blank in the same sentence appearing in a booklet in front of them, one sentence per page.

This sentence was also read to them, using the word "blank" where appropriate (requiring 3 secs.). Six additional seconds were allowed for the children to make their responses, followed after 1 sec. (for automatic slide-changing) by the presentation of the next projected sentence. The total interval from one slide to the next was thus 15 sec.

Twenty-five sentences were shown and reacted to in this way, one after another. In the Superordinate condition, every fifth slide was a topic sentence; in the Coordinate condition, a non-superordinate sentence related to the other five facts was substituted.

Retention was measured two days following the learning period. The children were reassembled in the same room, in two different groups, each of which included half the subjects of the two different learning conditions. The first of these groups took the test for retention of the five facts-to-be-remembered by being reminded orally of the superordinate sentence relevant to each page of the test. That is to say, the experimenter said, "Young howlers have many forms of play. Now go ahead with this page". In the second group, the topic sentence was not given, and the experimenter simply said, "Now go ahead with this page." Each page was to be completed by (1) checking the fact (paraphrased sentence with blank) that had been seen during the learning session, and then (2) completing the blank. For both groups, a true-false test of recognition of the other twenty context facts was administered last.

Results and Conclusions

Table 7-1 shows the percent remembered for the four groups of the experiment, measured in the following ways: (1) as recognition of the paraphrased sentences representing the five facts to be remembered; (2) as recall of the facts, indicated by correct completion of the paraphrased sentences; and (3) as a recognition score on a true-false test representing the other 20 context facts.

Table 7-1

Mean Percent Recall and Recognition Scores for the Four Groups of the Experiment (N = 11 in Each Group)

Score	Groups			
	Superordinate Sentence		Coordinate Sentence	
	SS in Test (SP)	No SS in Test (SA)	SS in Test (CP)	No SS in Test (CA)
Recognition of Fact	80	76	82	62
Recall of Fact	65	53	65	51
Recognition of Context Facts	83	78	75	82

The first row of the table, recognition of facts, indicates higher means for the two groups which had a superordinate sentence presented immediately before the retention test. This effect is a significant one ($F=5.19$;df 1,10;$p<.05$). The effect of the Learning Treatment variable was not significant ($F=1.68$;df 1,10), nor was the IQ variable ($F=0.97$;df 10,10). Thus, for recognition of facts, presenting a superordinate sentence just prior to retention brings about a significant increase in percent remembered. According to the means, this effect is particularly prominent in the group which has not previously experienced the superordinate sentence during learning.

The same trend is indicated by the second row of the table, pertaining to recall of the facts-to-be-remembered. While the means for learning treatments are insignificantly different ($F=2.45$;df 1,10), the recall treatment yields different scores in each case, higher when the superordinate sentence was used ($F=6.98$;df 1,10;$p<.05$). Again, no significant effect was found for the IQ variable ($F=0.74$;df 10,10). For recall, then, the results show that the introduction of a superordinate sentence at the time of retention testing has the effect of increasing the amount remembered. According to the means, this effect occurs to an equal extent whether or not the learners have experienced the same superordinate sentence during learning.

On the whole, the results obtained in this study are generally consistent with those of a previous investigation (Gagné, 1969). The percent remembered obtained with the three measures used here is entirely comparable to that found with another group of fourth and fifth grade children in a different school. There is, however, a discrepancy in results, in that the previous study found a significant effect of the superordinate sentence when presented as a part of the context materials during learning. Examination of the previous data indicates that the superordinate-coordinate difference is substantial in the fifth-graders of that study, but very small in the fourth-graders. A factor of language development is accordingly suggested as a possible cause of the difference in findings.

The present study extends the previous one in a single direction: finding a locus of effect for the use of a topic sentence as an "organizer" in remembering. The results obtained indicate that the topic sentence improves the remembering of facts when presented just before the retention test. Such a finding emphasizes the importance of the retrieval phase of the remembering process, and suggests the desirability of investigation of various kinds of "organizing" techniques designed to facilitate retrieval.

CHAPTER 5 - REFERENCES

Ausubel, D. P. A cognitive-structure theory of school learning. In L. Siegel (Ed.), Instruction: Some Contemporary Viewpoints. San Francisco: Chandler, 1967.

Ausubel, D. P. Educational Psychology. A Cognitive View. New York: Holt, Rinehart & Winston, 1968.

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology, 1962, 53, 243-249.

Ausubel, D. P., & Blake, E. Proactive inhibition in the forgetting of meaningful school material. Journal of Educational Research, 1958, 52, 145-149.

Ausubel, D. P. and Fitzgerald, D. The role of discriminability in meaningful verbal learning and retention. Journal of Educational Psychology, 1961, 52, 266-274.

Ausubel, D. P., & Fitzgerald, D. Organizer, general background, and antecendent learning variables in sequential verbal learning. Journal of Educational Psychology, 1962, 53, 243-249.

Ausubel, D. P. and Youssef, M. The role of discriminability in meaningful parallel learning. Journal of Educational Psychology, 1963, 54, 331-336.

AAAS Commission on Science Education. Science--A Process Approach, Third Experimental Edition, Part Five, Washington, D. C.: American Association for the Advancement of Science, 1966.

AAAS Commission on Science Education. Science--A Process Approach: Purposes, Accomplishments, Expectations. Washington, D. C.: American Association for the Advancement of Science, 1967.

Bloom, Benjamin S. Learning for mastery. Evaluation Comment. Los Angeles: University of California, Los Angeles, Center for the Study of Evaluation, Vol. 1, No. 2, May 1968.

Briggs, L. J., & Reed, H. B. The curve of retention for substance material. Journal of Experimental Psychology, 1943, 32, 513-517.

California Statewide Social Sciences Study Committee. Abridgement of the Proposed K-12 Social Sciences Framework. Sacramento: California State Department of Education, 1968.

Callantine, M. F., and Warren, J. M. Learning sets in human concept formation. Psychological Reports, 1955, 1, 363-367.

Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. American Journal of Psychology, 1941, 54, 1-20.

Coleman, Lois T. A study of the acquisition of the intellectual skill comparing in the area of social sciences. Ed. D. Dissertation, University of California, 1969.

Cox, R. C., and Graham, G. T. The development of a sequentially scaled achievement test. Paper delivered at Annual Meeting, American Educational Research Association, Chicago, 1966.

Davidson, R. E. Mediation and ability in paired-associate learning. Journal of Educational Psychology, 1964, 56, 352-356.

Dietze, A. G., and Jones, G. E. Faculty memory of secondary school pupils for a short article which they read a single time. Journal of Educational Psychology, 1931, 22, 586-598.

Dukes, W. F., and Bevan, W. Stimulus variation and repetition in the acquisition of naming responses. Journal of Experimental Psychology, 1967, 74, 178-181.

English, H. B., Welborn, E. L., and Killian, C. D. Studies in substance memorization. Journal of General Psychology, 1934, 11, 233-259.

Entwisle, D. R., & Huggins, W. H. Interference in meaningful learning. Journal of Educational Psychology, 1964, 55, 75-78.

Fenton, E. The New Social Studies. New York: Holt, Rinehart, and Winston, 1967.

Fergus, G. A. Human abilities. Annual Review of Psychology, 1965, 16, 39-62.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M. The Conditions of Learning. New York: Holt, Rinehart & Winston, 1965.

Gagné, R. M. The Conditions of Learning. 2nd Ed. New York: Holt, Rinehart & Winston, 1970.

Gagné, R. M. Context, isolation, and interference effects on the retention of fact. Journal of Educational Psychology, 1969, 60, 408-414.

Gagné, R. M. "Contributions of learning to human development." Psychological Review, 1968, 75, 177-191.

Gagné, R. M. "Curriculum research and the promotion of learning." In Perspectives of Curriculum Evaluation. Chicago: Rand-McNally, 1967.

Gagné, R. M. Learning hierarchies. Educational Psychologist, 1969, 6, 1-9.

Gagné, R. M., and Staff, University of Maryland Mathematics Project. Some factors in learning non-metric geometry. Monographs of the Society for Research in Child Development, 1965, 30, No. 1, 42-49.

Gagné, R. M., Mayor, J. R., Garstens, H. L., and Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs, 1962, 76 (Whole No. 526).

Gagné, R. M. and Bassler, O. C. Study of retention of some topics of elementary nonmetric geometry. Journal of Educational Psychology, 1963, 54, 123-131.

Gagné, R. M. and Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75 (Whole No. 518).

Gagné, R. M., and Wiegand, V. K. Effects of a superordinate context on learning and retention of facts. Journal of Educational Psychology, 1970, 71 (in press).

Gagné, R. M., and Wiegand, V. K. Some factors in children's learning and retention of concrete rules. Journal of Educational Psychology, 1968, 59, 355-361.

Gibson, Jeanne R. Transfer effects of practice variety in principle learning. Ph. D. dissertation, University of California, Berkeley, 1969.

Guttman, L. A basis for scaling qualitative data. American Sociological Review, 1944, 9, 139-150.

Hall, J. F. Retroactive inhibition in meaningful material. Journal of Educational Psychology, 1955, 46, 57-62.

Hively, W., II. Defining Criterion Behavior for Programmed Instruction in Elementary Mathematics. Cambridge, Mass.: Committee on Programmed Instruction, Harvard University, 1963.

Hull, C. L. Quantitative aspects of the evolution of concepts. Psychological Review Monographs, 1920, 28, No. 1.

Inhelder, Barbel, and Piaget, J. The Growth of Logical Thinking. New York: Basic Books, 1958.

Jensen, A. R. Individual Differences in Learning: Interference Factor. Berkeley, California: Institute of Human Learning, University of California, 1964. (Cooperative Research Project No. 1867).

Keppel, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 91-111.

Keppel, G. Verbal learning and memory. Annual Review of Psychology, 1968, 19.

King, D. J., and Russell, G. W. A comparison of rote and meaningful learning of connected meaningful material. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 478-483.

McGovern, J. B. Extinction of associations in four transfer paradigms. Psychological Monographs, 1964, 78 (Whole No. 593).

Melton, A. W. Individual differences and theoretical process variables: General comments on the conference. In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus: Charles E. Merrill, 1967.

Morrisett, L., and Hovland, Carl I. A comparison of three varieties of training in human problem solving. Journal of Experimental Psychology, 1959, 58, 52-55.

Newman, E. B. Forgetting of meaningful material during sleep and waking. American Journal of Psychology, 1939, 52, 65-71.

Newton, J. M., and Hickey, A. E. Sequence effects in programed learning of a verbal concept. Journal of Educational Psychology, 1965, 56, 140-147.

Okey, James R. A study of formative curriculum evaluation through the use of hierarchical analysis. Ph. D. dissertation. Tallahassee: Florida State University, 1968.

Postman, L. Short-term memory and incidental learning. In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Rothkopf, E. Z. Two scientific approaches to the management of instruction. In R. M. Gagné & W. J. Gephart (Eds.), Learning Research and School Subjects. Itasca, Ill.: Peacock, 1968.

Rothkopf, E. Z., and Bisbicos, E. Selective facilitative effects of interspersed questions on learning from written material. Journal of Educational Psychology, 1967, 58, 56-61.

Schutz, R. E., Baker, R. L., and Gerlach, V. S. Measurement Procedures in Programmed Instruction. Tempe, Arizona: Classroom Learning Laboratory, Arizona State University 1964.

Smith, W. I., and Moore, J. W. Learning Sets in Programed Instruction. Final Report, U. S. Office of Education Grant No. 7-48-0070-208, March, 1965.

Traub, R. E. Importance of problem heterogenity to programmed instruction. Journal of Educational Psychology, 1966, 57, 54-60.

Underwood, B. J., and Schulz, R. W. Meaningfulness and Verbal Learning. Chicago: Lippincott, 1960.

Underwood, B. J. Laboratory studies of verbal learning. In E. R. Hilgard (Ed.) Theories of Learning and Instruction. Sixty-third Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1964, 133-152.

Welborn, E. L., and English, H. B. Logical learning and retention: A general review of experiments with meaningful verbal materials. Psychological Bulletin, 1937, 34, 1-20.

Wiegand, Virginia K. A study of subordinate skills in science problem solving. Ph. D. dissertation, University of California, Berkeley, 1969.

Yoakam, G. A. The effect of a single reading. In Report of the Society's Committee on Silent Reading, NSSE Twentieth Yearbook, Part II. Chicago: National Society for the Study of Education, 1921.