DOCUMENT RESUME

ED 046 037	24	CG 006 128

AUTHOR	Romberg, Thomas A.; And Others
TITLE	Mastery Learning and Retention. Report from the Project on Analyses of Mathematics Instruction.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	TR-151
BUREAU NO	BR-5-0216
PUB DATE	Dec 70
CONTRACT	OEC-5-10-154
NOTE	21p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement, Learning Processes, *Mathematical Concepts, Memory, Performance, *Performance Factors, Recall (Psychological), *Retention, *Retention Studies, *Test Results

ABSTRACT

This paper reports results of two retention studies. In each, tests were given some time after instruction to a class of students whose initial level of performance was quite high. Results of these studies are summarized in terms of evidence related to the following questions: (1) to what extent are performances immediately following learning correlated with performances measured later? and (2) how much retention was there? For the first study on retention of probability concepts, the correlation between achievement scores immediately after learning and those obtained four weeks later was .78. Retention ratios were calculated for individuals (.60 to 1.05) for total test (.96), for each objective (.43 to 1.09), and for each item (.43 to 1.10). For the second study on recall of mathematical proofs the correlation between achievement scores immediately after learning and those obtained two weeks later was .75. Retention ratios were calculated for individuals (.33 to 1.25), for the total test (.93), for prerequisites (.97), for proofs (.79), and for correct steps within the proofs (.96). Results indicate that high initial performance may contribute to high retention. (Author)

MASTERY LEARNING AND RETENTION

RESEARCH
COGNITIVE LEARNING
DEVELOPMENT

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR
COGNITIVE LEARNING

ED0 46037

Technical Report No. 151

MASTERY LEARNING AND RETENTION

by

Thomas A. Romberg, Jack Shepler, and Irv King

Report from the Project on
Analyses of Mathematics Instruction
Thomas A. Romberg and John G. Harvey, Principal Investigators

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

December 1970

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

Center No. C-03 / Contract OE 5-10-154

SE 006 128

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Henry Chauncey
President
Educational Testing Service

Elizabeth Koontz
Wage and Labor Standards Administration, U.S. Department of Labor, Washington

Patrick Suppes
Professor
Department of Mathematics
Stanford University

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

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

Roderick McPhee
President
Punahou School, Honolulu

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

Francis S. Chase
Professor
Department of Education
University of Chicago

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

G. Wesley Sowards
Director, Elementary Education
Florida State University

RESEARCH AND DEVELOPMENT CENTER POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Russell J. Hosler
Professor, Curriculum and Instruction

Stephen C. Kleene
Dean, College of Letters and Science

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Donald J. McCarty
Dean
School of Education

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

Robert E. Grinder
Chairman
Department of Educational Psychology

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

Ira Sharkansky
Associate Professor of Political Science

M. Crawford Young
Associate Dean
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Robert E. Davidson
Assistant Professor, Educational Psychology

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

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Anne E. Buchanan
Project Specialist
R & D Center

Frank H. Farley
Associate Professor, Educational Psychology

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

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

Robin S. Chapman
Research Associate
R & D Center

FACULTY OF PRINCIPAL INVESTIGATORS

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor of Educational Psychology

James Moser
Assistant Professor of Mathematics Education; Visiting Scholar

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

Ted Czajkowski
Assistant Professor of Curriculum and Instruction

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

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Alan Voelker
Assistant Professor of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

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

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

Larry Wilder
Assistant Professor of Curriculum and Instruction

Gary A. Davis
Associate Professor of Educational Psychology

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

Thomas A. Romberg
Associate Director, R & D Center
Professor of Mathematics and of Curriculum and Instruction

Peter Wolff
Assistant Professor of Educational Psychology

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

Donald Lange
Assistant Professor of Curriculum and Instruction

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

MANAGEMENT COUNCIL

Herbert J. Klausmeier
Director, R & D Center
V.A.C. Henmon Professor of Educational Psychology

Thomas A. Romberg
Associate Director

James Walter
Director
Dissemination Program

Dan G. Woolpert
Director
Operations and Business

Mary R. Quilling
Director
Technical Development Program

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

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

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

CONTENTS

LIST OF TABLES AND FIGURES

I
THE PROBLEM OF RETENTION

Deese (1958) pointed out:

> There are two ways of looking at the problem of learning. We can examine the way in which behavior changes as a function of experience; ordinarily this means that we view improvement in performance as the result of practice. We can also take a fixed level of performance and see how that performance is retained over an intervening time interval. (p. 236)

The two studies here are of the latter type, retention studies, where the initial level of performance was quite high—children had reached specified levels of mastery.

Before discussing the studies, the two words, retention and forgetting, must be defined. *Retention* refers to the extent to which material originally learned is still retained and *forgetting* refers to the portion lost. Therefore,

> Amount forgotten = amount learned - amount retained. (Deese, 1958, p. 286)

Retention studies have shown a good deal of forgetting usually takes place soon after learning. Edwards and Scannell (1969) state, "From a functional standpoint we must expect forgetting and most of it will occur within a few hours after acquisition" (p. 320). However, McGeoch and Irion (1952) have shown that retention is better when the degree of original learning is high.

Mastery learning may be considered as one form of overlearning, where mastery learning is defined in terms of a class instructional procedure deliberately designed to get most students [perhaps over 90%] to reach some arbitrary and high [again, perhaps over 90%] level of performance on some task (Bloom, 1968). It can be argued that in order to reach such high levels of performance in a class many of the students must continue to practice after the task has been mastered, i.e., they are overlearning. Therefore, given students who actually mastered concepts and skills, one could hypothesize that there should be considerable retention of these concepts and skills over time.

One retention study, Gagné and Bassler (1963) could be considered as related to the studies reported here since instruction was based on learning hierarchies as are both these studies. Even though achievement on geometrical concepts was less than mastery (50.2% for five different groups), retention after 9 weeks was approximately the same (50.3%). However, four of the five groups actually had retention ratios ranging from 108% to 128%, while the other group fell off to 75%.

No study has been reported which examines retention after students have mastered mathematical concepts. Accordingly, the purpose of these two studies was to examine the extent concepts previously mastered were retained. Both studies are part of a sequence of exploratory studies which have been conducted in an attempt to gather information which will be helpful in determining the content and pedagogy of an emerging mathematics program being developed by the Analysis of Mathematics Instruction project of the University of Wisconsin Research and Development Center for Cognitive Learning (Romberg & Harvey, 1969).

II

STUDY #1: Retention of Probability Concepts

This study grew out of a successful mastery learning curriculum study, details of which appear in a previous Technical Report (Shepler, 1970).

The purpose of that study was two-fold: (1) To test the feasibility of teaching topics in probability and statistics to a class of Sixth Grade students; and (2) to construct a set of instructional materials and procedures in probability and statistics for Sixth Grade students.

The working paper of Shepler, Harvey, and Romberg (in preparation) and the developmental model of Romberg and DeVault (1967) were used to build the unit. Shepler, *et al.*, constructed a framework for the development of an instructional system in probability and statistics for use in the elementary school. That paper included a content outline, a task analysis of content, and specific grade recommendations for topics used in elementary school.

From parts of the task analysis an instructional analysis of the unit was undertaken to select or develop materials and procedures for teaching the unit. The goal of instruction was to demonstrate "mastery learning" of the behavioral objectives of the unit.

To achieve mastery learning of the objectives, recommendations from Bloom (1968) were incorporated into the instructional procedures. These were: (1) communicating the objectives of the unit by the use of a large goal chart, (2) mastery or nonmastery grading of quizzes and exercises (no A, B, C, D, or F grading), (3) opportunity for nonmasters of an exercise or quiz to achieve mastery by correcting the exercise or taking a parallel quiz, (4) using specific prescriptions to pinpoint errors and suggest what he could do to them, (5) extra help sessions for students who had not been successful on a set of goals, (6) assigning a consistent nonmaster learner to a master learner for extra help, (7) diagnosing whether most students had achieved prerequisite skills before introducing a task dependent on those skills, (8) further group instruction whenever fewer than 80% of students achieved mastery at first time of testing, and (9) using the extrinsic reward of awarding a diploma to a student who was a master of probability concepts in the elementary school. These instructional procedures were employed in teaching the unit.

These daily lessons were formatively evaluated. Guidelines set as a criteria for judging successful instruction were stated beforehand. An attempt was made to identify the weaknesses and strengths of each lesson. From the information gathered on the exercise that accompanied a lesson, on short quizzes, and on written observations by the teacher and the observer modifications of the unit and procedures were made.

In that study, 25 selected Sixth Grade students from Waunakee Elementary School, Waunakee, Wisconsin, were taught this unit on probability and statistics concepts. A test designed to measure 14 of the behavioral objectives was used both pre- and post-instruction to see what changes occurred as a result of instruction.

Overall, the group's mean performance changes from 27.28 (37.9%) on the pretest to 66.80 (92.8%) out of 72 items on the posttest; 11 of the 14 specified objectives were mastered; and 21 of the 25 students (84%) were considered master learners (answered correctly 90% or more of the items), while all had answered correctly at least 80% of the items.

THE RETENTION STUDY

To measure students' retention, the same test was administered exactly 4 weeks after the posttest. Other than brief comments after presenting the diplomas 1 week after posttesting, no instruction or practice was given the students concerning probability concepts.[1] During this 4-week period students studied percentages and decimals with their regular teacher in the mathematics class.

THE TEST

The same test was used as the pretest, the posttest, and the retention test. It consisted of 72 items measuring 14 behavioral objectives. Of the 72 items, 36 items were one-dimensional sample space problems; 19 items were two-dimensional sample space problems; 7 items were one-dimensional and two-dimensional sample space problems, and 10 items were on ordering of two fractions.

The models employed in these items were coins, dice, spinners, and boxes containing various objects. The items were based on one- and two-dimensional finite sample spaces generated by these models.

The behavioral objectives measured in the test and the number of items for each are as follows:

The child [illegible] be able, given the written test which used the previously mentioned models, to:

1. Distingu[illegible]ther an event is an instance of [illegible]nty, uncertainty, or impossibility [illegible] items).
2. Count the [illegible] of outcomes of an event [illegible].
3. Count [illegible]ber of possible outcomes of a sa[illegible]ce (7 items).
4. Specify [illegible]bability of a simple event (8 items[illegible]
5. Specify [illegible]robability of a compound event [illegible].
6. Specify [illegible]ability of a certain event (2 items).
7. Specif[illegible]ability of an impossible event [illegible]).
8. Specify the order of two fractions between 0 and 1 (10 items).
9. Identify the most likely event of two unequally likely events (10 items).
10. Identify two equally likely events as being equally likely (8 items).
11. Specify the estimated probability of an event, given the data from an experiment (1 item).
12. Identify the likely bounds of the frequency of an outcome of an experiment that has been done n times (1 item).
13. Identify an instance of the law of averages (4 items).
14. Identify an estimate of the true probability, given a set of data from an experiment (1 item).

The 72 items were divided into 14 parts labeled R1, R2, ..., R14.* The test was designed to be given in two testing sessions. Section A, to be administered first, consists of the even-numbered parts R2, R4, R6, ..., R14. Section B consists of the odd-numbered parts R1, R3, R5, ..., R13. Each section was assigned one of four random orders from a random number table. The items were then assembled to form Section A or Section B. This randomizing of the orders of Sections A and B was done to rule out the effect of taking a test in a specific order and to minimize the opportunities for cheating.

The same set of directions was attached to the front of each section. The directions were read to the students after the test was passed out. A more complete description of the test, its construction, and its characteristics appears in Shepler (1970).

CRITERIA

The decision as to whether class performance was successful is based on the number of students (k) and the number of items (n) used to measure an objective. For this study there are 25 students and 72 items on the complete test. Both the arbitrary 90/90 criterion level and the following practical criteria were used to judge overall success. The practical criteria used were that a criterion level was obtained if at least 22 of the 25 students achieved $\frac{n-1}{n}$ ($n \geq 5$) or n/n ($n < 5$) on an objective measured at least by n items.

RESULTS

The results of this study are summarized in terms of evidence related to the following questions:

---

[1]However, the items on the posttest concerned with question, "How many ways can you get ... ?" and item 13-1 on the approximate probability were discussed at that time. But, there was no blackboard to write on. Thus, the discussion was purely a verbal discussion of previously mentioned problems.

*There is no relationship between Objectives 1 to 14 and Labels R1 to R14.

1. To what extent are performances immediately following learning correlated with performances measured 4 weeks later?
2. How much retention was there?

For the total group of 25 subjects the correlation between achievement scores immediately after learning and those obtained 4 weeks later was .78. This high relationship indicates considerable consistency between what was learned and what was retained. The individual data (see Table 1) clearly support this consistency. From the posttest and retention test, 7 of the 25 students' scores improved, 2 scores remained the same, while 16 scores declined. Of the 16 scores which declined, 8 scores dropped 1 or 2 raw score points; 4 scores dropped 3 or 4 points; and 4 scores dropped more than 4 points. (Subject 1 dropped 24 points; Subjects 2 and 19 dropped 9 points; and Subject 14 dropped 6 points.)

It should be noted that Subjects 1 and 2 had the two lowest scores on both pretest and posttest. Thus it is not surprising that they also had the two lowest scores on the retention test. On the retention test, Subject 1 forgot to express the probability of an event as a ratio.

The 90/90 criteria for the complete test meant that at least 22 students should have answered 65 or more of the items correctly. In fact, only 21 of 25 students (84%) performed up to criteria on the posttest. However, the four who failed to reach 90% reached at least 80% (58 items answered correctly). A comparison of posttest and retention test levels of performance (see Table 2) show that of the 21 masters, 17 were still above a 90% level of performance 4 weeks later and that the other four masters were still above an 80% level of performance. Of the four non-masters two remained above 80%. Retention ratios (amount retained÷amount learned) also varied from 60% to 105% with 21 of 25 having retention ratios of .94 to 1.05.

Table 1. Individual Results on the Pretest, Posttest, and Retention Test on Probability (Number of Correct Responses on the 72-Item Test)

Subject	Pretest (before instruction)	Posttest (immediately after instruction)	Retention (4 weeks after instruction)	Retention Ratios
1	7	58*	34	.60
2	14	60*	51	.85
3	30	71	70	.99
4	21	69	70	1.01
5	22	68	67	.99
6	23	64*	64	1.00
7	25	69	66	.96
8	37	68	66	.97
9	26	67	70	1.04
10	34	67	67	1.00
11	37	67	70	1.04
12	28	66	64	.97
13	35	68	70	1.03
14	19	65	59	.91
15	23	70	68	.97
16	44	69	71	1.03
17	28	61*	64	1.05
18	22	69	67	.97
19	18	69	60	.87
20	34	67	64	.97
21	38	69	65	.94
22	26	65	67	1.03
23	22	69	67	.97
24	31	68	67	.99
25	34	69	68	.99

*Non-Masters on Posttest.

Table 2. Level of Performance from the Posttest and the Retention Test on Probability

		Level of Performance Retention Test					
Level of Performance Posttest		40-50	50-60	60-70	70-80	80-90	90-100
	80-90	1	0	0	1	2	0
	90-100	0	0	0	0	4	17

Table 3. Overall Results of the Pretest, Posttest, and Retention Test on Probability

	Pretest	Posttest	Retention Test
Mean (72 items)	27.28	66.80	64.40
Mean in terms of percentage	37.9%	92.8%	89.50%
Variance	74.13	11.17	56.95
Retention Ratio			.96

In all, the results indicate that how much was retained depended upon how much was learned.

AMOUNT OF RETENTION

The amount retained by the total group is reported in terms of the overall test, each objective, and each item.

For the complete test the 25 students in the study had a pretest mean of 27.28, a posttest mean of 66.80, and a retention test mean of 64.40. The results are summarized in Table 3. The variance was large for the pretest, small for the posttest, and somewhat larger again for the retention test. The high retention ratio (.96) indicates that most of what was learned was retained.

For behavioral objectives on the posttest, subjects attained the criteria level for 11 out of 14 objectives. Only for Objectives 2 (number of outcomes of an event), 8 (estimated probability), and 14 (estimate of probability) did the students not meet criteria.

On retention test, 7 out of the 11 previously attained objectives remained above the criteria levels (see Table 4). Objectives 2, 11, and 14 still did not meet the criteria level, and while performance on 8, 9, 10, and 12 now failed to reach criteria level they still had high level of performance. Thus, in general, the performance of children on the retention test by objectives was similar to their overall posttest performance. (See Figure 1.)

In addition, the retention ratios on the objectives were generally high, ranging from .43 to 1.09 with ratios of .80 or more for 11 of the 14 objectives.

The data on specific items also support the previous results. The change in item percentages for the 72 items from pretest to posttest to retention test was encouraging. All items had a positive change from pretest to posttest with 6 items having a 100% change, 10 items having an 80% to 99% change, 19 items having a 60% to 79% change, 12 items having a 40% to 59% change, 17 items having a 20% to 39% change, and 8 items having a 0% to 19% change.

However, there was little change from posttest to the retention test: 37 items decreased in item percentages (average decrease—8.6%), 12 items increased (average increase—5%) and 23 items stayed the same. Table 5 summarizes the item percentages. It should also be noted that the retention ratios for each item are, in general, very high, with ratios above .90 for 63 of the 72 items.

Table 4. Ratio of Children Reaching a Criterion on the Posttest and Retention Test on Probability

	POSTTEST		RETENTION TEST		
Behavioral Objective	Ratio of Children Reaching Test Criterion	Test Criterion	Ratio of Children Reaching Test Criterion	Test Criterion	Retention Ratios
1	* 24/25	4/5	* 23/25	4/5	.96
2	18/25	4/5	18/25	4/5	1.00
3	* 22/25	6/7	* 23/25	6/7	1.05
4	* 25/25	7/8	* 24/25	7/8	.96
5	* 23/25	7/8	* 22/25	7/8	.96
6	* 22/25	2/2	* 24/25	2/2	1.09
7	* 24/25	2/2	* 24/25	2/2	1.00
8	* 23/25	9/10	19/25	9/10	.83
9	* 23/25	9/10	20/25	9/10	.87
10	* 23/25	7/8	17/25	7/8	.74
11	13/25	1/1	7/25	1/1	.54
12	* 25/25	1/1	21/25	1/1	.84
13	* 22/25	3/4	* 22/25	3/4	1.00
14	7/25	1/1	3/25	1/1	.43

*Objectives reaching criteria.

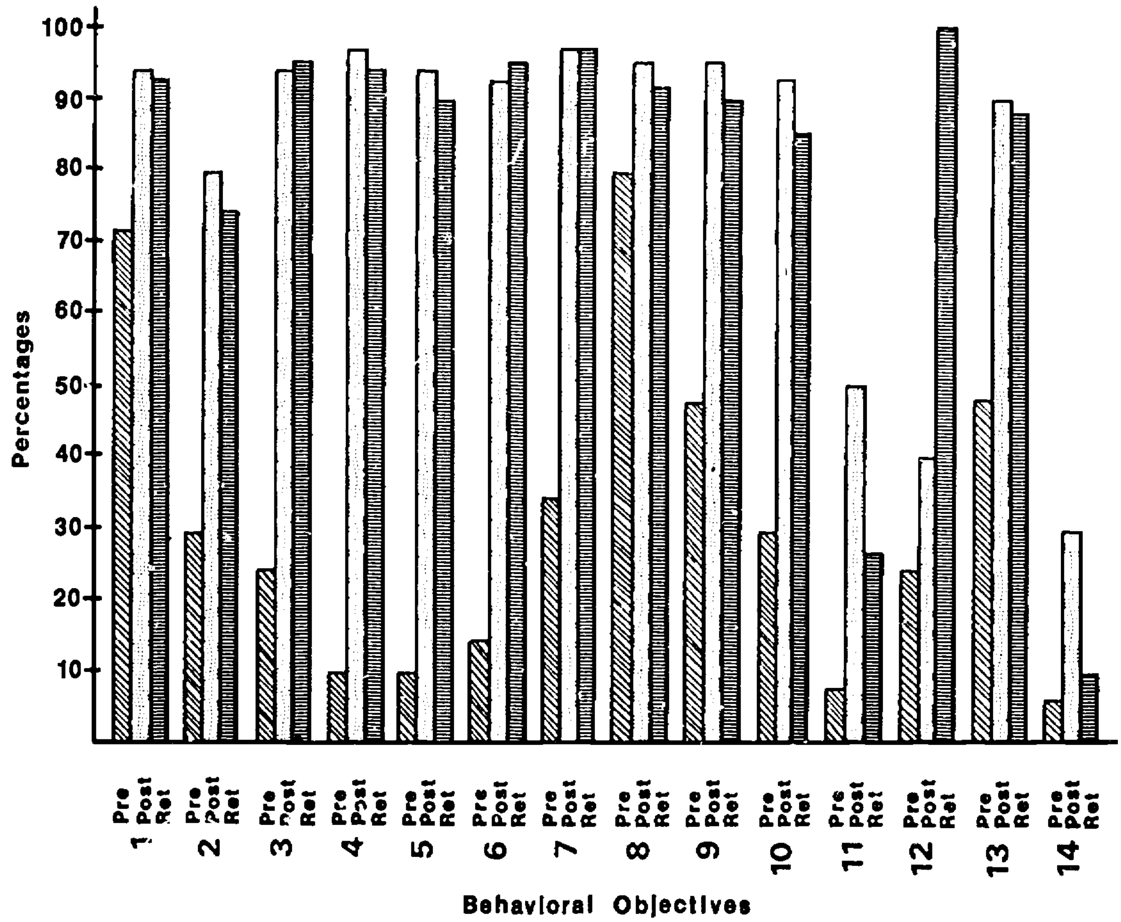

Figure 1. Percentage Graph of Behavioral Objectives as Measured by the Pretest, Posttest, and Retention Test on Probability.

Table 5. Item Percentages of the Pretest, Posttest, and Retention Test on Probability

		Total Number of Students = 25			
		PRETEST % of Correct Responses	POSTTEST % of Correct Responses	RETENTION TEST % of Correct Responses	Retention Ratio
R1	1	72	100	96	.96
	2	92	100	100	1.00
	3	72	92	92	1.00
	4	60	96	84	88
	5	56	88	88	1.00
	6	40	100	84	.84
	7	64	88	88	1.00
	8	60	88	88	1.00
R2	1	40	92	96	1.04
	2	32	100	92	.92
R3	1	96	100	100	1.00
	2	88	100	96	.96
	3	64	96	92	.96
	4	96	100	100	1.00
	5	96	100	100	1.00
	6	72	92	96	1.04
	7	68	96	72	.75
	8	88	100	100	1.00
	9	68	92	92	1.00
	10	52	84	76	.91
R4	1	36	96	76	.80
	2	28	88	72	.82
	3	40	96	96	1.00
	4	16	92	88	.95
	5	16	92	92	1.00
	6	8	92	80	.87
	7	52	80	84	1.05
R5	1	32	92	96	1.04
	2	60	96	100	1.04
	3	72	100	96	.96
	4	32	100	96	.96
	5	68	100	96	.96
	6	60	96	68	.71
	7	52	96	96	1.00
	8	52	100	96	.96
	9	52	100	92	.92
	10	28	92	84	.91
	11	32	92	92	1.00
R6	1	44	96	88	.92
	2	32	100	96	.96
	3	24	100	96	.96
	4	12	96	96	1.00
	5	48	100	100	1.00

(continued)

Table 5 (continued)

		Total Number of Students = 25			
		PRETEST	POSTEST	RETENTION TEST	
		% of Correct Responses	% of Correct Responses	% of Correct Responses	Retention Ratio
R7	1	48	100	100	1.00
	2	84	100	100	1.00
	3	32	100	96	.96
	4	32	100	96	.96
R8	1	4	100	96	.96
	2	16	64	64	1.00
	3	0	100	96	.96
	4	8	92	80	.87
R9	1	0	92	100	1.09
	2	24	88	88	1.00
	3	0	100	96	.96
	4	0	84	92	1.10
	5	20	96	96	1.00
R10	1	0	92	100	1.09
	2	0	100	96	.96
	3	0	100	96	.96
	4	4	100	96	.96
	5	20	92	96	1.04
R11	1	84	92	92	1.00
	2	28	72	76	1.06
	3	24	96	96	1.00
	4	16	92	88	.95
R12	1	0	88	92	1.05
	2	4	76	80	1.05
	3	0	100	92	.92
	4	0	96	88	.92
	5	0	100	88	.88
R13	1	8	52	28	.54
R14	1	4	28	12	.43

III
STUDY #2: Recall of Mathematical Proofs

This retention study was based on a successful exploratory study which examined the feasibility of presenting proofs of mathematical theorems to Sixth Grade students (King, 1970).

In conducting the learning study, a unit of instruction on mathematical proof was developed in accordance with the iterative, curriculum development model of Romberg and DeVault (1967). The model calls for a tentative unit to be written and piloted with a group of students. If satisfactory results are not obtained, the unit is rewritten and repiloted with another group of students. This iteration continues until a viable unit is developed. In that study, three formative pilot studies were completed before an effective unit on proof was developed.

The unit was built around the following six properties of whole numbers:

Theorem 1: If N divides A and N divides B, then N divides their sum (A + B).

Theorem 2: If N divides A and N divides B, then N divides their difference (A - B).

Theorem 3: If N divides A and N divides B and N divides C, then N divides their sum (A + B + C).

Theorem 4: If N divides A and N does not divide B, then N does not divide their sum (A + B).

Theorem 5: If N divides A and N does not divide B, then N does not divide their difference (A - B).

Theorem 6: Given any set of prime numbers {2, 3, 5, . . . , P}, there is always another prime number.

The unit included six major behavioral objectives of the following type: Given the statement of Theorem 1, the student can write a valid proof of the theorem. Each of these six objectives was subjected to a task analysis. The task analytic procedure was developed by Gagné (1961) to train human beings to perform complex tasks. Beginning with a particular objective, one asks: What must a person be able to do in order to perform this task? In this way a number of subordinate behaviors are identified. Each subordinate behavior is itself broken down into a set of subordinate behaviors. The process continues until a set of elementary behaviors are reached.

A hierarchy of behaviors is thus formed. If the hierarchy is properly constructed, the learner can begin at the bottom of the hierarchy with the simplest task and progress step-by-step up the hierarchy until he reaches the desired terminal behavior. The learning of each skill increases the probability of learning subsequent skills in the hierarchy.

Proving a mathematical theorem is a highly complex task. In addition to knowing a set of mathematical ideas which are the basic ingredients of the proof, the prover must also have a plan or strategy which permits him to link these ideas together to form a proof. The unit on proof stressed both aspects of proof, the basic mathematical ideas and the strategy to be used.

In the original study, ten students (mean IQ 117) received 17 days of instruction on proof. The concept of Mastery Learning was employed. Each student was told that he could earn an "A" for the unit if he mastered the material. Time was not to be a factor and each child would be given as much help as needed to master the content of the unit. This was operationally feasible because the class was limited to ten students, thus permitting the teacher to work with individuals as they needed assistance. Each step in the learning hierarchy was mastered before the students proceeded to the next step.

When instruction was completed, the students were given a 25-item test on both subordinate skills and the six proofs. The students responded correctly to 97.8% of the subordinate skills. The proofs were scored in two ways: (1) on an all-or-nothing basis, the proof being either valid or invalid; (2) on a partial credit basis, the student receiving credit for each part of the proof which is correct. Fifty-eight of the sixty proofs were valid (96.6%) and 498 of the 500 possible steps in the proofs were correct (99.6%).

THE TEST

The same test was used as a pretest, posttest, and the retention test. It consisted of 25 items; 19 measuring prerequisite skills and 6 proofs. A more complete description of the test, its construction, and its characteristics, appears in King (1970).

The separate scores were found for the prerequisite skills, proofs, and number of correct steps in each proof.

THE RETENTION STUDY

After the posttest was administered, the students returned to their regular Sixth-Grade mathematics class where they studied fractions. To the investigator's knowledge the students were not re-exposed to the ideas contained in the unit on proof. The 25-item posttest was readministered to the students after a period of 14 days had elapsed.

RESULTS

The results of the test provide evidence in the following four categories:

(1) To what extent are performances immediately following learning correlated with performances measured 2 weeks later?
(2) How much retention was there?
(3) Is the rate of forgetting a function of the learner's ability?
(4) As compared to the retention of proofs, what happens to the retention of subordinate skills?

For the total group of 10 students the correlation between achievement scores immediately after learning and those obtained 2 weeks later was .75. This high relationship indicates considerable consistency between what was learned and what was retained. The individual data (see Table 6) clearly support this consistency. From the posttest to the retention test 1 of the 10 students' scores was higher on prerequisites, 5 remained the same, and 4 scores declined. Similarly on proofs 1 score was higher, 5 scores remained the same, and 4 scores declined. Of the scores which declined on the prerequisites 2 scores dropped 1 point, 1 dropped 3 points, and 1 dropped four points; for the proofs the decline of correct steps where 1 score dropped 2 points, 1 dropped 4 points, 1 dropped 6 points, and the last dropped 7 points.

AMOUNT OF RETENTION

Table 7 summarizes the individual data for the pretest, the posttest administered 14 days later. For the complete test the 10 students in the study had a pretest mean of 5.70, posttest mean of 24.3, and a retention test mean of 22.5. The high total retention ratio (.93) indicates that most of what was learned was retained.

For the two parts of the test and the number of steps the amount retained is summarized in Table 8. The correct responses on subordinate skills dropped from 97.8% to 93.1% and the number of valid proofs dropped from 96.6% to 76.6%. However, when the proofs are scored on the basis of the total number of correct steps in the proofs, the percentage drop is considerably less, from 99.6% to 95.8%. Hence, while the total number of correct proofs dropped some 20%, the number of correct steps declined only about 4%.

RATE OF FORGETTING AND LEARNER'S ABILITY

Classical retention studies indicate that forgetting is not a function of the learner's ability. Underwood (1966) summarized the results as follows:

> Allowed enough time to study a list so that he can reproduce it as readily as a fast learner, a slow learner will score as high as a fast one in later tests of remembering the list.

Table 6. Individual Results on the Pretest, Posttest, and Retention Test on Reasoning
(Number of Correct Responses on 19 Prerequisite Items, 6 Proofs and 50 Steps)

	Pretest (before instruction)			Posttest (immediately after instruction)			Retention Test (2 weeks after instruction)			Retention Ratio		
Subject	Prereq.	Proofs	Steps	Prereq.	Proofs	Steps	Prereq.	Proofs	Steps	Prereq.	Proofs	Steps
1	4	0	0	19	6	50	16	6	50	.84	1.00	1.00
2	7	0	0	19	6	50	19	6	50	1.00	1.00	1.00
3	10	0	0	19	6	50	18	2	44	.95	.33	.88
4	4	0	0	17	4	48	16	5	48	.94	1.25	1.00
5	9	0	0	19	6	50	19	4	48	1.00	.67	.96
6	7	0	0	19	6	50	19	2	43	1.00	.33	.87
7	8	0	0	19	6	50	19	6	50	1.00	1.00	1.00
8	1	0	0	18	6	50	14	6	50	.78	1.00	1.00
9	3	0	0	18	6	50	18	6	50	1.00	1.00	1.00
10	4	0	0	18	6	50	19	3	46	1.06	.50	.92

Table 7. Overall Results of the Pretest, Posttest, and Retention Test on Reasoning

	Pretest	Posttest	Retention Test
Mean (25 items)	5.70	24.3	22.5
Mean in terms of Percentage	22.8%	97.2%	90.0%
Retention Ratio			.93

Table 8. Mean and Percentage of Correct Responses on the Posttest and the Retention Test on Reasoning

Subordinate Tasks	Posttest	Retention Test	Retention Ratios
Correct Responses			
Mean	18.5	17.9	.97
Percentage	97.8%	93.1%	
Correct Proofs			
Mean	5.8	4.6	.79
Percentage	96.6%	76.6%	
Correct Steps in the Proof			
Mean	49.8	47.9	.96
Percentage	99.6%	95.8%	

Since each student in the experimental class was given the time and help to master all of the concepts and proofs in the unit, the expectation is that the rate of forgetting is not related to the learner's ability. To examine this hypothesis, Henmon-Nelson IQ scores were used as a measure of "the learner's ability." The product-moment correlation coefficient between the IQ scores and the scores on the retention test was computed to be $r = -.71$. This is a fairly strong negative relationship, indicating that in this study the slower learner tended to forget less than the faster learner. Although this finding is only suggestive due to the small number of students, a possible explanation for this unexpected result may lie in the nature of the original study. The slower learners, knowing that grades would be given for mastery, worked very diligently to learn of the original study. In fact, on most occasions the slower learners studied at home and learned the proofs sooner than the better students. Whereas all of the students had mastered all of the proofs, it is quite possible that the slower students had overlearned the material to a greater extent, thus accounting for a better performance on the retention test.

If the foregoing analysis is correct, the results of this study indicate the mastery learning strategy is particularly viable for slower learners.

One further observation is worthy of mention. In the original study, the students employed several different strategies in memorizing proofs. Most of the students (eight of the ten) used the following approach: when asked to write a proof, they would first write down the number of steps in the proof, fill in all the statements (or reasons), and then complete the proof by filling in the reasons (or statements). The other two students attempted to reason through the proof each time they wrote it. Statement 1 was written first, then Reason 1 was written; then Statement 2, Reason 2, etc. In short, these two students were not learning the theorems by rote procedures as the rest of the class were apparently doing.

When the retention test was administered both of these students encountered more difficulty than the other students. They took longer to complete the test, and they did so only after many false and frustrating starts. Furthermore, their scores were the lowest in the class. Bartz (1968) points out that learning strategies affect amount of retention. The results of this study indicate that those students who learned the proofs by specific patterns were less likely to forget the proofs than those who learned the proofs by "thinking it through."

COMPARISON OF SUBORDINATE SKILLS AND PROOFS

Gagné and Bassler (1963) conducted a retention study in which the forgetting of subordinate learning sets occurred independently of, and without effect upon, retention of the total terminal task. The terminal tasks, defined as "specifying sets, interactions of sets, and separations of sets using points, lines, and curves" were remembered even though many of the subordinate skills had been forgotten.

The same phenomenon occurred in the present study. Forty-six of the sixty proofs were valid, yet in twelve of these proofs the student had missed one or more of the subordinate skills. More surprising was the fact that all of the subordinate items for each of the fourteen invalid proofs had been answered correctly. As in Gagné and Bassler's study (1963), the forgetting of subordinate tasks seem to occur independently of, and without effect upon, retention of the terminal task.

This raises a serious question. If a student can perform the terminal task without being able to perform a subordinate task, is that subordinate task really necessary? Gagné and Bassler (1963) argue that the subordinate tasks are necessary for learning the terminal task but, once the terminal task has been learned, the subordinate behaviors, having served their function, may be forgotten without adversely effecting the performance of the terminal task.

A similar explanation might apply to the present study. It is possible to memorize a proof without understanding it. In the original study a great deal of emphasis was placed upon understanding the proofs. In addition to reproducing the proofs, each student was required to explain and defend his proofs in an interview situation and to give numerical examples to illustrate various parts of the proofs. To this extent, the students understood the proofs when tested immediately after instruction. For example, nine of the ten students displayed an understanding of the proof of Theorem 1 immediately after instruction, but only three students were able to do so on the retention test. These results suggest that subordinate skills must be mastered if one is

GPO 820-359-3

to understand a proof, but that they are not necessary for reproducing a proof.

In summary, the data from this study clearly support the feasibility of teaching mathematical proofs to Sixth Grade students so that they are retained. In addition, this study raises several questions related to learning, over-learning, and forgetting, all of which suggest the design of subsequent studies.

IV
CONCLUSIONS

Two studies involving Sixth Grade pupils were undertaken to determine whether students could retain information on new mathematical topics (probability and proof) after instruction involving mastery learning. The ratios of mean scores on the two retention tests to mean scores on the posttests given immediately after instruction were .96 and .93. Retention tests were administered 4 and 2 weeks, respectively, after the posttests. These results indicate clearly that students maintained a high level of achievement after instruction. Although it seems unlikely that large item gains and high retention ratios could be due to test-treatment interaction, the design used in each study did not control for this effect. In fact, the same tests served as pretest, posttest, and retention tests. Further retention studies in which parallel tests are used are needed to determine whether retention ratios remain as high. If such studies corroborate these findings, then the further use of mastery learning principles with this age group is recommended.

REFERENCES

Bartz, Wayne H. *Memory*. Wm. C. Brown Company, Dubuque, Iowa, 1968, p. 18.

Bloom, Benjamin S. Learning for Mastery. *UCLA: Evaluation Comment, 1*, No. 2, May 1968.

Deese, James. *The Psychology of Learning.* New York: McGraw-Hill Book Co., 1958.

Edwards, Allen J. & Scannell, Dale P. *Educational Psychology the Teaching-Learning Process*. Scranton: International Textbook Company, 1969. Pp. 315-355.

Gagné, Robert M. & Bassler, Otto C. Study of Retention of Some Topics of Elementary Non-metric Geometry. *Journal of Educational Psychology*, 1963, *54*, No. 3, Pp. 123-131.

Gagné, Robert M. and Paradise, Noel E. "Abilities and Learning Sets in Knowledge Acquisition." *Psychological Monograph*, 1961, *75* (14, Whole No. 518).

King, Irv. A formative development of a unit on proof for use in the elementary school. *Technical Report from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin*, 1970, No. 111.

McGeoch, J. A. & Irion, A. L. *The psychology of human learning* (2nd. Edition) New York: Longmans, 1952.

Romberg, Thomas & DeVault, M. Vere, Mathematics Curriculum: Needed Research. *Journal of Research and Development in Education, 1*, No. 1, Fall 1967. Athens: The University of Georgia, 1967.

Romberg, Thomas A. & Harvey, John G. Developing mathematical processes: Background and projections. *Working Paper from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin*, 1969, No. 14.

Shepler, Jack L. A Study of Parts of the Development of a Unit in Probability and Statistics for the Elementary School. *Technical Report from the Wisconsin R & D Center for Cognitive Learning, University of Wisconsin*, 1970, No. 105.

Shepler, Jack L., Harvey, John G., and Romberg, Thomas A. Probability and Statistics. *Working Paper from the Wisconsin R & D Center for Cognitive Learning, University of Wisconsin*, (in preparation).

Underwood, Benton J. "Forgetting. *Frontiers of Psychological Research*. W. H. Freeman and Co. San Francisco, 1966, p. 156.