Best
Available
Copy

AD/A-000 576

ACQUISITION AND RESILIENCE UNDER TEST STRESS OF STRUCTURALLY DIFFERENT PROBLEM SOLVING PROCEDURES

Richard E. Mayer

Michigan University

Prepared for:

Air Force Office of Scientific Research
Advanced Research Projects Agency

May 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
DIST. AVAIL. and/or SPECIAL
A

THE HUMAN PERFORMANCE CENTER

DEPARTMENT OF PSYCHOLOGY

The Human Performance Center is a federation of research programs whose emphasis is on man as a processor of information. Topics under study include perception, attention, verbal learning and behavior, short- and long-term memory, choice and decision processes, and learning and performance in simple and complex skills. The integrating concept is the quantitative description, and theory, of man's performance capabilities and limitations and the ways in which these may be modified by learning, by instruction, and by task design.

The Center issues two series of reports. A Technical Report series includes original reports of experimental or theoretical studies, and integrative reviews of the scientific literature. A Memorandum Report series includes printed versions of papers presented orally at scientific or professional meetings or symposia, methodological notes and documentary materials, apparatus notes, and exploratory studies.

Security Classification

DOCUMENT CONTROL DATA - R & D AD/A-000 576

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) University of Michigan, Human Performance Center Department of Psychology, Ann Arbor, Michigan 48104	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3 REPORT TITLE ACQUISITION AND RESILIENCE UNDER TEST STRESS OF STRUCTURALLY DIFFERENT PROBLEM SOLVING PROCEDURES	
4 DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific Interim	
5 AUTHOR(S) (First name, middle initial, last name) Richard E. Mayer	

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO. OF REFS
May, 1973	95	25

8a. CONTRACT OR GRANT NO F44620-72-C-0019 b. PROJECT NO 1949 c. 61101D d. 681313	9a. ORIGINATOR'S REPORT NUMBER(S) 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR - TR - 74 - 1659
10 DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.	
11 SUPPLEMENTARY NOTES TECH, OTHER	12 SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research 1400 Wilson Boulevard (NL) Arlington, Virginia 22209

13 ABSTRACT

A procedure for solving binomial probability problems was taught in a three-part booklet that emphasized calculating with the formula (Sequence F) or that emphasized the meanings of the variables in the formula (Group G). In two main experiments, a multi-leveled transfer posttest was administered after S had read either one, two or all three parts of his instructional booklet, and the amount of stress during testing was varied. Two supplemental studies were also performed. The results indicated that, at all three points in learning, there was a clear pattern of treatment x posttest interaction (TPI) in which Group F Ss excelled on near transfer items and Group G Ss excelled on far transfer or interpretive posttest items. In addition the presence of stress during testing, while producing an overall quantitative decrease in performance, did not disrupt the pattern of TPI. The results were discussed with respect to the acquisition process for, and the apparent permanence under stress of, structurally different problem solving procedures.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

Unclassified
Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Instructional Sequence						
2. Multi-Leveled Transfer Posttest						
3. Time Stress						
4. Treatment x Posttest Interaction						

Unclassified
Security Classification

THE UNIVERSITY OF MICHIGAN

COLLEGE OF LITERATURE, SCIENCE AND THE ARTS

DEPARTMENT OF PSYCHOLOGY

ACQUISITION AND RESILIENCE UNDER TEST STRESS OF STRUCTURALLY DIFFERENT PROBLEM SOLVING PROCEDURES

Richard Edwin Mayer

HUMAN PERFORMANCE CENTER--TECHNICAL REPORT NO. 42

May, 1973

This research was supported by the Advanced Research Projects Agency, Department of Defense, and monitored by the Air Force Office of Scientific Research, under Contract No. F44620-72-C-0019 with the Human Performance Center, Department of Psychology, University of Michigan.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release; distribution unlimited.

ii b

PREFACE

This report is an independent contribution to the program of research of the Human Performance Center, Department of Psychology, on human information processing stress factors, supported by the Office of the Advanced Research Projects Agency, under Order No. 1949, and monitored by the Air Force Office of Scientific Research under Contract No. F44620-72-C-0019.

This report was also a dissertation submitted by the author in partial fulfillment of the degree of Doctor of Philosophy (Psychology) in the University of Michigan, 1973. The doctoral dissertation committee was: Drs. J. G. Greeno, Chairman, R. A. Bjork, S. Kaplan, and J. N. Payne.

TABLE OF CONTENTS

Preceding page blank

ABSTRACT

Previous studies (Mayer & Greeno, 1972; Greeno & Mayer, in preparation) have suggested that teaching Ss to solve certain kinds of problems by different instructional methods may result in learning outcomes which differ in structural or qualitative ways as indicated by a pattern of treatment x posttest interaction (TPI). Two new questions dealt with in this dissertation were: (1) How can the acquisition process for structurally different learning outcomes be characterized? (2) Once established, can the structure of learning outcomes be altered by post-experimental manipulations such as stress during testing?

In two main experiments, the concept of binomial probability was taught to S via expository four-lesson teaching booklets, by two instructional methods that differed in sequencing and emphasis. One instructional method (Sequence F) began each lesson with a formal statement of the rule (or sub-rule) and explained component variables only within the context of calculating with the formula; the other method (Sequence G) began each lesson by attempting to relate component variables to S's general experience, e.g., with "trials," "outcomes," "successes," etc., before presenting any formal statement of the rule. Learning was assessed by a multileveled transfer posttest which contained both near and far transfer items.

To provide information on the acquisition question, the posttest was administered at three points in learning for S's in both instructional groups. Some Ss were tested after the first two lessons (introduction and combinations), some after three lessons (introduction, combinations, joint probability), and some after the entire four-lesson booklet (introduction, combinations, joint probability, and binomial probability) had been presented. Both Experiment I and Experiment II used this procedure although the ordering of booklet lessons differed.

To provide information on the resilience question, the conditions of testing were varied with some Ss in each instructional group tested under more stress than others. In Experiment I, the stress variable was presence or absence of memory support during testing. In Experiment II the stress variable was presence or absence of time stress during testing. Two supplemental studies were also performed.

The main results with respect to acquisition were the following:
(a) The overall percent correct on the posttest increased for both instructional groups as the amount of instruction was increased.
(b) The same treatment x posttest interactions were generally found at all three points in learning, with Sequence F excelling on near transfer items and Sequence G excelling on far transfer items. Thus, there was no evidence of any structural change in "what is learned"

as learning progressed. (c) The combinations lesson was far more important in producing increases in posttest performance than the joint probability lesson. However, this was especially true for the Ss in Sequence F, suggesting that Ss in Sequence G were better able to create solutions after being exposed to only a part of the material than were Ss in Sequence F. Findings were discussed in conjunction with the notion that different kinds of existing knowledge were activated early in learning by Ss in different instructional groups, and that those different sets were used consistently throughout learning in order to assimilate the content material.

The main results with respect to the resilience question were the following: (a) Stress during testing, both memory load and time stress, had an overall quantitative effect on problem solving performance. (b) Test stress did not seem to alter the pattern of TPI for the two instructional groups, and hence, gave no evidence of altering the structural properties of established learning outcomes. Findings were discussed in relation to the apparent permanence of the effects of instructional method.

CHAPTER I

INTRODUCTION

There has been much written about the importance of "how" someone is taught to solve problems, especially mathematical problems, of a given class. The "how" issue has often centered on an attempt to show that one instructional method (e.g., "discovery method") is better or worse overall than another method (e.g., "reception method"). Wittrock (1966), Hermann (1969), and Mayer (1972) have reviewed this line of research; and in 1966, Shulman and Keisler edited a book entitled <u>Learning by Discovery: A Critical Appraisal</u> which summarized much of the thinking in this area.

Since that time, attention seems to have shifted from locating the quantitative effects (i.e., "how much" is learned) of how material is taught, to an interest in the cognitive question of "what" is learned under different instructional methods (e.g., Roughhead & Scandura, 1968). This has necessitated an analysis of <u>S</u>'s cognitive activity during learning, under the assumption that the outcome of learning--"what" is learned--is the product of both the content of material presented and the particular assimilative set or receptive style used to encode it.

In other words, as the pervasive emphasis on "how much" is learned under different instructional methods gives way to the question of "what kind" of learning occurs, the need for an understanding of the acquisition process and especially, of the role of <u>S</u>'s internal cognitive activities during acquisition, has become apparent. At least two kinds of theories of the acquisition process seem possible. (1) A fairly simple idea, one that follows from "how much" questions, is that apparent differences in what is learned are due to some Ss acquiring more of one kind of content and less of another relative to other Ss. (2) A more complex proposition, one that most theorists interested in "what kind" questions seem to hold, is that different kinds of learning outcomes are due to acquisition processes in which content material is encoded within different assimilative sets by different Ss. Although the first proposal requires only an analysis of the amount of material presented, most recent theories have relied on modified versions of the second proposal in which S's cognitive activity or set as well as the material presented must be analyzed.

For example, Rothkopf (1970) writes of the influence of instructional method on S's "mathemagenic activities" during learning, activities that he relates to such concepts as set, attention, orienting reflex, information processing, cognition, and rehearsal. Rothkopf (p. 325) argues: "The proposition is simple. In most instructional situations, what is learned depends largely on the activities of the student."

Gagné (1965, 1966) has outlined a series of "internal conditions" for problem solving which include "search and selection" of existing knowledge. He has suggested that "what is learned" involves both "external events" such as instructional materials, instructions, and direction, and "internal events" such as the nature of this "trial and error," "hypothesis selection," or "search and selection" activity.

Taking a somewhat different approach, Ausubel (1961, 1968) speaks of S's "learning set" and suggests that learning outcomes are determined both by "content conditions" (i.e., presentation of the to-be-learned material) and "set conditions" (i.e., the existing structures S uses to assimilate the content). Ausubel proposes that S may store content material in either a rote or meaningful way depending on S's ability to relate subject matter content to existing cognitive structures, i.e., depending on whether the "content" is encoded into a "rote learning set" or into a wider "meaningful learning set." The point is summarized as follows (1961, p. 95): "As long as the set and content conditions of meaningful learning are satisfied, the outcome should be meaningful and the advantages of meaningful learning (economy of learning effort, more stable retention and greater transferability) should accrue irrespective of whether the content to be internalized is presented or discovered, verbal or nonverbal."

This general theory of the acquisition process has been summarized by Mayer & Greeno (1972, p. 165): "...different instructional procedures could activate different aspects of existing cognitive structure. And since the outcome of learning is jointly determined by new material and the structure to which it is assimilated, the use of different procedures could lead to the development of markedly different structures during the learning of the same new concept."

For the "how much" theories of the acquisition process, there is little need to analyze the various possible kinds of cognitive sets, since it is the amount and type of information in the teaching material which is the main determinate of learning outcome. However, a crucial question for the "what kind" theories is: How can the internal cognitive set and activities of the learner be characterized? A number of defining factors have been noted with respect to these internal events during learning, and generally the distinctions among various types of internal activity during learning involve: (1) S actively participates in the discovery of the to-be-learned principle or material vs. S passively receives the material in final form, (2) S stores and organizes the material in his own way vs. S stores and organizes the material as E has organized it, (3) S assimilates the material to a wide range of existing cognitive structures vs. S accomodates his existing structures to the material, (4) S acquires a high level, general rule or strategy vs. S strives to acquire discrete, specific responses to specific situations.

Bruner (1961, p. 24) in his classic paper entitled "The Act of Discovery" discusses this distinction, especially the first part of it, between types of internal events during learning:

> Very generally, and at the risk of oversimplification, it is useful to distinguish two kinds of teaching: that which takes place in the expository mode and that which takes place in the hypothetical mode. In the former, the decisions concerning mode and pace and style of exposition are principally determined by the teacher as expositor, the student as listener...in the hypothetical mode the student is in a more cooperative position ...the student is not a bench-bound listener, but is taking part in the formulation...and may even take an 'as if' attitude.

The second aspect of the distinction between internal activities is reflected in Bower's (1970) separation between "experimenter-imposed groupings" (E-codes) and "subject-imposed groupings" (S-codes). The assimilation-accomodation component of the distinction (i.e., the third part) follows with slight modification, from Piaget's (1970) original usages. The final aspect of the distinction deals with S's interpretation of what should be learned as discussed by Rosenthal (1966) and Orne (1962).

One way of summarizing the possible distinctions in internal activity or set, supposedly evoked by different instructional methods, is to differentiate the degrees to which S searchs through existing knowledge in order to map presented material into superordinate organizing systems. Unfortunately, however, although it is the theme of many a theory there are very few experimental studies to deal directly with the relationship among instruction method, internal cognitive activity, and learning outcome.

One set of experiments has been carried out by Scandura (1966, 1967), who has attempted to delineate experimental variables influencing how broadly S encodes mathematical rules or algorithms. For example, in a problem solving task, Ss given the solution algorithm in conjunction with very specific applications performed significantly better on near transfer items than Ss not given the algorithm, but performed significantly worse on far transfer than Ss given the algorithm with more general applications. In another study, Ss learning problem solution rules in symbolic notation could apply them just as well as Ss learning the same rules in plain English, only if they had received pretraining in what the symbols meant.

The evidence supports the claim that although all Ss are taught the same content (i.e., problem solving rules), internal factors such as a broader assimilative "set" do influence transfer performance.

In another study, Gagné and Smith (1962) found that Ss who were forced to verbalize a rationale for each step during learning to solve the "disc problem" performed significantly better on transfer to harder problems of the same type than Ss who did not verbalize during learning. Gagné and Smith (p. 17) state the influence of verbalization on internal cognitive activity as follows: "Requiring verbalization somehow forced the Ss to think."

Such findings seem to support the idea that differences in what is learned are due to differences in S's assimilative set which are activated (at least in part) by instructional method. However, techniques for producing and assessing structural or qualitative differences in what is learned have not been well developed, and there is still relatively little empirical information about the acquisition process.

In the present study a technique was employed that was developed in Greeno's laboratory (Mayer & Greeno, 1972; Egan & Greeno, 1973; Greeno & Mayer, in prep.) which seems to be capable of producing and assessing structural differences in learning outcomes. To produce qualitiative or structural differences in how material is encoded and, hence, in "what is learned"--rather than just an overall quantitative effect--two short, expository teaching booklets were developed which varied the emphasis and sequencing of instruction.

One booklet (Sequence G) attempted to activiate S's general experience with probability situations--such as "trial," "outcome," "success," "outcome sequence," etc.--before presenting the formula. The other booklet (Sequence F) attempted to activate S's experience with arithmetic operations and calculating with a formula by beginning with a statement of the entire formula and explaining each component variable in terms of how it fit into using the formula. The difference between the two booklets was whether the formula and sub-formulas were presented before (Sequence F) or after (Sequence G) discussions of component variables and whether component variables, when discussed, **where** related to S's general experience (Sequence G) or explained only within the context of the formula (Sequence F).

Having attempted to induce differential encoding and hence qualitatively different learning outcomes, a multileveled transfer test consisting of both near and far transfer items was used to assess the structure of learning outcomes. The reason for this was that structurally different learning outcomes could be inferred from a pattern of results in which subjects in one group excell on some kinds of posttest items and subjects in another group excell on other kinds, yielding what has been called a treatment x posttest interaction or TPI.

The focus of the present experiments was on two questions. The first question was to determine how the structural differences in learning outcomes which are observed at the end of learning develop over the course of learning. One fairly straightforward possibility is that Ss in one instructional group simply learn more of one kind of content and Ss in another instructional group learn more of another kind. A contrasting view takes S's internal cognitive set or encoding style into consideration and suggests that different assimilative sets are activated by different instructional methods; hence, content material is encoded differently by Ss in different instructional groups and structurally different learning outcomes result.

The second problem was to determine whether the cognitive structure, once established, could be affected by testing conditions such as stress or memory support. Such manipulations represent an attempt to force Ss who had learned by different instructional methods to process existing structures in the same way. If structural differences "disappear" under these circumstances, the importance of the original teaching method set in establishing the structure of learning outcomes would be diminished.

In order to provide information on how structural differences develop over the course of learning, the present study included between-subject tests for TPI at three points during learning--after presenting approximately one-third (introduction and combinations lesson), two-thirds (introduction, combinations and joint probability lessons), or all of the material (introduction, combinations, joint probability and binomial lessons). Attention was to be paid to results suggesting the nature of the acquisition process which produces structurally different learning outcomes, especially results distinguishing (1) a process of adding more and more content material of different kinds under the two instructional methods, from (2) a process of embedding material in different assimilative sets under the two treatments.

In order to provide information on how structural differences persist once they have been established, the present study varied the amount and kind of stress during testing. Attention was paid to results suggesting whether subjects who have structured a problem-solving rule in different ways can be forced to process this knowledge in the same way. This question was tested by comparing the pattern of TPI produced by Ss from both instructional groups who are tested under stress with Ss tested without stress.

CHAPTER II

EXPERIMENT I

This study attempted to assess the consequences of varying the sequencing and amount of instruction for a mathematical concept, and the amount of stress during testing. In each instructional sequence, Ss were given varying amounts of instruction. All Ss took the same 30-item posttest, containing several kinds of test items, and the condition of testing was varied.

Method

Subjects and Design

Subjects were 117 University of Michigan students who had volunteered to participate in psychological experiments at the Human Performance Center for pay. Nine Ss served in each cell of a 2 x 3 x 2 factorial design, with a thirteenth group of Ss serving as controls.

The first factor that was varied was the emphasis and sequencing of instruction. Some Ss (Sequence F) learned the binomial distribution concept by a method that began with a statement of the formula and explained the variables of the formula in relation to the mechanical operations involved. Other Ss (Sequence G) received a teaching booklet that related the variables of the formula to concepts that were presumably part of S's general knowledge and then presented the formula.

The second factor that was varied was the amount of instruction. Each teaching booklet was divided into four parts consisting of an introduction (Part I), a lesson on combinations (Part II), a lesson on joint probability (Part III), and a lesson on binomial probability (Part IV). Some Ss (Amount I) received just Part I and Part II; some Ss (Amount II) received Parts I, II, and III; and some Ss (Amount III) received all four parts. In Experiment I, all Ss received the booklets in the order described above, namely introduction followed by combinations followed by joint probability followed by binomial probability (Ordering ABC). (In Experiment II, all Ss received the booklets with the positions of the combinations and the joint probability parts reversed, Ordering BAC.)

The third factor that was varied in Experiment I was the presence or absence of memory support, a sort of stress relief, during the transfer test. Some Ss (Condition Memory Support) were allowed to refer to their teaching booklets during the test, while other Ss (Condition No Memory Support) were not. In Experiment I, all Ss were allowed as much time as necessary to complete the test (Condition No Time Stress).

Materials

The two instructional sequences (Sequence F and Sequence G) were incorporated into two four-part typewritten booklets with six to eight pages in each booklet. The lessons were converted, with slight revision, from the CAI frames used by Greeno and Mayer (in prep.). Both groups learned the concept expressed by the formula,

$$P(R,N) = C(N,R) \times P^{R} \times (1-P)^{N-R}$$

where N is the number of trials, R is the number of successes, P is the probability of success, C(N,R) is the number of combinations, and P(R,N) is the probability of R successes in N trials.

The ideas expressed in the booklet for Sequence G were:

Part I (Introduction): Introduction and example of "trial" and "outcome," introduction and example of "success," introduction and example of "probability of success," introduction and example of "failure" and "probability of failure," introduction and example of outcome "sequence" and "probability of a sequence," formal notation for N and R.

Part II (Combinations): Re-introduction of "number of trials," "number of successes," and outcome "sequence;" example of how to generate number of combinations with N = 3 and R = 2; verbal definition of "number of combinations;" formal definition of C(N,R); two examples relating verbal explanation of how to generate combinations to the operations involved in using the formula for N = 4, R = 1 and N = 5, R = 3.

Part III (Joint Probability): Re-introduction of "trial," "outcome," "sequence," and "probability of success;" example of how to generate joint probability for a 4 term sequence; two examples of how to generate joint probability for 5 and 6 term sequences and their relation to the formal definition; introduction of notation for P and 1-P; another 5 term example relating generated joint probability to formal definition; statement of formal definition of $P^R \times (1-P)^{N-R}$.

Part IV (Binomial Probability): Introduction and verbal definition of bionomial probability; re-introduction of "probability of a sequence" and "number of combinations;" formal definition of $C(N,R) \times P^R \times (1-P)^{N-R}$; example of how to use formula, $P(R,N) = C(N,R) \times P^R \times (1-P)^{N-R}$.

The ideas expressed in the booklet for Sequence F were:

Part I (Introduction): Introduction of formal definition for binomial probability as $P(R,N) = C(N,R) \times P^R \times (1-P)^{N-R}$; explanation and example of formula as finding the probability of R successes in N trials; restatement of formula and explanation of $P(R,N)$ as product of $C(N,R)$, P^R, and $(1-P)^{N-R}$; verbal definition of $C(N,R)$, R, P, $(1-P)$, and $(N-R)$.

Part II (Combinations): Restatement of formula; re-introduction of $C(N,R)$ as first major step in calculating; formal definition of $C(N,R)$ as $N!/(R! \times (N-R)!)$; enumeration of steps to calculate $C(N,R)$; two examples of steps to calculate for $N = 4$, $R = 1$ and $N= 5$, $R = 3$.

Part IIIa (Joint Probability - P^R Term): Re-introduction of P^R as second major step in calculating with the formula; verbal definition and example of P; enumeration of steps to calculate P^R; two examples of steps to calculate for specific values.

Part IIIb (Joint Probability - $(1-P)^{N-R}$ Term): Restatement of formula; re-introduction of $(1-P)^{N-R}$ as the third major step; enumeration of steps to calculate $(1-P)^{N-R}$ and example; verbal statement of P as probability of success, $(1-P)$ as probability of failure, R as number of successes, with N-R as number of failures; two examples of steps to calculate $(1-P)^{N-R}$ for specific values.

Part IV (Binomial Probability): Restatement of formula; restatement of steps to calculate P(R,N) as calculating C(N,R), calculating P^R, calculating $(1-P)^{N-R}$, and finding the product of these three values; two examples for specific values.

All examples were presented either in terms of N, R and P or a die rolling situation, and an attempt was made to present both groups with the same examples. The two booklets are reproduced in Appendix A.

A test set consisted of 30 typewritten cards representing five problem types, two problem formats, and three problem content areas. The 5 x 2 x 3 design yielded 30 cells, each represented in one of the 30 test items.

The first transfer dimension was the type of test item. Familiar problems (F-type) were presented in the same way as example problems during training; transformed problems (T-type) required a transformation, usually of an algebraic nature, to be put into familiar form; so-called Luchins problems (L-type) presented a complicated looking situation which could be solved quite easily if S would take a moment to "think;" question items (Q-type) asked a question about the variables in the formula rather than requiring a solution value; and unanswerable problems (U-type), although looking very much like Type F problems, presented either insufficient or inconsistent information.

The second dimension was format of test problems. Formula problems (F-format) were stated in terms of N, R and P--the formal notation used in presenting the formulae in the teaching booklets; story problems (S-format) were stated in terms of some situation not discussed in the teaching booklets such as sampling peanuts from a barrel in which some proportion is rotten.

The third dimension was the content of test problems. Combinations problems (C-content) asked for or dealt with C(N,R) or the number of combinations; joint probability problems (J-content) asked for or dealt with $P^R \times (1-P)^{N-R}$ or the probability of a specific sequence; binomial probability (B-content) asked for or dealt with the theme of the teaching booklets, finding P(R,N) or the probability of R successes in N trials.

Examples of the test problems and answers are given below:

Familiar Type, Formula Format, Binomial Content: N = 4, R = 3, P = .20. What is P(R,N)? The correct answer requires plugging the values of N, R and P into the formula to get,

$$P(R,N) = C(4,3) \times (.20)^3 \times (.80)^1 = 16/625.$$

Transformed Type, Formula Format, Joint Content: P = 3(1-P), N = 6, R = N-R. What is $P^R \times (1-P)^{N-R}$? The correct answer requires solving for P and R before plugging into the joint probability formula to get,

$$P^R \times (1-P)^{N-R} = (3/4)^3 \times (1/4)^3 = 27/4096.$$

Luchins Type, Story Format, Combinations Content: There are 10 different sequences that have exactly two successes. All the sequences have the same length. How long are they? The correct answer requires finding a value of N to fit C(N,R) = 10 and R = 2, as shown:

$$\frac{N!}{(N-2)!2!} = 10, \; N = 5$$

Question Type, Story Format, Binomial Content: Is there a difference between the probability that two dice rolled at once both come up 6 and the probability that one die rolled twice comes up 6 both times? The answer requires an understanding of independence of events; hence the subject should answer "no" or "no difference."

Unanswerable Type, Story Format, Combinations Content: How many different sequences have the same number of successes as failures? The answer requires the recognition of insufficient information, i.e., no value of N is given, and hence the correct answer is "no answer."

The entire test set is reproduced in Appendix B.

Additional materials used in the experiment were a subject record consisting of questions to determine the extent of S's experience in statistics and probability, and a pretest designed to determine whether S had sufficient computational skill to master the material in the teaching booklets. The pretest consisted of 10 items: 2 dealt with factorials (e.g., 5! = 5 x 4 x 3 x 2 x 1 = ____); 4 dealt with exponentiation of fractions or decimals (e.g., $(1/4)^3$ = ____); and 4 dealt with multiplication of fractions or decimals (e.g., 2/3 x 1/4 = ____).

Procedure

Subjects were run in small groups averaging four per session. First, the subject record and pretest were administered. Subjects indicating no relevant experience with the binomial but making no more than 3 computational errors were randomly assigned to treatment groups.

The E read the instructions and again asked S to indicate familiarity with the binomial. Then, each S was given the appropriate teaching booklet.

Subjects were instructed to read their booklets silently and at their own rate and to try to understand the explanations and examples. Subjects were allowed to take notes or figure on a blank sheet of paper. Subjects were told that they would have 20 minutes, and all Ss finished within that limit. Subjects who finished earlier were asked to sit quietly until the test began.

Immediately following the reading period, E collected the teaching booklets and notes (except for Condition Memory Support Ss), read the instructions for the test, and wrote the formula for P(R,N) on the chalkboard. Each S was given a pile of 30 problem cards face down, and a blank answer sheet. On E's signal S was to turn up the first card, copy its code number, show his work and circle his final answer, and then go on to the next card. There was no time limit but all Ss were told that once they began a new card they could not go back to work on any previous card. All Ss were told to write "no answer" if they felt a problem was unanswerable.

Nine different orders of presenting the test cards were constructed. The orders were random except for the constraint that, of the nine orderings, each item had to appear in the first one-third (i.e., 1 through 10) in three of the sets, in the second one-third (i.e., 11 through 20) in another 3 of the sets, and in the last one-third (i.e., 21 through 30) in the other 3 sets.

After the test, Ss were paid and informally questioned by E. No Ss who participated in the experiment indicated previous familiarity with the material at this time.

Results and Discussion

In Experiment I, five Ss failed to receive a score of seven or more on the pretest, four Ss expressed direct familiarity with the material in the teaching booklets, and two failed to properly follow directions. These Ss were eliminated from the experiment and new Ss were run in their places.

The performance on the transfer test was scored with each item marked either correct or incorrect. Answers were marked correct if they were in proper form even though computation may have been incorrect or not carried out. The control Ss performed at very low levels, i.e., an average of less than 10% correct, thus indicating that the experimental treatments had a substantial effect. An analysis of variance was performed on the data of all experimental Ss. There was only a marginal difference in the overall performance of the two experimental groups ($F = 2.92$, $df = 1/96$, $p < .100$), thus frustrating the question of "which method is best."

Development of Structural Differences

In earlier studies, where Ss were tested after receiving the entire booklet, there was evidence for inferring that Ss in different instructional groups had acquired learning outcomes which differed in a structural or qualitiative way (Mayer & Greeno, 1972; Egan &

Greeno, 1973; Greeno & Mayer, in prep.). Structurally different learning outcomes could be inferred, it was argued, if Ss in one treatment condition excelled on certain kinds of posttest (i.e., transfer test) problems and subjects in another treatment condition excelled on other kinds of posttest problems. Thus the inference of structural differences between learning outcomes is based on what has been called a treatment x posttest interaction or TPI--a disordinal interaction (Bracht, 1970) between treatment condition (e.g., sequencing of instruction) and kind of posttest item (e.g., format, type or content).

In the present experiment, a comparison of Ss in the two instructional sequence groups who received Amount III reveals the same general pattern of TPI as produced in previous studies. The right panels of Figures 1, 2, and 3 show that Ss in Sequence F who received the entire booklet excelled on near transfer problems such as F-format (formula), F-type (familiar), and B-content (binomial), while Ss in Sequence G excelled on far transfer problems such as F-format (story), Q- and U-type (question, unanswerable), and C- and J-content (combination, joint probability). These TPIs indicate that differences in the structure of learning outcomes have been produced.

A major new question asked in the present experiment is: How do the structural differences in learning that are observed at the end of the lesson develop over the course of learning? This "acquisition question" has at least two possible answers of theoretical interest. One possibility is that over the course of learning Ss in one group are acquiring more of one kind of content while Ss in another group are acquiring more of another kind of content. This process should manifest itself as an even increase in overall performance by both groups accompanied by a gradual emergence of TPI, being very weak early in learning where little content has been presented and strong at the end where the maximum differences in content are possible. A second possibility is that the general organization and emphasis of each testing sequence encourages the activation of different kinds of assimilative sets very early in, and across all, learning. This process should manifest itself either as a difference in overall performance early in learning that disappears by the end or as a fairly strong and consistent TPI at all points in learning, at the beginning as well as the end of the lesson.

In order to provide some data on this question, posttests were administered at three distinct point in learning: After Amount I for some Ss, after Amount II for other Ss, and after Amount III for other Ss. As could be expected, a significant overall effect due to amount of instruction was found (F = 4.49, df = 2/96, p < .025), suggesting that adding more sections increased performance for both instructional sequences.

Before presenting the various figures concerning treatment x posttest interaction, it should be noted that the x-axis of each graph is a continuum, representing the degree of transfer, only in a very general sense. Although an attempt has been made to place near transfer tasks on one side (e.g., familiar type on the left) and interpretive, far transfer tasks on the other (e.g., unanswerable and question types on the right), in some cases the ordering is arbitrary (e.g., the respective placements of Q-type and U-type, or of C-content and J-content). Thus the lines connecting the dots are included to aid in presenting the data and in making comparisons across panels, but should not be interpreted as suggesting a strict continuum for degree of transfer that would allow interpolation between points or judgments based on the absolute shape of the lines.

Figure 1 shows the instructional sequence by problem format interaction after each amount of instruction. The two-way interaction between sequencing and format is reliable ($F = 14.72$, df = 1/96, $p < .001$) indicating structural differences in what was learned. As can be seen, this TPI is clearly evidenced in all three panels of Figure 1 and there is no reliable sequence by amount by format interaction ($F = 1.00$, df = 2/96, $p > .20$) required to reject the hypothesis that TPI is the same at each point in learning. Although the performance of each sequence group reliably increased overall from Amount I to Amount II to Amount III, the absence of an amount by format interaction suggests a proportional quantitative increase for both problem formats rather than a structural change as the amount of instruction was increased.

Figure 2 shows the instructional sequence by problem type interaction after each amount of instruction. The two-way interaction between sequencing and type is reliable ($F = 6.28$, df = 4/384, $p < .001$), again suggesting structural differences similar to those found in earlier studies. As can be seen, even after Amount I there is evidence of this characteristic TPI with Sequence F slightly ahead for F-type and Sequence G superior on Q- and U-type. This pattern seems fairly consistent across all three amounts of instruction and there is no reliable treatment

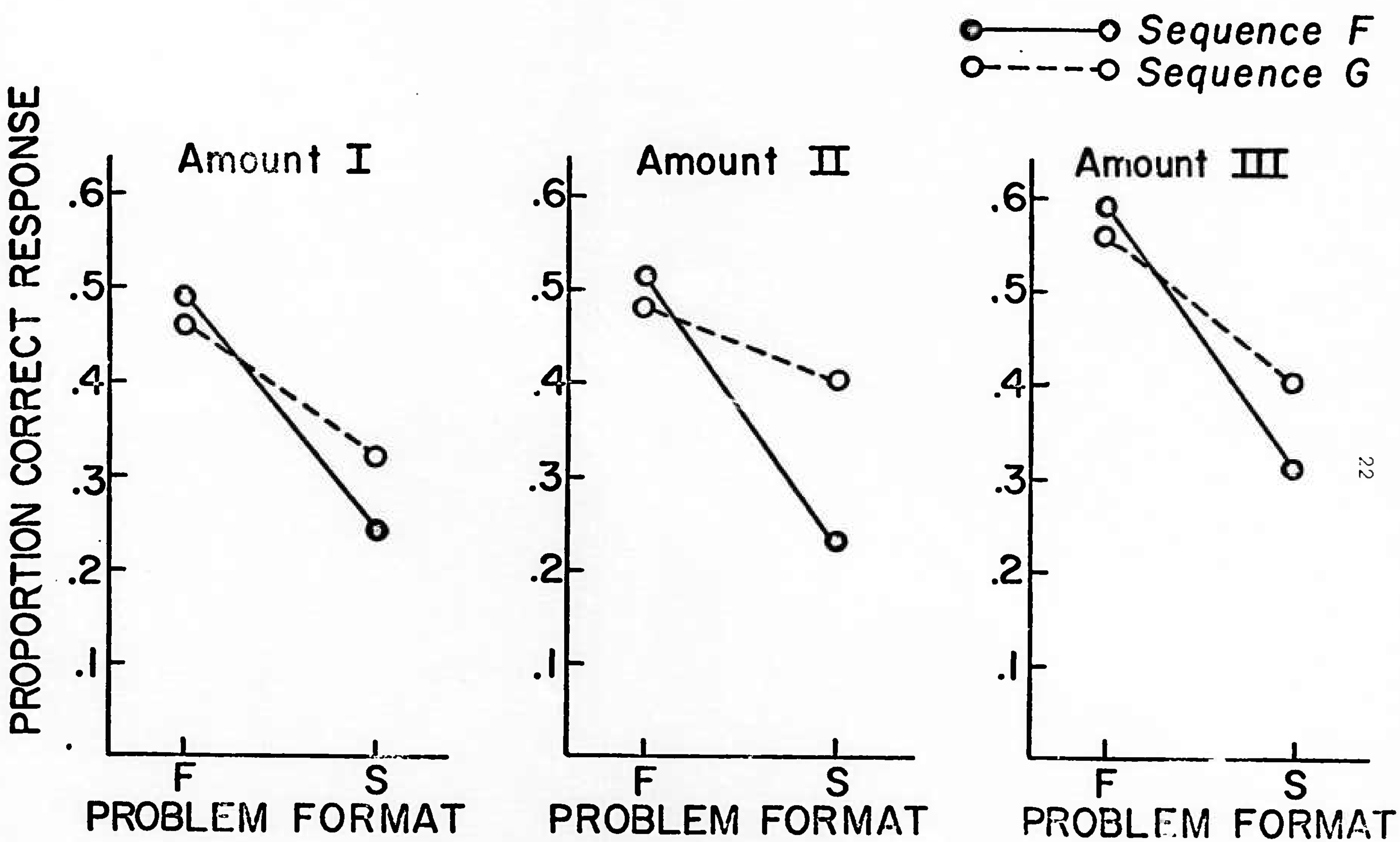

Fig. 1. Proportion correct response for two instructional groups by problem format for three amounts of instruction--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: F-format (formula) represents near transfer and S-formula (story) represents far transfer.

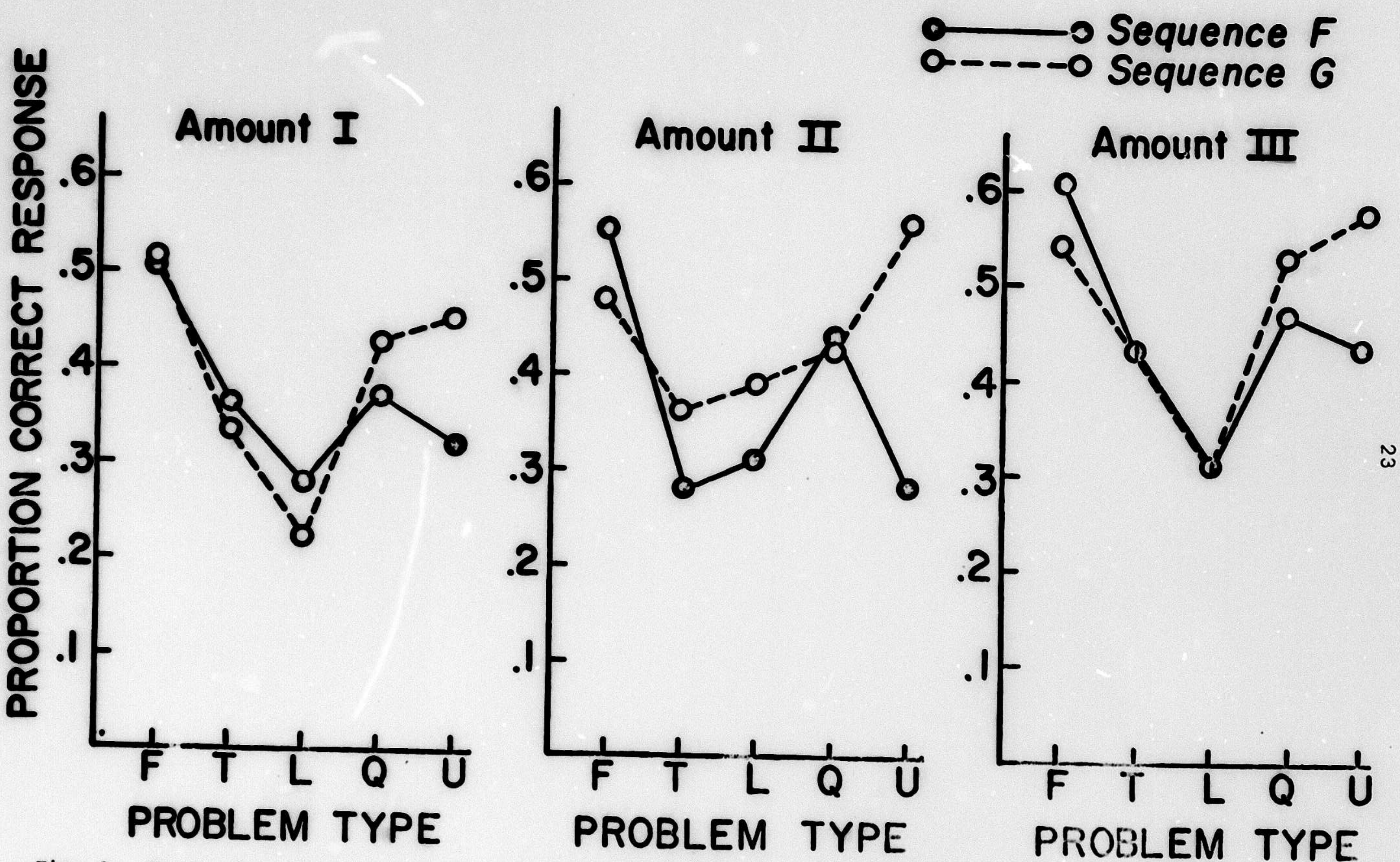

Fig. 2. Proportion correct response for two instructional groups by problem type for three amounts of instruction--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: F-type (familiar) and T-type (transformed) represent near transfer tasks, while L-type (Luchins) represents intermediate transfer, and U-type (unanswerable) and Q-type (question) represent far transfer.

by type by amount interaction ($F = 1.03$, df = 8/384, $p > .20$) required to reject this hypothesis. Although the performance of each treatment group reliably increased overall from Amount I to Amount II to Amount III there is no reliable amount by type interaction ($F = 1.05$, df = 8/384, $p > .20$) again suggesting that the increase is proportional across all problem types. These findings are consistent with the notion that the structure of learning outcomes remains constant across all three amounts of instruction.

Figure 3 shows the sequence by problem content interaction after each amount of instruction. The two-way interaction between sequence and content is significant ($F = 10.19$, df = 2/192, $p < .001$) again indicating a difference in the structure of what was learned. As can be seen, this TPI is hinted after Amount I, but does not become strong until after Amount II and III. However, there is no reliable treatment by content by amount interaction ($F = 1.77$, df = 4/192, $p > .15$) thus it is not possible to reject the notion that the structural differences among the instructional groups remain constant at each of the three points in learning that were tested. A reliable amount by problem content interaction ($F = 2.54$, df = 4/192, $p \approx .050$) suggests that, not only did performance increase overall with the addition of more sections, but it also increased disproportionately more--as might be expected--on the content of the material covered.

In summary, the three possible sequence by posttest by amount interactions represent three separate tests of the "acquisition question." In all three cases there is reliable two-way TPI, but there is no evidence of any difference in TPI among Amount I, Amount II, and Amount III.

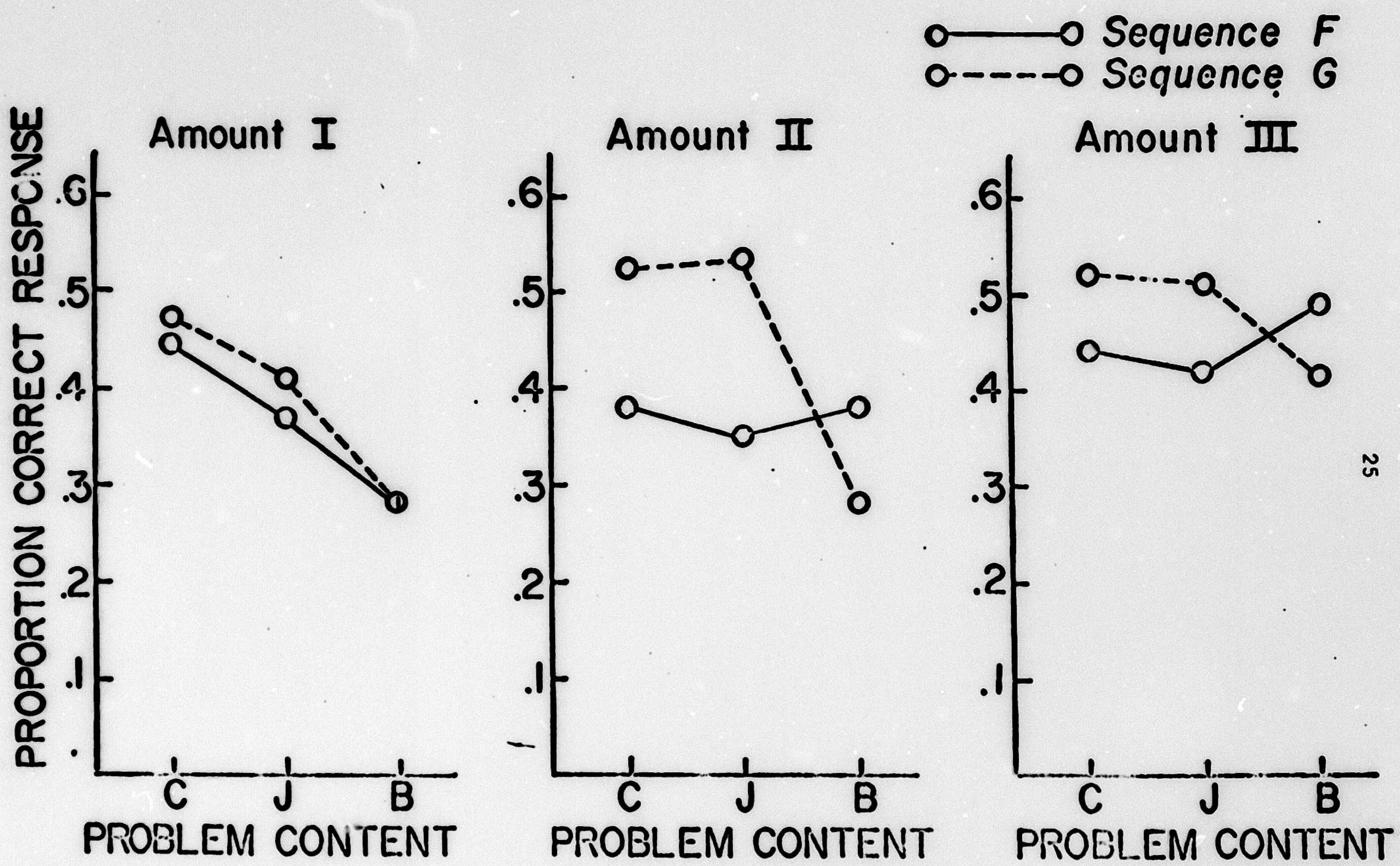

Fig. 3. Proportion correct response for two instructional groups by problem content for three amounts of instruction--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: B-content (binomial) represents near transfer, and C-content (combination) and J-content (joint) represent far transfer.

These results suggest that structural differences occur early in learning and remain fairly constant throughout.

In previous attempts to describe the structural differences in what is learned by subjects in different instructional groups it has been useful to postulate two dimensions of cognitive structuring--internal connectedness and external connectedness. Internal connectedness refers to the degree to which the variables of the formula are related to one another in S's cognitive structure, e.g., P is related to R by exponentiation. External connectedness refers to the degree to which variables of the formula are related to knowledge already existing in S's cognitive structuring, e.g., P is related to past experience with probability such as weather forecasts (20% chance of rain), dice (each number has 1/6 probability), batting average (.333 means the batter gets a hit 1/3 of his times at bat).

In the present situation, there is evidence that Sequence F Ss developed structures with strong internal connectedness but weak external connections while Sequence G Ss developed structures with strong external connectedness but weak internal connections. This hypothesis is consistent with observed TPI (with each of the three posttest dimensions) which generally showed that Sequence F Ss excel on near transfer items which require an exact application of the formula, and that Sequence G Ss excel on far transfer items which require a more sophisticated understanding and interpretation of the component variables.

The present study adds some information on how these different structures develop. The results of the present study seem consistent with the notion that different receptive sets are activated early in learning. Furthermore, the differences in how the material is encoded are established early and remain constant throughout learning. Greeno (1972) provides a framework for analyzing problem solving behavior which relies on a distinction between two kinds of knowledge stored in semantic memory--"propositional knowledge" and "algorithmic knowledge." Propositional knowledge refers to relational and conceptual information such as hierarchies of classes and subsets (collies are dogs), properties of classes of things (dogs have tails), and facts (April 17 - 23 is national Dog Week). Algorithmic knowledge refers to rules or operations such as the procedures followed in doing "long division."

In the present example, it seems that the organization and emphasis of Sequence F encourages the use of a narrow, assimilative set concerned with mathematical operations and calculations--what could be called algorithmic knowledge--and that Sequence G encourages the activation of a broader, more integrative set made up of S's general experience--what could be called propositional knowledge. Since what is learned is the product of both the presented material and the assimilative set S uses to encode it, different learning outcomes are possible. What these results seem to show is that the algorithmic kind of set is activated quite early in learning for Sequence F and the propositional kind of set is activated quite early in learning for Sequence G; further, there is support for the idea that the structural differences observed at the end of learning have their roots in this embedding of material into two different kinds of assimilative sets or knowledge which begins early and continues throughout learning.

Resilience of Structural Differences

A second new question asked by the present experiment was: Once structural differences have been established by different instructional methods, can TPI be disrupted by testing conditions that vary the amount of stress during testing? This "resilience question" depends on an attempt to create testing conditions that could force Ss in different instructional groups (and hence with different learning outcomes) to process their cognitive structures in the same way and thus eliminate TPI.

This question is of interest because it gives some information about the permanence of effects generated by teaching method. If structural differences appear during and immediately after learning, as has been shown, but "disappear" or are strongly influenced by testing condition, then the long range importance of what method is used can be minimized. In this case, instructors could rely on post-learning manipulations to alter the structure of learning outcomes. However, if varying testing conditions fail to alter the characteristic pattern of TPI, this would provide more evidence for the importance of selecting an appropriate teaching method. Such results would suggest that postlearning manipulations to process learned material in a certain way have negligible effect on the structure of outcome performance.

In Experiment I, the testing variable of "memory support"--a sort of stress relief--was used. Subjects in the memory support condition had their teaching booklets available for reference and the binomial formula on the blackboard during the test while Ss in the no memory support condition had only the formula on the blackboard.

The performance for the two sequence groups under memory support and no memory support conditions is given by problem format in Figure 4, by problem type in Figure 5, and by problem content in Figure 6. As can be seen by comparing the right and left panels of each figure, "open book" testing seems to have helped to slightly increase posttest scores overall for both instructional groups, although the increase failed to reach a statistically significant level ($F = 1.44$, df = 1/96, $p > .20$).

For both instructional groups there was some hint that memory support helped to disproportionately increase performance on harder problems--"far" transfer items--especially for S-format relative to F-format items. Marginally significant type by testing condition ($F = 2.12$, df = 4/334, $p < .100$), format by testing condition ($F = 4.79$, df = 1/96, $p < .050$), and content by testing condition ($F = 2.57$, df = 2/192, $p < .100$) interactions offer only weak support for this observation.

Comparing the right and left panels of each figure also indicates that TPI was present under both memory support and no memory support conditions. A failure to find any reliable three way interaction among instructional sequence, type of posttest item, and testing condition ($F < 1.00$, df = 4/84, $p > .20$) among sequence, problem format and testing condition ($F = 2.79$, df = 1/96, $p \approx .100$) nor among sequence, content and testing condition ($F < 1.00$, df = 2/192, $p > .20$) indicates that the attempt to alter the structural differences between the instructional sequences was not successful.

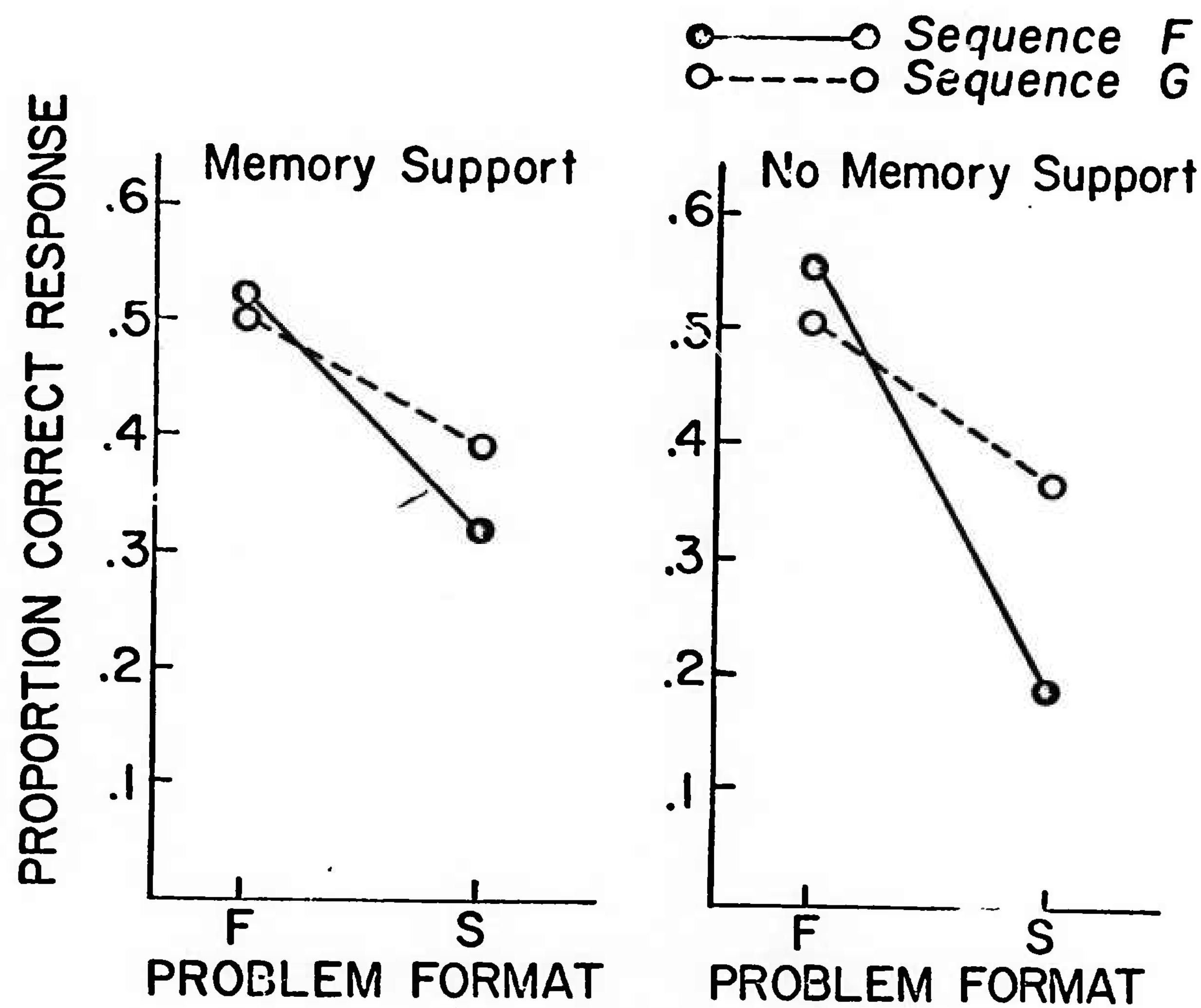

Fig. 4. Proportion correct response for two instructional groups by problem format for two amounts of stress during testing--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: F-format (formula) represents near transfer and S-formula (story) represents far transfer.

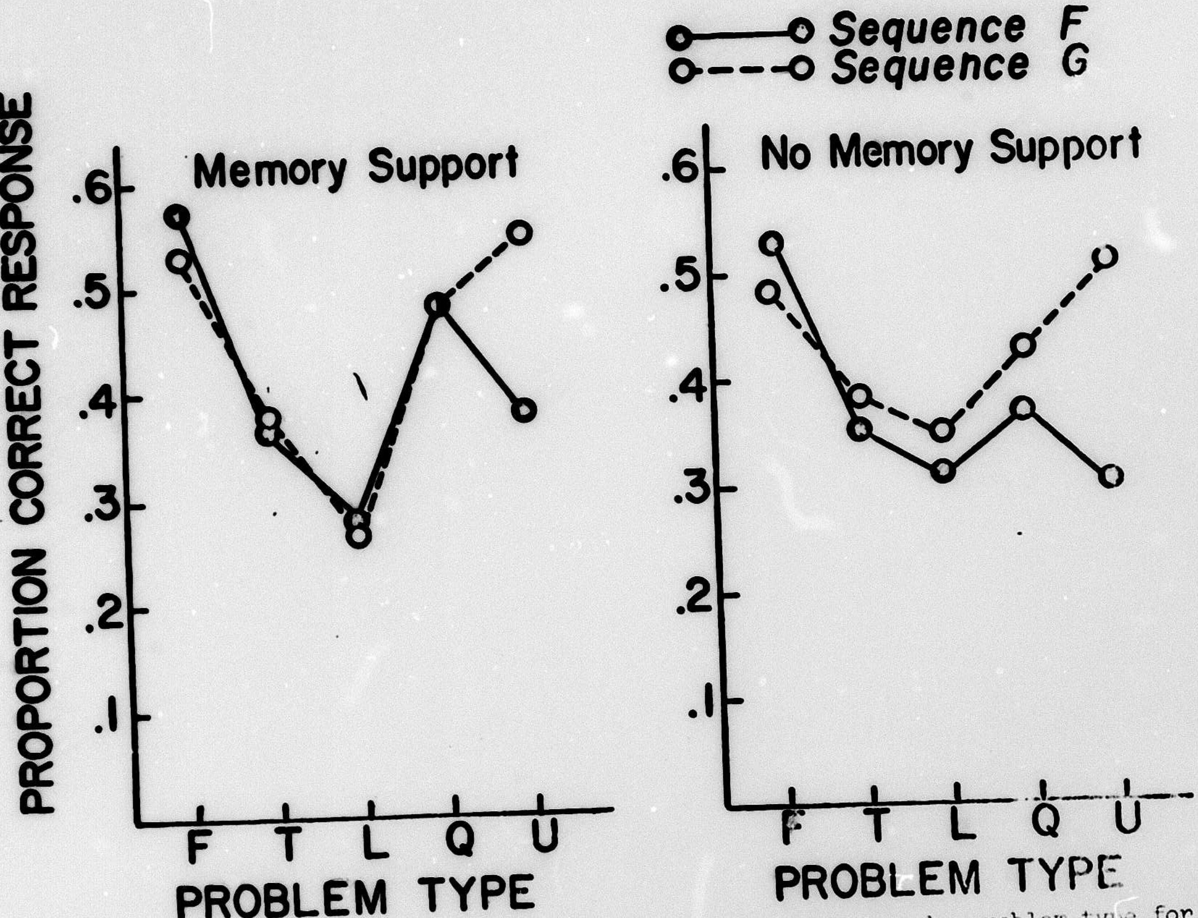

Fig. 5. Proportion correct response for two instructional groups by problem type for two amounts of stress during testing--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: F-type (familar) and T-type (transformed) represents near transfer tasks, while L-type (Luchins) represents intermediate transfer, and U-type (unanswerable) and Q-type (question) represent far transfer.

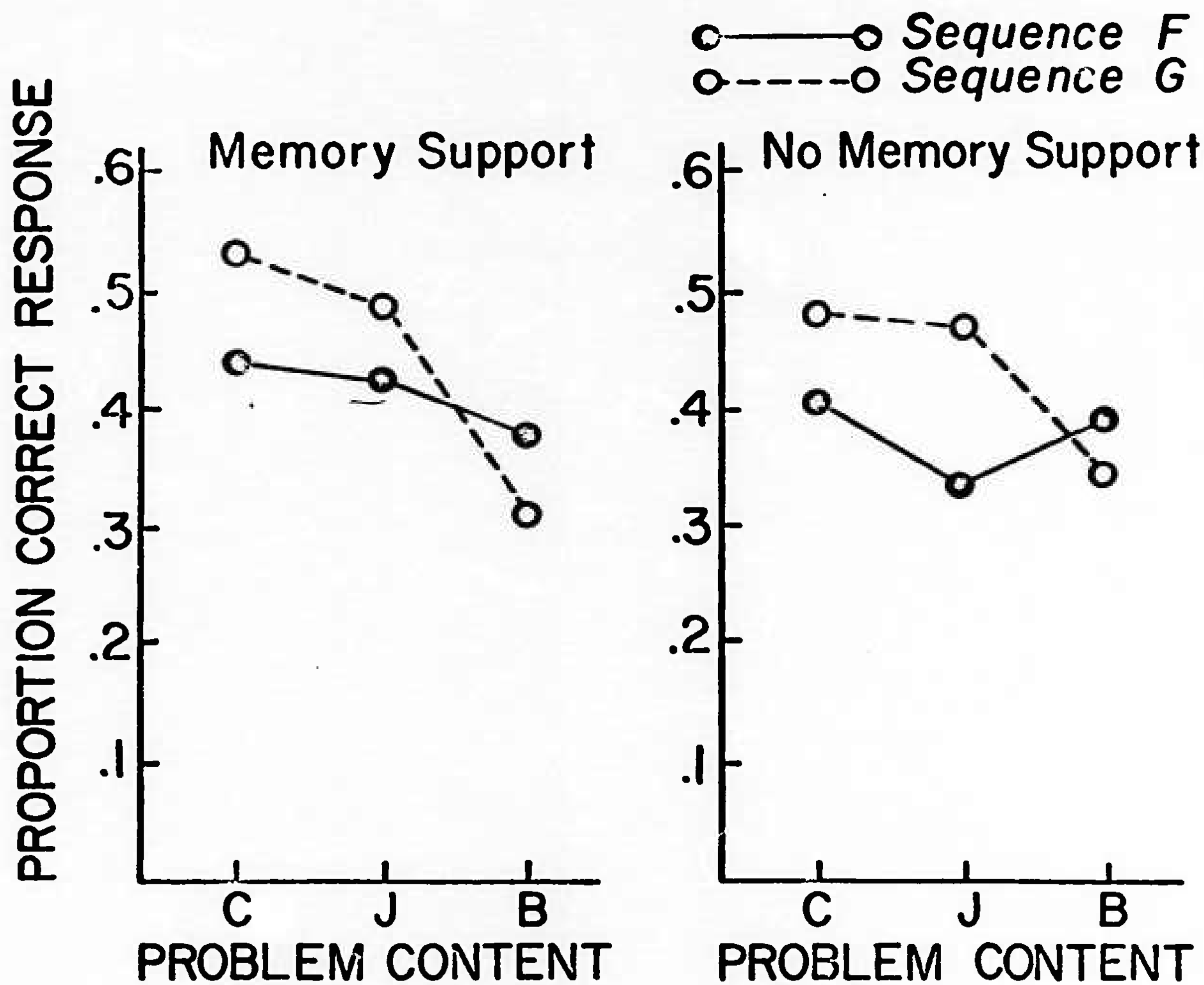

Fig. 6. Proportion correct response for two instructional groups by problem content for two amounts of stress during testing--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: B-content (binomial) represents near transfer, and C-content (combinations) and J-content (joint) represent far transfer.

Open-book testing seems to have had very little effect on either the amount or the structure of problem solving performance. The fact that the test followed learning almost immediately, and that the formula was on the blackboard for all Ss may have reduced the impact of open-book testing. Apparently, Ss in the no memory support condition were able to generate problem solutions quite well with only their existing knowledge and the formula on the blackboard. In any case, the present results do indicate that structural differences are resilient at least within this small range of stress manipulation, and that the attempt to force similiar kinds of memory processing failed to alter evidence for pre-established differences in the structure of learning outcomes.

CHAPTER III

EXPERIMENT II

Experiment II was intended to replicate and extend the results of Experiment I.

Method

Experiment II was the same as Experiment I except that all Ss received the BAC ordering of booklet parts (instead of the ABC ordering) and, during the test, half of the Ss were required to go on to the next card every 90 sec (Time Stress Condition) and half were allowed to go at their own rate as in Experiment I (No Time Stress Condition). In Experiment II, no Ss received memory support during testing.

As in Experiment I, nine Ss served in each cell of a 2 x 3 x 2 design with the factors being instructional sequence (Sequence F and Sequence G), amount of instruction (Amount I, II and III), and testing condition (time stress and no time stress). The 108 Ss were recruited from a pool of University of Michigan students who had volunteered to participate in psychological experiments at the Human Performance Center for pay.

Because of the change in ordering, Ss receiving Amount I studied the introduction section and joint probability section; Ss given Amount II had the introduction followed by joint probability and then, combinations, and Ss given Amount III received all four sections in the order: introduction, joint probability, combinations, binomial probability.

Results and Discussion

In Experiment II, six Ss failed to receive a score of seven or more on the pretest, three Ss indicated direct familiarity with the material in the teaching booklets, and four failed to follow directions. The data for these Ss were eliminated from the experiment and new Ss were run in their places. The transfer test performance was scored and analyzed as in Experiment I. As in Experiment I, there was no reliable difference in overall performance between the two treatment groups ($F = 1.80$, $df = 1/96$, $p > .15$).

The Effect of Ordering

A variable that was changed from Experiment I to Experiment II was the ordering of the parts or sections in the instructional booklets. For example, Amount I contained information about an introduction and combinations in the ABC ordering of Experiment I but Amount I contained information about an introduction and joint probability in the BAC ordering of Experiment II. Some indication of the effect of ordering and the respective importance of the combinations and joint probability sections is provided by comparing performance across experiments, although an interpretation of these data is made more difficult by the fact that the two experiments were not run concurrently nor under the same testing conditions.

The overall performance with Amount I, Amount II, and Amount III for Ss in Experiment I and Experiment II is given in Table 1. In both experiments a significant effect due to amount of instruction was

TABLE 1

PROPORTION CORRECT RESPONSE BY AMOUNT OF INSTRUCTION FOR EXPERIMENTS I AND II

Amount of Instruction	Experiment I	Experiment II
Control	.10	-
Amount I	.38	.25
Amount II	.40	.34
Amount III	.46	.43

obtained (for Exp. I: $\underline{F} = 4.47$, df = 2/96, $\underline{p} < .025$; for Exp. II: $\underline{F} = 22.10$, df = 2/96, $\underline{p} < .001$). As can be seen, however, a major difference is that Ss in Experiment I showed relatively good performance after Amount I and very little improvement from Amount I to Amount II, while Ss in Experiment II showed relatively poor performance after Amount I and very much improvement from Amount I to Amount II.

Apparently, the combinations section (introduced in Amount I of Experiment I and Amount II of Experiment II) was much more important in increasing performance than the joint probability section (introduced in Amount II of Experiment I and in Amount I of Experiment II). Perhaps the concept of combinations is less intuitive or less familiar than the concept of joint probability; but whatever the reason, the present study indicates that in teaching the concept of binomial probability the most important component to teach is the concept of combinations. The ordering of presentation, however, apparently had little or no effect on final outcome.

Development of Structural Differences

As in Experiment I and as can be seen in the right panels of Figures 7, 8, and 9, the characteristic patterns of TPI are present, suggesting that the final learning outcomes of Ss in Sequence F and Sequence G are structurally different. Some information about how these structural differences develop over the course of learning is also provided by Experiment II.

Again, special attention should be paid to whether TPI is observed at all points in learning (as found in Experiment I) and whether both treatment groups perform at equal levels at each point in learning (as found in Experiment I). The even growth of learning outcomes for the two instructional groups, accompanied by a gradual emergence of TPI would provide support for the notion that Ss in different groups are simply adding more and more of different kinds of information. Either early differences in the apparent amount of learning displayed by the two groups that disappear by the end of learning, or a constantly observable TPI, or a combination of the two would suggest that Ss are assimilating the material to different cognitive sets throughout learning.

Figure 7 shows the performance of the two instructional groups by problem format et each of three points in learning. The two-way interaction between sequence and format--in which Sequence F Ss excel on F-format and Sequence G Ss perform better on S-format items--is reliable ($F = 21.65$, $df = 1/96$, $p < .001$) again suggesting a structural difference in what is learned by Ss in the two treatment groups. The sequence by amount

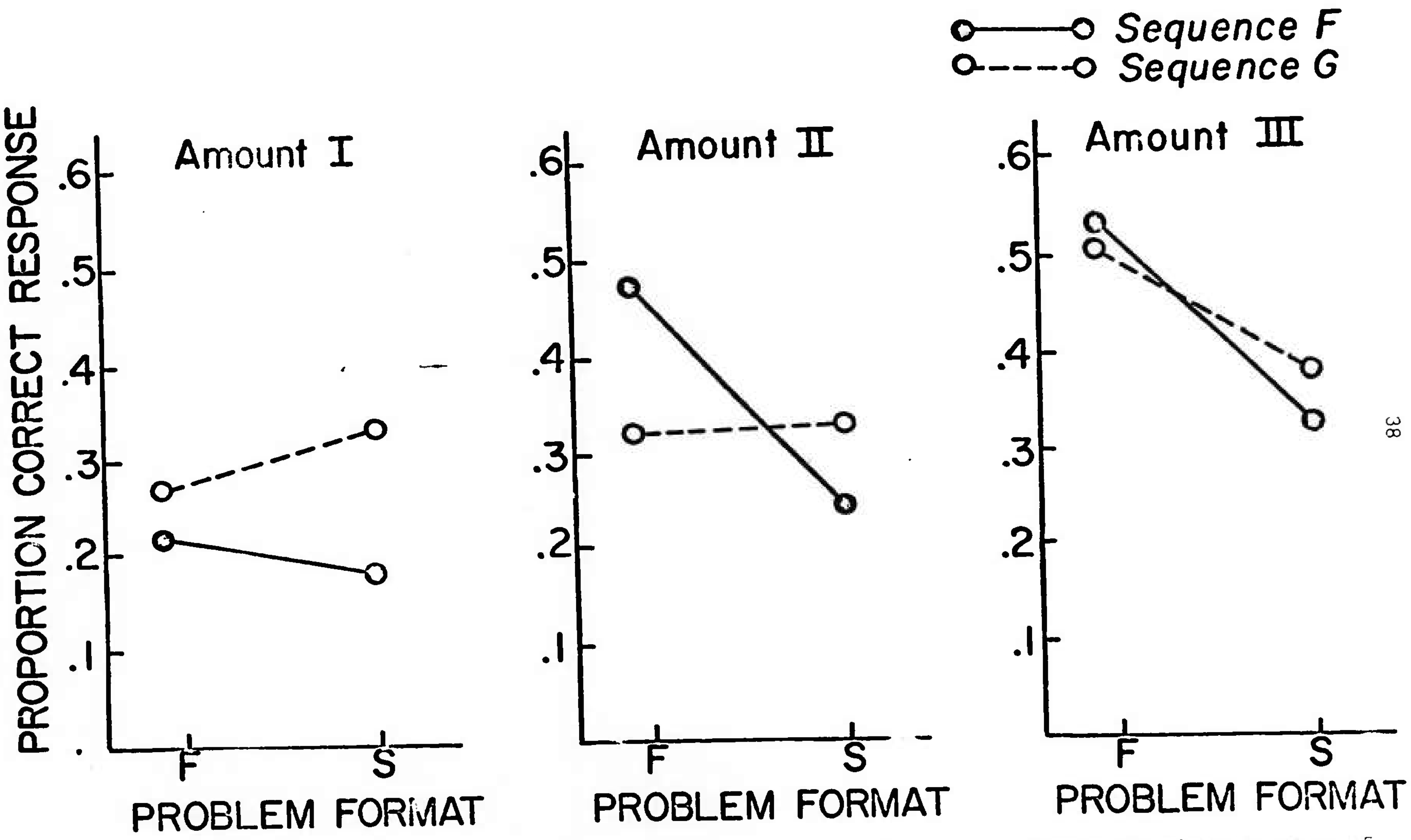

Fig. 7. Proportion correct response for two instructional groups by problem format for three amounts of instruction--Experiment I. The abscissa roughly represents the degree of transfer constituted as follows: F-format (formula) represents near transfer and S-format (story) represents far transfer.

interaction is manifested in the fact that Sequence G outperforms Sequence F for both formats after Amount I but evens out to the usual TPI after Amount II and III. Reliable amount by format ($F = 11.54$, df = 2/96, $p < .001$), amount by format by content ($F = 3.46$, df = 4/192, $p < .025$), and sequence by amount by format by content ($F = 7.10$, df = 4/192, $p < .001$) interactions support the observation that adding the combinations section (Amount II) helps most on near transfer items (F-format) in the content area just covered, especially for Sequence F Ss, and adding the final, binomial section (Amount III) helps most on near transfer items (F-format) in the content area just covered, expecially for Sequence G Ss (see Table 2).

TABLE 2

PROPORTION CORRECT RESPONSE FOR TWO INSTRUCTIONAL GROUPS BY PROBLEM FORMAT AND CONTENT FOR THREE AMOUNTS OF INSTRUCTION--EXPERIMENT II

Instructional Sequence	Amount of Instruction	Problem Content and Format					
		C-Content		J-Content		B-Content	
		F	S	F	S	F	S
F	I	.07	.24	.41	.17	.19	.12
G	I	.11	.23	.54	.56	.17	.21
F	II	.53	.14	.49	.24	.38	.34
G	II	.32	.44	.42	.36	.20	.18
F	III	.61	.22	.54	.34	.42	.39
G	III	.42	.46	.54	.44	.45	.24

Although the inclusion of specific content material was important (especially for near transfer items) under both instructional methods, it seems that Sequence G Ss were developing a structure that was far more capable of "adlibing" when the assimilation of essential content material was incomplete, i.e., after only Amount I was given. The broad assimilative set that is made up of propositional knowledge, apparently used by Ss in Sequence G, could be responsible for their ability to solve problems not specifically dealt with in the teaching booklet. These results do not suggest, however, that Sequence G Ss do not require presentation of specific content material; the huge increase in performance on F-format items by Sequence G Ss especially from Amount II to Amount III clearly indicates that presentation of content material is important especially for near transfer performance in both treatment groups.

In addition, the same pattern of TPI can be observed at each of three points in learning, although the interaction is not disordinal after Amount I largely due to the poor performance of Ss in Sequence F. The failure to obtain a reliable three-way interaction among sequence, format and amount (F = 2.49, df = 2/96, p > .100) allows us to retain the hypothesis stated in Experiment I that structural differences begin quite early and do not change much throughout learning. However, due to differences in the breadth of assimilative sets used, lack of specific content material is more detrimental for Sequence F Ss.

Figure 8 shows the performance of the two instructional groups by problem type at each of three points in learning. Although Sequence F Ss performed better on F-type problems while Sequence G Ss performed better on Q- and U-type problems as was found in Experiment I, the two-way interaction between sequence and type failed to reach a statistically significant level ($F = 1.63$, df = 4/384, $p < .150$). An investigation of the marginally significant three-way interactions--sequence by type by content ($F = 2.14$, df = 8/768, $p < .50$) and sequence by type by format ($F = 2.40$, df = 4/384, $p < .050$)--indicates that the sequence by type interaction was strongest for B-content and F-format.

Since TPI is complicated by content and format, and is not statistically significant in this case, there is little point in trying to locate where TPI begins; however, there is no evidence that the pattern of TPI is influenced by the amount of instruction. The failure of the sequence by amount by type interaction to reach statistical significance ($F = 1.17$, df = 3/384, $p > .200$) is consistent with the results of Experiment I indicating no structural change in how Ss in the two instructional groups encoded material over the course of learning.

An interesting fact, however, is that Sequence G is superior to Sequence F on all problem types after Amount I but they tend to even out after Amount II and III. A reliable sequence by amount interaction ($F = 3.30$, df = 2/96, $p < .050$) verifies this observation and permits

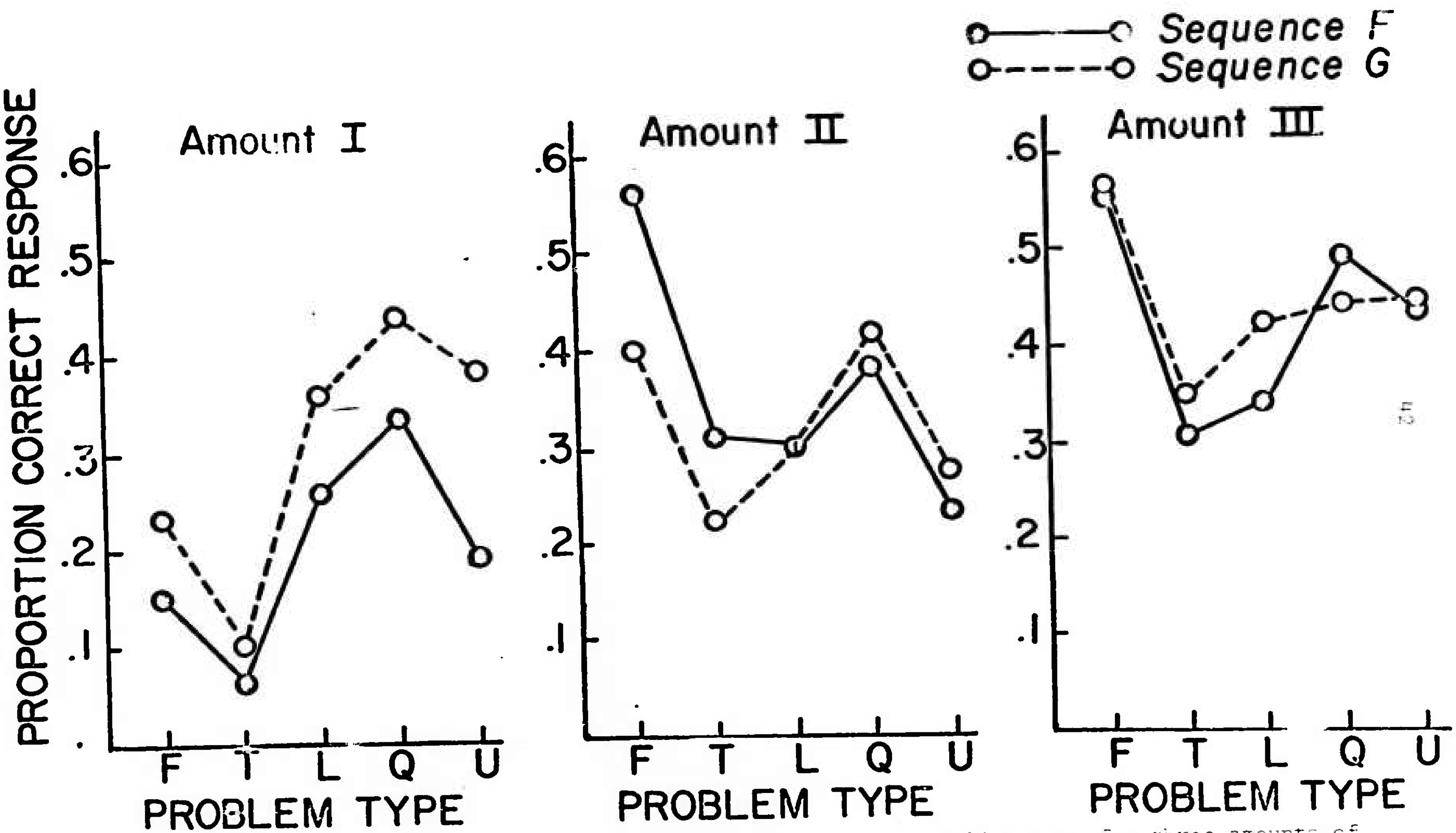

Fig. 8. Proportion correct response for two instructional groups by problem type for three amounts of instruction--Experiment II. The abscissa roughly represents the degree of transfer constituted as follows: F-type (familiar) and T-type (transformed) represent near transfer tasks, while L-type (Luchins) represents intermediate transfer, and U-type (unanswerable) and Q-type (question) represent far transfer.

the claim that adding the combinations section (Amount II) helped Sequence F Ss more than Sequence G Ss.

Reliable amount by type (F = 7.87, df = 8/384, p < .001), amount by type of content (F = 2.74, df = 16.768, p < .010), and sequence by amount by type by content (F = 2.10, df = 16/768, p < .050) interactions suggest that adding new content helped most on direct, "near" transfer items (F- and T-type) in the content area just covered with the addition of Amount II especially helping Sequence F Ss and the addition of Amount III especially helping Sequence G Ss (see Table 3). Again, these findings are consistent with the notion that, when specific content material is incomplete, the ability to use a wide band of propositional knowledge already existing in memory can benefit the Sequence G Ss while Sequence F Ss encode material in a way that is far more dependent on encoding specific material in a narrow band of algorithmic knowledge. Again, however, the importance of specific content for Sequence G Ss is manifested in the jump in near transfer performance (F- and T-type) for Amount III.

Finally, the performance of the two instructional groups by problem content at each of three points in learning is shown in Figure 9. As in Experiment I, reliable TPI suggests structural differences were obtained with Sequence F Ss superior on B-content and Sequence G superior to C- and J-content items (F = 13.09, df = 2.192, p < .001).

Again, the sequence by amount interaction is manifested in the fact that Sequence G outperformed Sequence F for all content areas

TABLE 3

PROPORTION CORRECT RESPONSE FOR TWO INSTRUCTIONAL GROUPS BY PROBLEM TYPE AND CONTENT FOR THREE AMOUNTS OF INSTRUCTION--EXPERIMENT II

Instructional Sequence	Amount of Instruction	Problem Content and Type														
		C-Content					J-Content					B-Content				
		F	T	L	Q	U	F	T	L	Q	U	F	T	L	Q	U
F	I	.03	.14	.06	.22	.33	.42	.06	.31	.47	.19	.00	.00	.42	.31	.06
G	I	.03	.08	.06	.14	.56	.07	.22	.75	.78	.33	.00	.00	.28	.39	.28
F	II	.50	.50	.14	.22	.33	.53	.06	.50	.61	.14	.67	.36	.25	.31	.22
G	II	.44	.36	.25	.36	.50	.61	.14	.44	.56	.19	.14	.17	.19	.33	.11
F	III	.50	.33	.25	.33	.61	.56	.11	.47	.72	.36	.58	.42	.31	.42	.31
G	III	.56	.52	.28	.22	.61	.67	.25	.64	.72	.44	.47	.28	.33	.39	.28

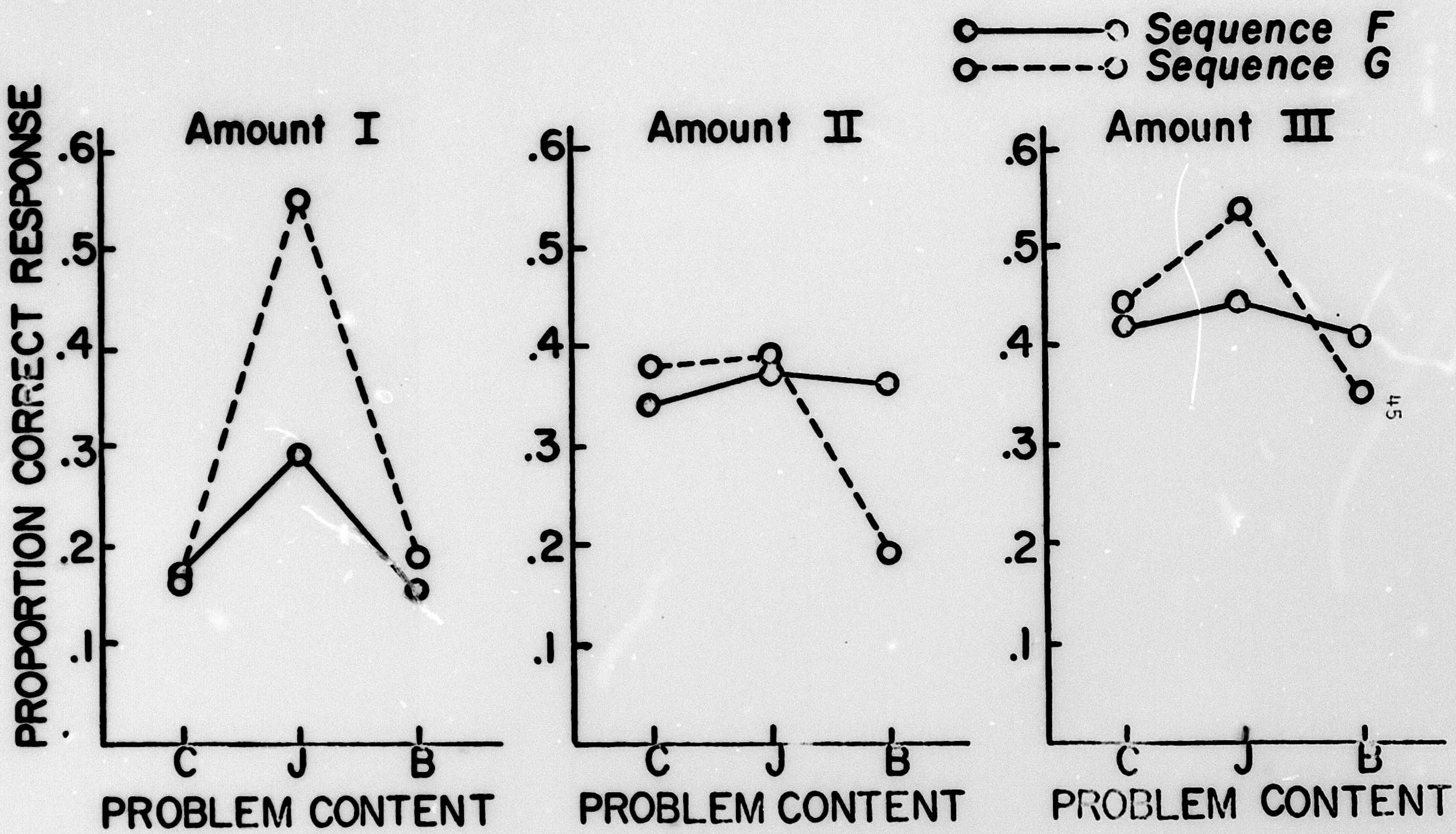

Fig. 9. Proportion correct response for two instructional groups by problem content for three amounts of instruction--Experiment II. The abscissa roughly represents the degree of transfer constituted as follows: B-content (binomial) represents near transfer, and C-content (combinations) and J-content (joint) represent far transfer.

after Amount I but evened out after Amount II and III. Again, reliable amount by content ($F = 7.68$, df = 4/192, $p < .001$) and sequence by amount by content ($F = 2.62$, df = 4/192, $p < .050$) interactions suggest that adding extra sections helped most--as might be expected--on test items about the content just covered and especially helped Sequence F Ss after Amount II and Sequence G Ss after Amount III.

As in Experiment I, the same general pattern of TPI seems to be present, at least in some form, at each of three points in learning. However, unlike Experiment I, a reliable sequence by amount by content interaction indicates that the structural differences changed as the amount of instruction was increased. Apparently, the presentation order used in Experiment II in which the more necessary content was not presented until Amount II, helped focus on the disproportionate importance of specific content for incomplete structures of Sequence F Ss relative to Sequence G Ss.

One interpretation that is consistent with the analysis of acquisition in terms of different assimilative sets established early in learning is that Sequence F Ss simply had less to work with after Amount I--i.e., very little content and mainly algorithmic knowledge--while Sequence G Ss had activated a broad range of existing propositional knowledge, useful in inventing solutions in the absence of specific information. However, the fact that performance jumped on near transfer items (B-content) for Sequence G Ss with addition of

Amount III again points to the importance of specific content for answering specific, near transfer problems.

Resilience of Structural Differences

This study also provides some information about the resilience of TPI under varying testing conditions. In Experiment I, the variable of memory support during testing--a sort of stress relief--failed to produce any major change in the structural differences among learning outcomes and only slightly influenced overall performance. Experiment II used the variable of time stress during testing--as a sort of stress evoker--in another attempt to disrupt both the qualitative (or structural) and quantitative (or overall) features of the performance of Ss who had acquired different cognitive structures.

Figure 10 gives the performance of the two instructional groups by problem format for the stress and no stress conditions. A reliable overall effect due to stress ($F = 12.5$, df = 1/96, $p < .001$) is manifested in the observation that the curves in the right panel are shifted down from those in the left panel. However, as can be seen in comparing the two panels, the structural differences as manifested by the usual sequence by format interaction (TPI) seem to be present both with and without stress; the failure to obtain reliable stress by format ($F < 1.00$, df = 1/96, $p > .20$) or sequence by stress by format interactions ($F < 1.00$, df = 1/96, $p > .20$) suggests that there is no evidence that stress had any effect on the structure of learning outcomes.

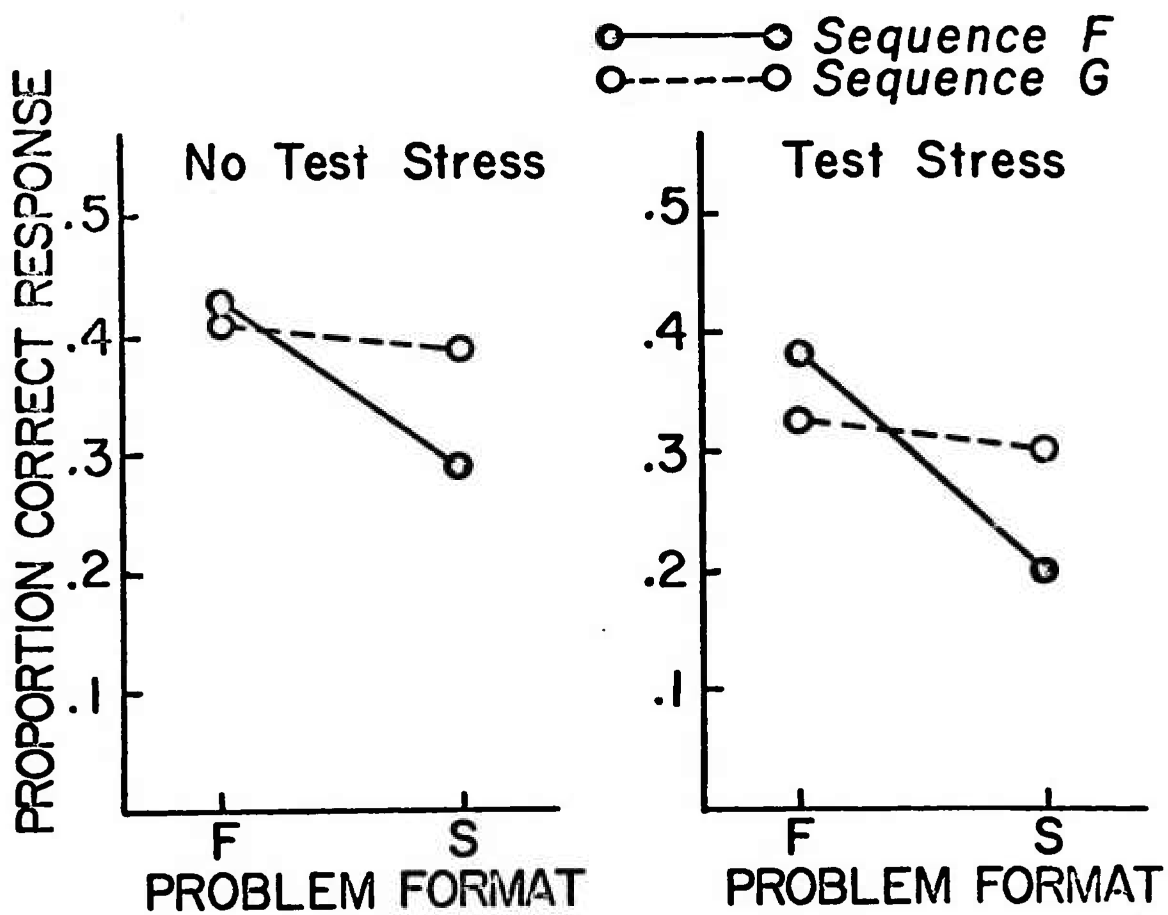

Fig. 10. Proportion correct response for two instructional groups by problem format for two amounts of stress during testing--Experiment II. The abscissa roughly represents the degree of transfer constituted as follows: F-format (formula) represents near transfer and S-format (story) represents far transfer.

Figure 11 shows the performance of the two instructional groups by problem type for the time stress and no time stress conditions. As can be seen, and as is indicated by a reliable stress effect ($\underline{F} = 12.50$, $df = 1/96$, $\underline{p} < .001$) but no reliable stress by type ($\underline{F} = 2.13$, $df = 11/384$, $\underline{p} > 1.00$) or sequence by stress by type ($\underline{F} < 1.00$, $df = 4/384$, $\underline{p} > .20$) interaction, test stress generally had the effect of reducing overall performance for both instructional groups and for all types of test problems. The structural differences between the two instructional groups, though not reaching statistical significance, do not seem altered by the introduction of test stress and there is no statistically reliable evidence that stress had anything but an overall quantitative effect on performance in this case.

Figure 12 gives the performance of the two instructional groups by problem content for the time stress and no time stress conditions. As with problem type and problem format, a reliable overall effect due to stress can be observed in the present case with the curves in the right panel simply shifted down from those in the left panel. Again, the structural differences between the two instructional groups is clearly present under both stress and no stress conditions, and the failure to obtain reliable stress by content ($\underline{F} < 1.00$, $df = 2/192$, $\underline{p} > .20$) or sequence by stress by content ($\underline{F} < 1.00$, $df = 2/192$, $\underline{p} > .20$) interactions provides no evidence that stress had any effect on the structure of learning outcomes.

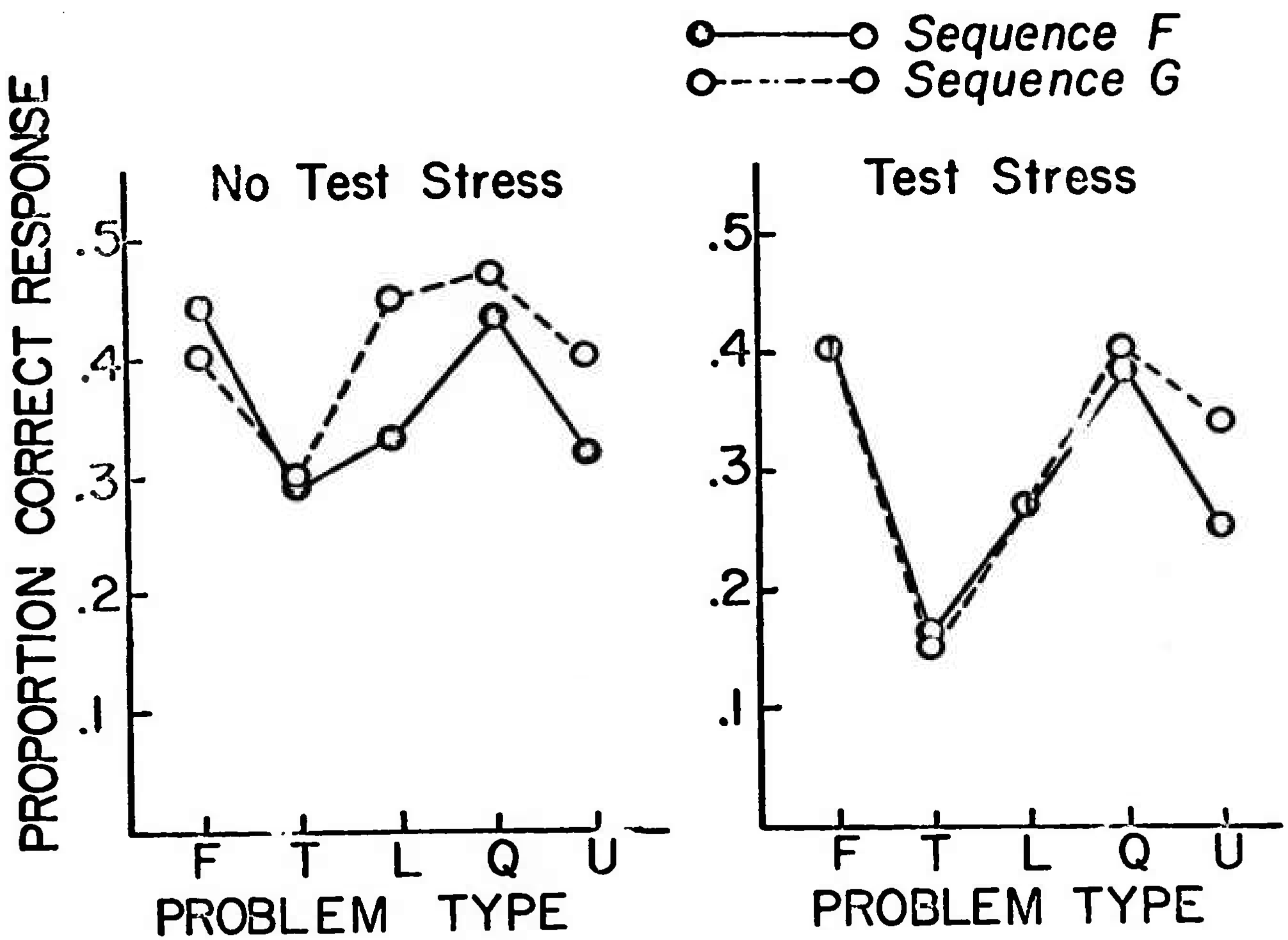

Fig. 11. Proportion correct response for two instructional groups by problem type for two amounts of stress during testing--Experiment II. The abscissa roughly represents the degree of transfer constituted as follows: F-type (familiar) and T-type (transformed) represent near transfer tasks, while L-type (Luchins) represents intermediate transfer, and U-type (unanswerable) and Q-type (question) represent far transfer.

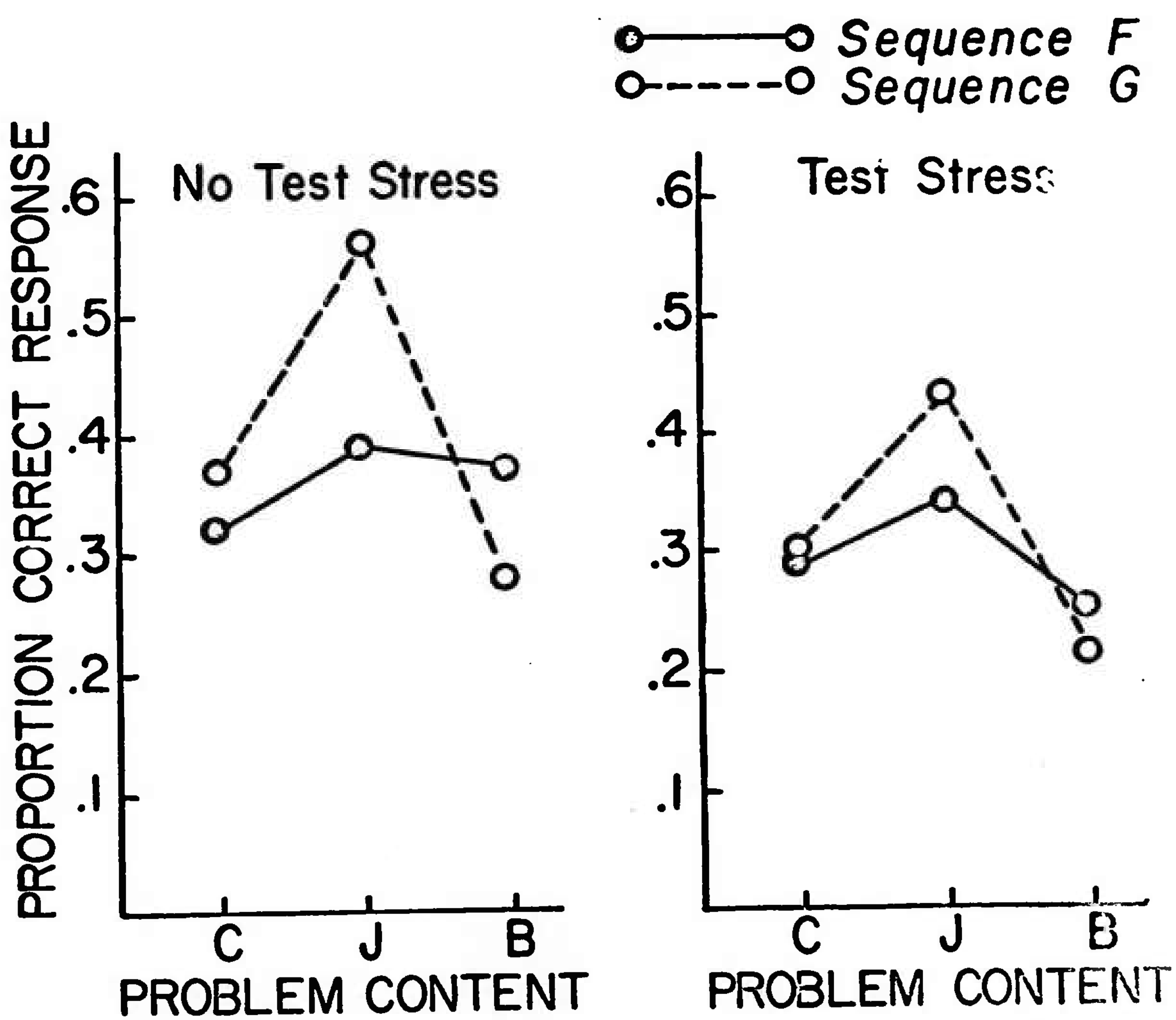

Fig. 12. Proportion correct response for two instructional groups by problem content for two amounts of stress during testing--Experiment II. The abscissa roughly represents the degree of transfer constituted as follows: B-content (binomial) represents near transfer, and C-content (combinations) and J-content (joint) represent far transfer.

As in Experiment I, three separate tests of the "resilience question," allowing for the possibility of TPI under one testing condition but not another, were conducted and failed to produce this result. These findings give some information about the effect of stress on problem solving that requires a pre-learned rule. Apparently, stress may have an overall inhibitory effect on the amount of performance, but there is no evidence that stress had any effect on the quality of performance. The encoding process, i.e., the assimilative set into which the rule is embedded, seems to be the major determinant of the structure of problem solving performance regardless of the amount of stress during testing. The amount of stress seems to influence only the absolute level of performance in this case.

CHAPTER IV

SUPPLEMENTAL STUDIES

In addition to the two main experiments, two smaller, supplemental studies were conducted. The supplemental studies used the same instructional materials and procedures as the main studies, but varied the testing situation in order to provide more information about the resilience and development of differences in learning outcomes.

Supplemental Study I

One study focused on retention performance in an attempt to test the endurance over time of the structural differences established by different instructional sequences.

Method

Nine Ss received the Sequence F, Amount III, Order ABC booklet and nine other Ss received the Sequence G, Amount III, Order ABC booklet. The 30-item transfer test was given--with no memory support and no time stress--two days later.

Results and Discussion

The performance of the two instructional groups by problem format, by problem type, and by problem content is shown in Figure 13. As can be seen, the two day retention interval does not seem to have destroyed the structural differences in learning outcomes as indicated by TPI. In partial replication of previous findings, Sequence F Ss excelled on near transfer items such as F-type, T-type, F-format and B-content while

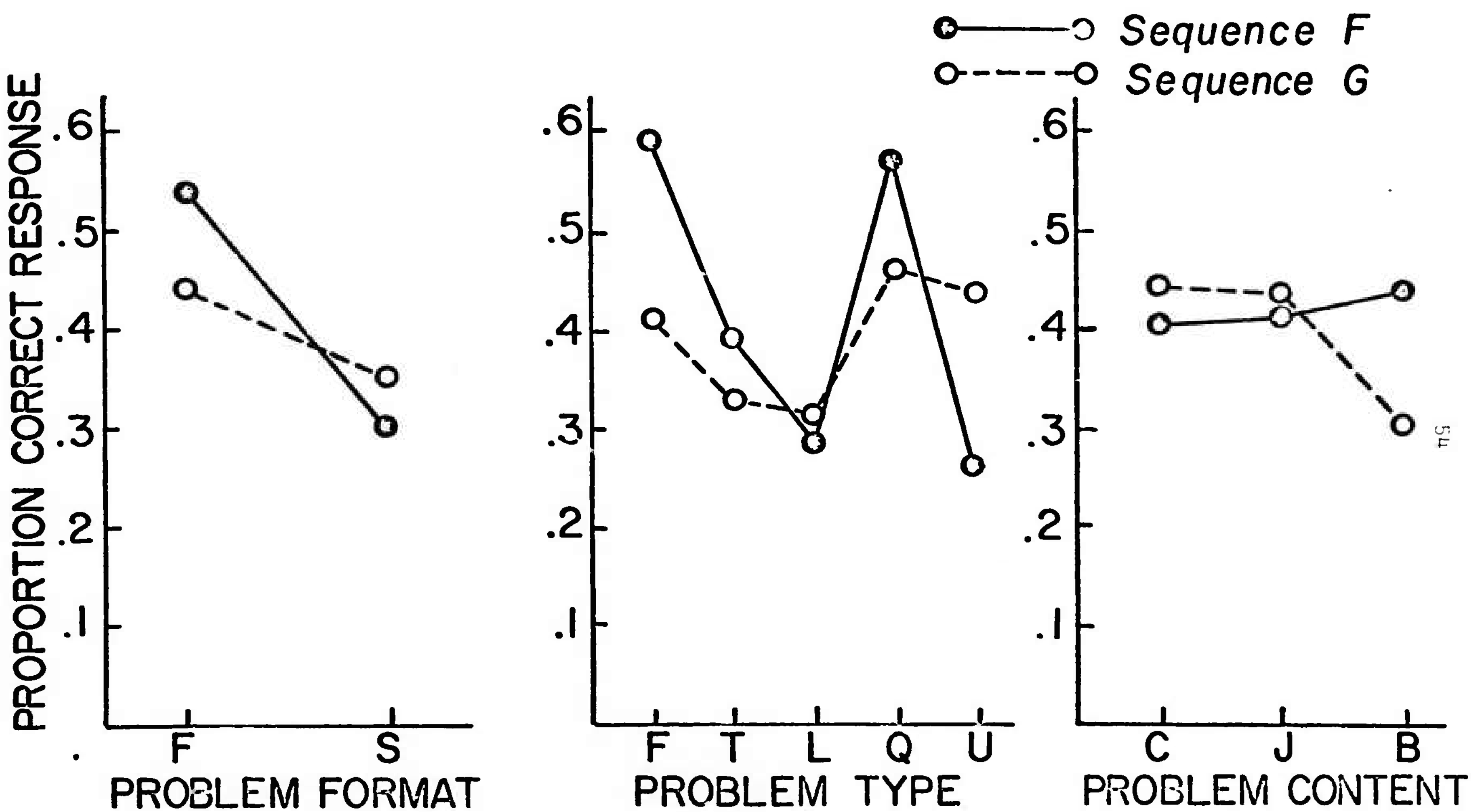

Fig. 13. Proportion correct response after two-day retention interval for two instructional groups by problem format, problem type, and by problem content--Supplemental Study I. The abscissa roughly represents the degree of transfer constituted as follows: F-format (formula), F-type (familiar), T-type (transformed) and B-content (binomial) represent near transfer, while S-format (story), U-type (unanswerable), Q-type (question), C-content (combinations), and J-content (joint) represent far transfer.

Sequence G excelled on far transfer items such as U-type, S-format, and C- and J-content.

As in the main studies, an analysis of variance revealed no significant difference between the groups in overall performance, and found TPI for type ($F = 3.58$, df = 4/64, $p < .025$) and TPI for format ($F = 5.60$, df = 1/16, $p < .050$) to be at statistically reliable levels. The sequence by content interaction failed to reach significant levels ($F = 2.09$, df = 2/30, $p > .10$); however, this may be a reflection of the low number of Ss involved. A significant three-way interaction among sequence, type and format ($F = 5.41$, df = 4/64, $p < .001$) indicates that the sequence by type interaction was much stronger for F-format items than for S-format items. As with the memory support and time stress variables, there was no strong evidence that the retention interval altered the structure of learning outcomes.

Supplemental Study II

The other supplemental study used a modified "method of reproduction" and questionnaire in order to assess S's judgment of what was important or what was supposed to be learned at each of three points in learning for the two instructional groups.

Method

Six Ss each were presented with either the Sequence F, Amount I, Order ABC booklet; the Sequence F, Amount II, Order ABC booklet; the Sequence F, Amount III, Order ABC booklet; the Sequence G, Amount I, Order ABC booklet; the Sequence G, Amount II, Order ABC booklet; or the Sequence G, Amount III, Order ABC booklet. All Ss were then given an immediate retention test and questionnaire. The test was to write down, with no memory support or time stress, what had been in the teaching program as if S were trying to teach it to another S coming in for the next session.

The traditional method of reproduction was modified in that S was encouraged to "make sense" out of the material as if he were explaining what he had just learned to someone who did not already know it, rather than to reproduce the material word for word as given. It was hoped this procedure, following Piagetian (Piaget, 1969) techniques with children, would provide some rudimentary information about how S stored the content material at each point in learning. In addition, S was asked to make up three test items to cover the material he had just written.

The questionnaire was designed to assess S's interpretation of the motives of the author of the teaching booklet, i.e., to assess S's ability to pick up the cues in the teaching booklet concerning what should be learned. Sixteen questions were constructed, each trying to get at whether S saw his task as (1) learning how to use the formula and the mechanical steps involved, or as (2) learning what the formula means and understanding the component concepts.

Representative questions are given below:

Pro-Calculating/Anti-Understanding: Based on my experience with the teaching booklet I would say that the formula does not necessarily have to make sense; all you have to do is learn how to use it.

Anti-Calculating/Pro-Understanding: To use the teaching program to best advantage you have to spend time thinking about how each concept in the formula relates to what you already know.

Subjects were asked to rate their agreement-disagreement on a 6-point scale, although for purposes of analysis 1, 2 or 3 were counted as disagree and 4, 5 or 6 were counted as agree.

Results and Discussion

The protocols from each S were coded for number of words and number of symbols, with each operator, each number and each formal notation character counting as a symbol (e.g., N! counts as 2 symbols; 15/2 counts as 3 symbols). The test items suggested by S were coded for their type, format, and content. A questionnaire score was obtained by summing the number of pro-calculating (or anti-understanding) agreements and pro-understanding (or anti-calculating) disagreements and dividing by the total number of questions.

Table 4 shows the average number of words, the average number of symbols, and the average number of both words and symbols given at each point in learning by Ss in the two instructional groups. Separate analyses of variance were performed on these data yielding some preliminary information about how S stores the material and S's ability to detect the emphasis of his teaching booklet at each of three points in learning. The protocols show that Ss in Sequence F used significantly more symbols overall ($F = 12.97$, df = 1/30, $p < .005$) and significantly fewer words overall ($F = 9.25$, df = 1/30, $p < .005$) in explaining what was taught than Sequence G Ss.

TABLE 4

AVERAGE NUMBER OF ELEMENTS IN REPRODUCTION FOR TWO INSTRUCTIONAL GROUPS AND THREE AMOUNTS OF INSTRUCTION--SUPPLEMENTAL STUDY II

Average Number of Words				
Instructional Sequence	Amount of Instruction I	II	III	Average
F	110	152	243	168
G	259	278	206	247
Ave.	184	215	225	

Average Number of Symbols				
Instructional Sequence	Amount of Instruction I	II	III	Average
F	95	187	190	157
G	111	107	88	102
Ave.	103	147	139	

Average Number of Words and Symbols				
Instructional Sequence	Amount of Instruction I	II	III	Average
F	205	339	433	326
G	370	385	293	350
Ave.	287	362	363	

One implication is that Ss were sensitive to the emphases of their respective teaching booklets. Sequence F may foster the attitude that what is to be learned is the mechanics of calculating with the formula, i.e., the building of internal connections (hence many symbols), while Sequence G may foster the attitude that what is to be learned is a conceptual understanding of what the components of the formula mean, i.e., the building of external connections (hence many words). Apparently, Ss were able to identify these emphases and use them as a key to encoding the material.

A more striking finding is that as the amount of instruction increased, both the number of words and the number of symbols output by Ss in Sequence F increased while the output by Ss in Sequence G stayed about the same or went down. Reliable sequence by amount interactions for words (F = 5.10, df = 2/30, p < .025), for symbols (F = 5.63, df = 2/30, p < .010) and for both (F = 6.91, df = 2/30, p < .005) support this observation. It appears that Sequence F Ss were adding more and more discrete pieces of information while Sequence G Ss were forming a tighter, more streamlined structure--less dependent on material presented by E because more already stored knowledge could be used.

The results concerning test items suggested by S also provide some information about what S thought was important or the goal of learning. Table 5 shows the average number of F- and T-type (as opposed to Q-type), F-format (as opposed to S-format), and B-content (as opposed to C- and J-content) problems given at three points in learning by Ss in the two instructional groups. Separate analyses of variance were performed.

Subjects in Sequence F suggested significantly more F- and T-type (F = 10.80, df = 1/30, p < .005) and F-format (F = 12.25, df = 1/30, p < .005) items than Ss in Sequence G, indicating an emphasis on using the formula on direct, near transfer; conversely, Sequence G Ss offered signficantly more Q-type and S-format items, indicating an

TABLE 5

AVERAGE NUMBER OF TEST ITEMS SUGGESTED BY METHOD AND AMOUNT OF INSTRUCTION--SUPPLEMENTAL STUDY II

Instructional Sequence	Amount of Instruction I	II	III	Average
Average Number of F-Format Items				
F	1.50	1.50	1.67	1.56
G	.33	.17	.00	.17
Ave.	.92	.83	.83	
Average Number of B-Content Items				
F	2.17	2.83	2.67	2.56
G	2.67	2.33	2.50	2.50
Ave.	2.42	2.58	2.58	
Average Number of F- and T-Type Items				
F	1.17	2.67	2.67	2.17
G	.83	1.33	1.33	1.17
Ave.	1.00	2.00	2.00	

emphasis on less direct, far transfer applications. The fact that there were no significant sequence by amount interactions for format ($\underline{F} < 1.00$, df = 2/30, $\underline{p} > .20$), for content ($\underline{F} < 1.00$, df = 2/30, $\underline{p} > .20$) nor for type ($\underline{F} = 1.20$, df = 2/30, $\underline{p} > .20$) suggests that S can pick up and retain the emphasis of his booklet equally well after Amount I, II or III.

Finally, Table 6 shows the proportion of pro-calculating/anti-understanding response on the questionnaire given by Ss in the two instructional groups at three points in learning. As can be seen, and as is revealed by an analysis of variance, Sequence F Ss were significantly more agreeable with pro-calculating statements and more disagreeable with pro-understanding statements than Sequence G Ss ($F = 14.97$, df = 1/30, $p < .001$). However, no reliable sequence by amount interaction ($F < 1.00$, df = 2/30, $p > .20$) was obtained thus suggesting that the overall difference between the groups occurred consistently at each point in learning. Again, support is provided for the claim that S can interpret and retain the emphasis or expectation of his teaching booklet, even only after Amount I or II.

TABLE 6

PROPORTION "PRO-CALCULATING" OR "ANTI-UNDERSTANDING" RESPONSE ON QUESTIONNAIRE, BY METHOD AND AMOUNT OF INSTRUCTION--SUPPLEMENTAL STUDY II

Instructional Sequence	Amount of Instruction I	II	III	Average
F	.62	.64	.56	.60
G	.39	.38	.48	.42
Ave.	.51	.51	.52	

CHAPTER V

SUMMARY AND CONCLUSIONS

Development of Structural Differences

In Experiment I, Experiment II, and the supplemental studies there was clear evidence that, of those Ss presented with the entire lesson (i.e., Amount III), Ss in different instructional groups had acquired learning outcomes that differed in a qualitative or structural way. This inference was based on a pattern of results which partially replicated three earlier studies (Mayer & Greeno, 1972; Egan & Greeno, 1973; Greeno & Mayer, in prep.): Ss in Sequence F and Sequence G performed at fairly equal levels overall, but TPI was displayed in that Sequence F Ss excelled on "near" transfer items such as F-format, F-type, and C-content while Sequence G Ss excelled on "far" transfer items such as S-format, Q- and U-type, and C- and J-content.

The final learning outcomes of the two groups were characterized in terms of internal and external connections with Ss in Sequence F acquiring structures with strong internal connections (e.g., P is linked to R by the operation of exponentiation) and weak external connections (e.g., P is related to general experience with probability of an event), and Ss in Sequence G acquiring structures with weak internal connections and strong external connections. This kind of analysis seems to be useful in describing differences in cognitive structure that allow Sequence F to excel on direct plug-in-to-the-formula problems and allow Sequence G to excel on more interpretative tasks such as recognizing unanswerable problems, understanding how the formula relates to story problems, or dealing with logical components of the formula.

A new goal of the present set of experiments was to generate some useful information concerning how these structural differences develop over the course of learning, and especially, to help choose which of two descriptions of the acquisition process seems more appropriate: an acquisition process in which more and more of different kinds of material (e.g., internal links vs. external links) is being added to the cognitive structures of different Ss, or an acquisition process in which the content material is being embedded within different areas of existing knowledge (e.g., algorithmic knowledge vs. propositional knowledge) by different Ss.

The two main studies provide empirical support for the second idea, that different instructional sequences activate different kinds of existing knowledge (i.e., different receptive sets) quite early in learning, and that structurally different learning outcomes result from content material being encoded within the context of these different "sets" throughout learning. The concepts of algorithmic and propositional knowledge were found useful in describing the difference in Ss' assimilative sets: Sequence F seems to activate only a narrow band of experience with mathematical operations and calculating with a formula (fitting the description of algorithmic knowledge), while Sequence G seems to activate a much wider band of experience with probability situations (fitting the description of propositional knowledge).

Empirical support for the proposal that the structure of learning outcomes is determined both by the content material (in this case, the same mathematical rule for all Ss) and the kind of existing knowledge to which it is assimilated (in this case, either algorithmic or propositional knowledge) comes from several findings. In Experiment I, there was no evidence that TPI (and by inference, the structure of learning outcomes) was any different after Amount I, Amount II or Amount III. If Ss had been adding more and more of different kinds of material under the two instructional methods, one would have expected a weak TPI after Amount I that got stronger after Amount II and strongest after Amount III since each added section should add more of different kinds of content to Ss' cognitive structure. However, the fact that TPI was constant, i.e., equally strong and observable, across all three amounts of instruction is consistent with the idea that different assimilative sets were evoked quite early in learning and used throughout learning to encode incoming information. Thus although the overall level of performance was increased for both instructional groups as more material was presented, the differences in how it was structured were constant throughout learning.

The results of Experiment II seem to replicate Experiment I except that Sequence F Ss performed reliably worse than Sequence G Ss after Amount I; the two groups performed at equal levels of overall performance for more amounts of instruction (as in Experiment I). Apparently, the ordering of instruction in Experiment II allowed for a substantial amount of content material to be lacking in Amount I. The fact that the lack of specific material hurt Sequence F far more than Sequence G is consistent with the idea that Sequence G Ss had access to a wide range of already existing knowledge that could be used in inventing solutions to problems not yet "taught" while Sequence F Ss had to depend on a much narrower, more specific kind of knowledge. However, the importance of specific content for Sequence G Ss is indicated by a large increase in "near transfer" performance with the addition of Amount III. Apparently, the activation of propositional knowledge in of itself--though helpful in creating solutions to far transfer items--is not sufficient for good problem solving performance and especially needs the inclusion of specific content for solving straightforward near transfer items. Although beyond the scope of the present study, it remains for further studies to investigate whether it is possible to construct a teaching sequence that insures the advantages of both of the present teaching booklets--i.e., both good near and far transfer.

Finally, the results of Supplemental Study II show that Ss were easily able to recognize the differences in emphasis of the two booklets. This inference is supported by the fact that Sequence F Ss more often report (on a post-experimental questionnaire) the purpose of the lesson as memorizing how to use a rule and Sequence G Ss report the purpose as understanding what the rule means, and the fact that these differences do not reliably change from Amount I to Amount II to Amount III. Secondly, the fact that Sequence F Ss gave more symbols in their reproduction protocols and Sequence G gave more words also indicates an ability not only to recognize but to reproduce differences in the emphases of the teaching booklets. Another interesting finding, consistent with the foregoing description of the acquisition process, is that the reproduction protocols of Sequence F Ss got longer and longer as the amount of instruction was increased, but the length of protocols of Sequence G Ss fell or stayed about the same across all amounts of instruction. Apparently, Sequence F Ss were far more dependent on adding specific pieces of information to cognitive structure while Sequence G "saved space" by integrating the information into existing structures.

Resilience of Structural Differences

There were two separate experimental attempts to disrupt the characteristic patterns of TPI by varying the amount of stress during testing and forcing, it was assumed, Ss with structurally different learning outcomes to process their cognitive structures in the same way. Both attempts failed, suggesting that post-learning manipulations aimed at S's internal processing style do not have the same effect on outcome performance as do such manipulations during learning. Apparently, once the problem solving rule is encoded, test stress can influence the absolute level of performance but has negligible effect on the structure of problem solving performance.

For example, in Experiment I Ss tested under the memory support condition (a sort of stress relief) performed at a slightly higher overall level relative to no memory support Ss, but also exhibited a pattern of TPI indistinguishable from Ss tested under no memory support conditions. In Experiment II, Ss tested under time stress, although showing a relative decrement in overall performance, also exhibited a pattern of TPI indistinguishable from Ss tested under no time stress conditions. Finally, Ss in Supplemental Study I gave evidence of the typical patterns of TPI after a two day retention interval, although there were no control groups in this study upon which to base comparisons.

The effect of stress on the application of a pre-learned rule to problem solving situations seems to be mainly quantitative (Experiment II) and in some cases not even that (Experiment I). There was some evidence that stress relief allowed better performance on the harder, "far" transfer problems (Experiment I) but stress affected both kinds of cognitive structures about the same in this respect. However, the qualitative differences in learning outcomes were not strongly influenced by stress or retention intervals, and this finding--while requiring further study--seems to demand the tentative conclusion that the structure of learning outcomes, once established, may be a fairly permanent feature of S's cognitive life.

Relation to Discovery Learning and Creative Problem Solving

These results also give some hints about the prerequisites for discovery learning and creative problem solving. For the kind of teaching whose goal is creative problem solving, such as displayed by Sequence G, Amount I Ss in Experiment II (i.e., inventing solutions with very little specific content material having been given), it is clear that a substantial bank of what has been called "propositional knowledge" must be available. For Ss who do not have a well integrated set of general experiences in the required area (e.g., in this case, in probability of events), it seems that attempts to achieve learning outcomes that can support creative problem solving will fail. This is so because the kind of learning outcome that supports creative problem solving (e.g., as displayed by Sequence G Ss) is acquired by embedding the problem solving rule in a bank of the appropriate propositional knowledge. Instead of demanding that Ss who lack the appropriate prerequirite knowledge learn a given problem solution rule, and thereby insuring either no learning or a very specific kind of encoding, it seems a better educational practice to first provide Ss with the necessary prerequisite concepts. Previous findings (Egan & Greeno, 1973; Greeno & Mayer, in prep.) which show individual differences in specific prerequisite knowledge to be far more important for "discovery" learning methods than for "rule" learning methods support this argument.

In short, the results seem to indicate that for those kinds of learning whose goal is a quick, efficient ability to perform a given set of operations (e.g., arithmetic operations) only a narrow set of existing knowledge is necessary, but for learning that supports creative problem solving (e.g., reconstructing a procedure for new problems) the need to make sure S possesses the prerequisite knowledge is essential. These results suggest that discovery teaching to Ss who lack the appropriate propositional knowledge will not result in discovery learning or creative problem solving, and may result in no learning at all.

Limitations

The generality of these findings is limited by a number of factors including the pecularities of the binomial as the to-be-taught rule. The concept of binomial probability is expressed as a fairly complicated and seemingly arbitrary mathematical rule made up of a set of conceptual variables (with potential meaning in terms of S's existing propositional knowledge) and a set of mechanical operators (with relevance for a narrow range of experience with arithmetic calculations). Thus it seems that the conclusions concerning the binomial are most likely to apply in the learning of problem solving procedures which are expressed as complicated rules and which can be structured in memory in terms of meaningful variables or in terms of algorithmic operations. A main instructional variable in such cases is the degree to which the potential "meanings" of the component variables are emphasized versus the degree to which the mechanical operations are emphasized. For example, in an unpublished study using this instructional variable in conjunction with teaching Bayes' Theorem evidence for the expected structural differences was obtained (Stiehl, 1973). It seems clear, however, that more studies using a wider range of problem materials, as well as different Ss, and long-term experimental designs would be useful.

APPENDIX A

COMPLETE TEACHING PROGRAM TEXTS FOR SEQUENCE F AND G

Sequence F - Part I (Introduction)

$P(R|N) = C(N,R) \times P^R \times (1-P)^{N-R}$

This is called the binomial formula. It can be used to find the probability that something called "success" occurs R times in N trials.

The symbol P(R|N) stands for the probability of R successes in N trials. R is the number of successes and N is the number of trials. For example, if R = 2 and N = 4, P(R|N) is the probability of 2 successes in 4 trials. We might roll a die 4 times and define success as rolling a 5 or 6. If R = 2 then P(R|N) is the probability of rolling a 5 or 6 twice in 4 rolls.

To find the value of P(R|N) we need to find three values. The formula is:

$P(R|N) = C(N,R) \times P^R \times (1-P)^{N-R}$.

Thus, P(R|N) is the product of three terms:

(1) C(N,R)

(2) P^R

(3) $(1-P)^{N-R}$

You will need to remember how to find the value of each of the three terms in the formula for P(N|R). The first term is C(N,R). As we will see later the value of C(N,R) depends on N--the number of trials-- and on R--the number of successes. The second term in the formula is P^R. R is the number of successes and P is the probability of success. The third term in the formula is $(1-P)^{N-R}$. 1-P is the probability of failure and N-R is the number of failures.

Sequence F - Part II (Combinations)

As we said earlier, the first step to finding out what P(R|N) equals is to calculate the value of C(N,R). To find the value of C(N,R) use the formula:

$$C(N,R) = \frac{N!}{R! \times (N-R)!}$$

(Remember that the "!" symbol stands for factorial. To calculate N!--read "N factorial"--multiply N times N-1 times N-2 and so on down to 1. That is, N! = N(N-1)(N-2)(N-3) ... (1). For example 5! = 5 x 4 x 3 x 2 x 1 = 120 or 3! = 3 x 2 x 1 = 6. Also note that by definition 1! = 1 and 0! = 1.)

To find the value of C(N,R):

(a) Take N ....N
(b) Make it N! ....N!
(c) Take R ....R
(d) Make it R! ....R!
(e) Take N minus R ....N-R
(f) Make it (N-R)! ....(N-R)!
(g) Multiply R! times (N-R)! .....R!(N-R)!
(h) Divide R!(N-R)! into N! ......N!/R!(N-R)!)

As an example of the formula C(N,R) = N!/(R! x (N-R)!), think of R = 1 and N = 4. Then,

$$C(N,R) = \frac{4!}{1!3!} = \frac{4x3x2x1}{1x3x3x1} = 4.$$

To find the value of C(4,1) the steps are:

(a) N = 4
(b) N! = 4 x 3 x 2 x 1 = 24
(c) R = 1
(d) R! = 1
(e) N-R = 3
(f) (N-R)! = 3 x 2 x 1 = 6
(g) R!(N-R)! = 1 x 3 x 2 x 1 = 6
(h) N!/(R!(N-R)!) = 4 x 3 x 2 x 1/1 x 3 x 2 x 1 = 24/6 = 4.

As another example if we define successes as 5 or 6 and roll a die five times getting three successes, we have $N = 5$ and $R = 3$, and the value of $C(N,R)$ is:

$$C(N,R) = \frac{5!}{3! \times 2!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 2 \times 1} = 10.$$

You can check this answer by carefully going through each step of the formula for $C(N,R)$.

Sequence F - Part IIIa (Joint Probability-P^R Term)

Recall that $P(R|N) = C(N,R) \times P^R \times (1-P)^{N-R}$. You know how to find $C(N,R)$.

To find P^R we must find P, the probability of success. In general P is the proportion of trials on which a success would occur if there were a very large number of trials. For example, if success is defined as rolling a 5 or 6 on a die, then $P = 1/3$.

To find P^R, simply take P and raise it to the Rth power. (Note that $P^1 = P$ and $P^0 = 1$.) To find the value of P^R:

(a) Take P
(b) Raise it to the Rth power.

For example, if $P = 1/3$ and $R = 3$, then $P^R = (1/3)^3 = 1/27$. To find the value of P^R the steps are:

(a) $P = 1/3$
(b) $P^R = (1/3)^3 = 1/27$.

As another example, if success is rolling a 5 or 6 on a die and the number of successes is two, then $P = 1/3$, $R = 2$ and $P^R = (1/3)^2 = 1/9$. The steps to finding P^R are just to take P and then raise it to the Rth power.

Sequence F - Part IIIb (Joint Probability-$(1-P)^{N-R}$ Term)

Remember the formula is $P(R|N) = C(N,R) \times P^{R} \times (1-P)^{N-R}$. You now know how to find $C(N,R)$ and P^{R}.

To find the value of $(1-P)^{N-R}$ there are three steps:

(a) Take P ....P
(b) Subtract P from one 1-P
(c) Take N ....N
(d) Subtract R from N ....N-R
(e) Take 1-P and raise it to the (N-R)th power ..$(1-P)^{N-R}$

For example, if success is a 5 or 6 then $P = 1/3$ and $1-P = 2/3$. If there are five trials and two successes then $N = 5$, $R = 2$, and $N-R = 3$. In this case,

$$(1-P)^{N-R} = (1 - 1/3)^{5-2} = (2/3)^{3} = 8/27.$$

The probability of success is P and the probability of failure is 1-P. Similiarly, the number of successes is R and the number of failures is (N-R). In our example, the probability of failure is 2/3 and the number of failures is 3. Thus,

$$(1-P)^{N-R} = (2/3)^{3} = 8/27.$$

Remember, if $P = 1/3$, $N = 5$, and $R = 2$ then the steps to finding the value of $(1-P)^{N-R}$ are:

(a) $P = 1/3$
(b) $1-P = 2/3$
(c) $N = 5$
(d) $N-R = 3$
(e) $(1-P)^{N-R} = (2/3)^{3} = 8/27.$

Sequence F - Part IV (Binomial Probability)

Now we will summarize what you have learned. The binomial formula is:

$P(R|N) = C(N,R) \times P^R \times (1-P)^{N-R}$.

You find the value of $P(R|N)$ in four stages:

(1) First, find $C(N,R)$ using the formula, $C(N,R) = N!/(R! \times (N-R)!)$.
(2) Next, find P^R. Just take P and raise it to the Rth power.
(3) Next, find $(1-P)^{N-R}$. Subtract P from one to get 1-P, then subtract R from N to get N-R, and then raise 1-P to the N-R power, giving $(1-P)^{N-R}$.
(4) When you have found the values of the three terms, find their product. This gives $P(R|N)$, the probability of R successes in N trials.

For example, suppose that $P = 3/4$, $N = 3$ and $R = 2$.

(1) First, find $C(N,R)$. $N = 3$ and $R = 2$ so $C(N,R) = 3!/(2! \times 1!) = 3$.
(2) Next, find P^R. $P = 3/4$ and $R = 2$ so $P^R = (3/4)^2 = 9/16$.
(3) Next, find $(1-P)^{N-R}$. $1-P = 1/4$ and $N-R = 1$ so $(1-P)^{N-R} = (1/4)^1 = 1/4$
(4) The probability of 2 successes in 3 trials with $P = 3/4$ is: $P(R|N) = 3 \times 9/16 \times 1/4 = 27/64$.

Here is another example. If success is defined as rolling an even number on a die, the probability of three successes in five trials is $P(R|N)$ with $P = 1/2$, $R = 3$ and $N = 5$.

(1) First, $C(N,R)$ has $N = 5$ and $R = 3$. $C(N,R) = 5!/(3! \times 2!) = 10$.
(2) Next, $P = 1/2$, $R = 3$ so $P^R = (1/2)^3 = 1/8$.
(3) Next, $1-P = 1/2$ and $N-R = 2$ so $(1-P)^{N-R} = (1/2)^2 = 1/4$.
(4) Finally, $P(R|N) = 10 \times 1/8 \times 1/4 = 10/32 = 5/16$.

Sequence G - Part I (Introduction)

Two important concepts are trials and outcomes. A trial is something we do. The outcome of a trial is just what happens on the trial. Usually there are several possible outcomes of a trial.

For example, imagine rolling a die. Rolling the die is a trial. The possible outcomes are the numbers 1, 2, 3, 4, 5, and 6. The number that comes up when we roll the die is called the outcome of the trial.

Another important concept is success. We define a success as one or more of the possible outcomes. Then if one of those outcomes occurs, we have a success.

In rolling a die we could decide to define success as rolling a 5 or a 6. Then if 5 or 6 comes up, a success has occurred.

We can define success in different ways. A success might be rolling an even number. Then if the outcome is 2, 4, or 6 we have a success. Or a success might be not -4. Then success outcomes are 1, 2, 3, 4, 5, and 6.

The next concept to learn is the probability of success. The probability of success is the proportion of trials on which a success would occur if there were a large number of trials. For example, if a success is an outcome of 5 or 6 in rolling a die, then the probability of success is 1/3. This is so because we expect a 5 or 6 to come up on about one-third of the times the die is rolled.

In other words, the probability of a success is the number of success outcomes divided by the total number of outcomes (including success outcomes) if all the outcomes have an equal chance. If a success is 5 or 6 and all the possible outcomes have an equal chance, then probability of success is 2/6 or 1/3 because there are two possible outcomes with six outcomes in all.

On the other hand, a failure occurs whenever success does not occur. For example, if we say a success is either of the outcomes 5 or 6 in rolling a die, then the outcome 1, 2, 3, or 4 could be a failure. The probability of failure is: Probability of success subtracted from one. If success is rolling 5 or 6, probability of success is 1/3 and probability of failure is 2/3.

The next concept to learn is a sequence. A sequence is what happens when we conduct several trials, one after the other. Suppose each trial is rolling a die and we define a success in some way. If we roll the die five times we might obtain the sequence (failure, failure, success, failure, success), or more simply (F, F, S, F, S).

Any sequence has a probability. The probability of a sequence is the product of the probabilities of the individual events. For example, if probability of success is 1/3, then the sequence (success, failure, success) has probability $(1/3)\times(2/3)\times(1/3) = 2/27$. The sequence (F, F, S, S, F) has probability $(2/3)\times(2/3)\times(1/3)\times(1/3)\times(2/3) = 8/243$.

We let the number of trials in a sequence be symbolized by the letter N and the number of successes in those trials is called R. The number of failures is N-R.

N = number of trials
R = number of successes
N-R = number of failures

For example, in the sequence (S, F, S) there are three trials and two successes. Therefore, N = 3, R = 2, and N-R = 1.

Sequence G - Part II (Combinations)

As we said earlier, when a trial is repeated several times we get a sequence of outcomes, consisting of a pattern of successes and failures. But different outcome sequences can have the same number of successes. For example, in three trials we can have two successes in three different ways: (S,S,F), (S,F,S), and (F,S,S).

The number of different sequences having R successes in N trials is called the number of combinations, denoted by C(N,R). Think of C(N,R) as the number of ways R successes can occur in N trials.

To find the number of ways R successes can occur in N trials take N! (read "N factorial") divided by the product R! x (N-R)!. (Remember that "!" means take the number before it times itself minus one, times that number minus one, and so on down to one, e.g., 5! = 5x4x3x2x1 = 120 or 3! = 3x2x1 = 6. Note that by definition 1! = 1 and 0! = 1).

Use the formula:

$$C(N,R) = \frac{N!}{R!x(N-R)!} = \text{number of ways R successes can occur in N trials.}$$

$$\frac{N!}{R!x(N-R)!} = \frac{N(N-1)(N-2)\ldots.1}{(R(R-1)(R-2)\ldots.1)x(N-R)(N-R-1)(N-R-2)\ldots.1)}$$

As an example of finding C(N,R), think of one success in four trials. There are four sequences: (S,F,F,F), (F,S,F,F), (F,F,S,F), (F,F,F,S). This agrees with the formula:

$$C(N,R) = \frac{N!}{R!x(N-R)!} = C(4,1) = \frac{4!}{1!x3!} = \frac{4x3x2x1}{1x3x2x1} = 4.$$

If we define success as 5 or 6 and roll a die five times, the number of ways to get three successes is:

$$C(N,R) = C(5,3) = \frac{5!}{3!x2!} = \frac{5x4x3x2x1}{3x2x1x2x1} = 10.$$

As a check you can note that the 10 possible sequences are: FFSSS, SFFSS, SSFFS, SSSFF, FSFSS, SFSFS, SSFSF, FSSFS, SFSSF, FSSSF. The C(N,R) formula is just a quick way of finding out how many ways R successes can occur in N trials.

Sequence G - Part III (Joint Probability)

Now you know how to find the number of different sequences having R successes in N trials. In order to find the probability of any single sequence, you have to recall that if a sequence has N trials, then the probability of the sequence has N terms. For example, if probability of success is 1/3, the sequence (S,F,F,F) has probability (1/3)x(2/3)x(2/3)x(2/3). Note that N = 4 and so the probability has four terms.

If probability of success is 1/3, the sequence (S,F,S,S,F) has probability (1/3)x(2/3)x(1/3)x(1/3)x(2/3). Again N = 5 and the number of terms in the probability is 5. Actually, this way of expressing the probability is the same as $(1/3)^3x(2/3)^2$. Keep in mind that R = 3 and N-R = 2 for this sequence.

In the sequence (F,F,S,S,F,F), N = 6, R = 2, N-R = 4. If probability of success is 1/5, the probability of the sequence is $(4/5)x(4/5)x(1/5)x(1/5)x(4/5)x(4/5) = (1/5)^2x(4/5)^4 = (1/5)^Rx(4/5)^{N-R}$. Note that these last two expressions are just shorter ways of expressing the probability. They all equal 256/15625.

The probability of success is symbolized by the letter P. The probability of failure is 1-P. For example, if a trial is rolling one die and success outcomes are 5 or 6, then P = 1/3 and 1-P = 2/3.

P = probability of success
1-P = probability of failure

The probability of any sequence having R successes and N-R failures is $P^R x(1-P)^{N-R}$. For example, the probability of the sequence (S,F,F,S,S) is $P^3 x(1-P)^2$. Note that this agrees with the longer version: P x (1-P) x (1-P) x P x P. If P = 1/2 then the probability of the sequence equals $(1/2)^3 x(1/2)^2 = 1/32$. Note that this agrees with the longer method of multiplying (1/2)x(1/2)x(1/2)x(1/2)x(1/2) = 1/32.

As a quick way of finding the probability of a sequence use the formula:

$P^R x(1-P)^{N-R}$ = probability of R successes and N-R failures in any single sequence

(Note: By definition $P^0 = 1$, $(1-P)^0 = 1$, $P^1 = P$, and $(1-P)^1 = (1-P)$.)

Sequence G - Part IV (Binomial Probability)

Finally you will learn about the probability of R successes in N trials. There are a number of sequences having R successes in N trials (that is, the number of combinations). Each of these sequences has the same probability: $P^R x(1-P)^{N-R}$.

The probability of R successes in N trials is the probability of getting one of the sequences in the set of sequences having R successes. For example, if N = 3 and R = 2, there are three possible sequences:

(SSF), (SFS), (FSS). If P = 3/4, each sequence has probability $(3/4)^2$x $(1/4)^1$ = 9/64. We have two successes in three trials if any of the three sequences occurs.

Going on, if there are three sequences in the set with two successes, and each sequence has probability 9/64, then the probability of obtaining one of those three sequences is 3 x 9/64 = 27/64. We know that the number of sequences having R successes in N trials is C(N,R). Also, the probability of each such sequence is $P^R x(1-P)^{N-R}$. Thus the probability of R successes in N trials is:

$C(N,R) \times P^R x(1-P)^{N-R}$ = probability of R successes in N trials.

We use the symbol P(R|N) to stand for the probability of R successes in N trials. This gives a formula called the binomial formula:

$P(R|N) = C(N,R)xP^R x(1-P)^{N-R}$.

For example, the probability of one success in four trials is denoted as P(1|4) since R = 1 and N = 4. If P = 1/3, then P(1|4) = C(4,1)x$(1/3)^1$x $(2/3)^3$. C(4,1) = 4. $(1/3)^1$x$(2/3)^3$ = 8/81. So P(4|1) = 4 x (8/81) = 32/81 or .395. Note that this equation is just a simple way of saying that the probability of R successes in N trials equals the number of possible sequences with R successes in N trials times the probability of any one of those sequences.

APPENDIX B

PROBLEMS GIVEN IN 30-ITEM TRANSFER POSTTEST

Letters in parenthesis after each number indicate problem format, type, and content respectively.

1-(F,F,C): $P = 1/2$, $N = 8$, $R = 5$ What is $C(N,R)$?

2-(F,T,C): $N = 3^2$, $R = (2/3) \times N$, $P = 1/2$ What is $C(N,R)$?

3-(F,L,C): $C(N,R) = 10$, $R = 3$ What is N?

4-(F,Q,C): Can $C(N,R)$ be smaller than R? Can $C(N,R)$ be smaller than N?

5-(F,U,C): $N = 2$, $R = 3$, $P = 1/2$ What is $C(N,R)$?

6-(S,F,C): A coin is flipped six times, giving a sequence of heads and tails. How many different sequences contain two heads and four tails?

7-(S,T,C): If a fair die is rolled 4 times and the number of successes equals half the number of trials, how many different sequences could be generated?

8-(S,L,C): There are 10 different sequences that have exactly two successes. All the sequences have the same length. How long are they?

9-(S,Q,C): If a die is rolled 100 times and success is defined as a 1 or 2, does the number of possible sequences with 47 successes equal the number of sequences with 47 failures?

10-(S,U,C): In flipping a coin, how many different sequences have the same number of successes as failures?

11-(F,F,J): $P = .75$, $N = 9$, $R = 4$ What is $P^R \times (1-P)^{N-R}$?

12-(F,T,J): $P = 3(1-P)$, $N = 6$, $R = N-R$ What is $P^R \times (1-P)^{N-R}$?

13-(F,L,J): $P^R \times (1-P)^{N-R} = 1/2$, $N = R$ What is P^R?

14-(F,Q,J): Can P be greater than 1-P? Can 1-P be greater than P?

15-(F,U,J): $N = 5$, $R = 2$, $P = 3/2$ What is $P^R \times (1-P)^{N-R}$?

16-(S,F,J): One-tenth of the peanuts in a barrel are rotten. If you take five peanuts, what is the probability that the first four are good and the fifth one is rotten?

17-(S,T,J): A bag has red and blue chips with twice as many reds as blues. What is the probability of picking 4 reds in a row and then one blue?

18-(S,L,J): A large jar is filled with red and black marbles. When you take a sample of three marbles the probability of having only red marbles in the sample is 1/64. What proportion of marbles in the jar are red?

19-(S,Q,J): A fair coin is flipped six times with the following outcome: (heads, heads, heads, heads, heads, heads). Is the probability that the next flip will be tails equal to 1/2, greater than 1/2, or less than 1/2?

20-(S,U,J): When a die is rolled four times, what is the probability of the sequence: (success, success, failure, failure)?

21-(F,F,B): $N = 4$, $R = 3$, $P = .20$ What is $P(R|N)$?

22-(F,T,B): $N-R = 2$, $P = R/2$, $N = 3$ What is $P(R|N)$?

23-(F,L,B): $P(R|N) = .25$, $P = 1/4$, $R = 1$ What is N?

24-(F,Q,B): Can $P(R|N)$ be greater than P? Can $P(R|N)$ be less than $P^R \times (1-P)^{N-R}$?

25-(F,U,B): $R = 4$, $N = 5$, $P = R-N$ What is $P(R|N)$?

26-(S,F,B): Define a success as rolling a 1 or a 2 on a die. If the die is rolled five times, what is the probability that there are successes on exactly two of the trials?

27-(S,T,B): A die is rolled six times. What is the probability of evens coming up twice as many times as odds?

28-(S,L,B): In a game you toss a coin four times and you win if the number of heads in the sequence equals a certain number. If your probability of winning is 6/16 what is your winning number?

29-(S,Q,B): Is there a difference between the probability that two dice rolled at once both come up 6 and the probability that one die rolled twice comes up 6 both times?

30-(S,U,B): Suppose that two people out of every nine in a certain town like John Wayne movies. If a sample is taken, what is the probability that two people in the sample like John Wayne movies?

BIBLIOGRAPHY

Ausubel, D. P. Educational psychology: A cognitive approach. New York: Holt, Rinehart & Winston, 1968.

Ausubel, D. P. In defense of verbal learning. Educational Theory, 1961, 15-25.

Bower, G. H. Organizational factors in memory. Cognitive Psychology, 1970, 1, 18-46.

Bracht, G. H. Experimental factors related to aptitude-treatment interactions. Review of Educational Research, 1970, 40, 627-645.

Bruner, J. S. The act of discovery. Harvard Educational Review, 1961, 31, 21-32.

Egan, D. E., & Greeno, J. G. Acquiring cognitive structure by discovery and rule learning. Journal of Educational Psychology, 1973, in press.

Gagné, R. M. Human problem solving: Internal and external events. In B. Kleinmuntz (Ed.), Problem solving: Research, method and theory. New York: Wiley, 1966, 128-148.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1965.

Gagné, R. M., & Smith, E. C. A study of the effects of verbalization on problem solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Greeno, J. G. The structure of memory and the process of solving problems. University of Michigan: Human Performance Center, Technical Report No. 37, 1972.

Greeno, J. G., & Mayer, R. E. Structural and quantitative interaction among aptitudes and instructional treatments. Journal of Educational Psychology, in preparation.

Hermann, G. Learning by discovery: A critical review of studies. Journal of Experimental Education, 1969, 38, 58-72.

Mayer, R. E. Dimensions of learning to solve problems. University of Michigan: Human Performance Center, Memorandum Report No. 15, 1972.

Mayer, R. E., & Greeno, J. G. Structural differences between learning outcomes produced by different instructional methods. Journal of Educational Psychology, 1972, 63, 165-173.

Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.

Piaget, J. Science of education and the psychology of the child. New York: Orion Press, 1970.

Piaget, J. The language and thought of the child. Cleveland: Meridian, 1969.

Rosenthal, R. A. Experimental effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

Rothkopf, E. Z. The concept of mathemagenic activities. Review of Educational Research, 1970, 40, 325-336.

Roughhead, W. G., & Scandura, J. M. 'What is learned' in mathematical discovery. Journal of Educational Psychology, 1968, 59, 283-289.

Scandura, J. M. Learning verbal and symbolic statements of mathematical rules. Journal of Educational Psychology, 1967, 58, 356-364.

Scandura, J. M. Algorithm learning and problem solving. Journal of Experimental Education, 1966, 34, 1-6.

Shulman, L. S., & Keisler, E. R. (Eds.) Learning by discovery. Chicago: Rand-McNally, 1966.

Stiehl, C. C. Personal communication, 1973.

Wittrock, M. C. The learning by discovery hypothesis. In L. S. Shulman and E. R. Keisler (Eds.), Learning by discovery. Chicago: Rand-McNally, 1966, 33-75.

RECENT TECHNICAL REPORTS

24. Adams, R. A. S. Interference in short-term retention of discrete movements. September 1970.
25. Armstrong, T. R. Feedback and perceptual-motor skill learning: A review of information feedback and manual guidance training techniques. August 1970.
26. Armstrong, T. R. Training for the production of memorized movement patterns. August 1970.
27. Collins, A. M. The effect of visual stimulus traces on memory. August 1970.
28. Cohen, V. V. R. Short-term memory for quantitative information from three kinds of visual displays. June 1971.
29. Gelfand, H. Organization in free recall learning: Output contiguity and interresponse times as a function of presentation structure. June 1971.
30. Roberts, K. H. An investigation of paraphrasing: The effects of memory and complexity. June 1971.
31. Thomas, J. C. An analysis of behavior in the Hobbits-Orc problem. August 1971.
32. Bennett, I. F. Spatial effects in visual selective attention. August 1971.
33. Bennett, R. W. Theoretical implications of proactive interference in short-term memory. August 1971.
34. Chananie, J. E. Memory effects in visual search: The graticule. December 1971.
35. Wichawut, C. Encoding variability and the effect of spacing of repetitions in continuous recognition memory. January 1972.
36. Walter, D. A. The role of acoustic and semantic dimensions of memory in sentence memory and comprehension. March 1972.
37. Greeno, J. G. The structure of memory and the process of solving problems. June 1972.
38. Greeno, J. G. & Bjork, R. A. Mathematical learning theory and the new "mental forestry." July 1972.
39. Polzella, D. J. Remembering the functional sentence. August 1972.
40. Malin, J. T. An analysis of strategies for solving certain substitution problems. April 1973.

MEMORANDUM REPORTS AFTER 1968

10. Martin, E. Associative interference theory and spontaneous recovery. November 1969.
12. Greeno, J. G. Theoretical entities in statistical explanation. October 1970.
13. Greeno, J. G. Technical and informal theories about mental entities. October 1970.
14. Halff, H. M. The differential effects of stimulus presentation during error- and success-feedback intervals in concept identification. August 1971.
15. Mayer, R. E. Dimensions of learning to solve problems. February 1972.
16. Martin, E. Serial learning theory. February, 1973.