DOCUMENT RESUME

ED 062 667 CG 007 199

AUTHOR Egan, Dennis E.; Greeno, James G.
TITLE Acquiring Cognitive Structure by Discovery and Rule Learning.
INSTITUTION Michigan Univ., Ann Arbor.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE 70
GRANT OEG-0-9-320447-4194
NOTE 35p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Aptitude; *Cognitive Processes; *Discovery Processes; *Learning; Learning Activities; *Learning Processes; Problem Solving; *Teaching Methods; Thought Processes

ABSTRACT

The study concerned (1) identifying component processes of discovery and rule learning; (2) describing differences in learning outcomes produced by the two instructional methods, and; (3) optimizing learning. It was believed that understanding the effects of aptitude, instructional methods, and their interaction is important in the study of learning and problem solving. Two experiments were performed to investigate the effects of aptitude and instructional methods on learning concepts of probability. The two methods were learning by discovery and learning by rule versions of programmed instruction. Results supported the hypothesis that the outcome of discovery is the structural integration of previously known concepts, while the outcome of rule learning is the addition, subject scoring low on tests of relevant abilities performed better by every measure when instructed by the rule method. The data indicates that the result of learning by discovery is a well integrated cognitive structure because subjects can solve problems that require relating principles previously learned. (BW)

6.5
N -NPU
CG

Acquiring Cognitive Structure by Discovery and Rule Learning

Dennis E. Egan and James G. Greeno

The University of Michigan

Abstract

The study concerned (1) identifying component processes of discovery and rule learning; (2) describing differences in learning outcomes produced by the two instructional methods; and (3) optimizing learning. In two experiments subjects acquired concepts of probability by discovery or rule versions of programmed instruction. Descriptions of learning by discovery and rule were based on reliable aptitude-treatment interactions involving several problem solving skills. Results also supported the hypothesis that the outcome of discovery is the structural integration of previously known concepts, while the outcome of rule learning is the addition of new structure. Finally, subjects scoring low on tests of relevant abilities performed better by every measure when instructed by the rule method.

Acquiring Cognitive Structure by Discovery and Rule Learning[1]

Dennis E. Egan and James G. Greeno

The University of Michigan

Understanding the effects of aptitude, instructional method, and their interaction (the aptitude-treatment interaction or ATI) is important in the study of learning and problem solving for at least three reasons. First, a thorough understanding of these effects may make it possible to assign Ss of differing ability to optimal instructional methods (Cronbach, 1967). Second, the process of acquiring cognitive structure can be analyzed in terms of the skills that are more or less relevant to success under different instructional methods. In this case, aptitude becomes a theoretical process variable (Melton, 1967). Third, the characteristics of cognitive structure acquired by different instructional groups can be inferred from group differences in terminal performance (Mayer & Greeno, in press).

Two experiments were performed to investigate the effects of aptitude and instructional method on learning concepts of probability.

Experiment I

Learning by discovery and learning by rule are contrasting instructional methods that appear important for applications and promising for analysis of process and structural distinctions. These methods have, in one form or another, been the focus of much research (Ausubel, 1961; Bruner, 1961; Corman, 1957; Gagné & Brown, 1961; Guthrie, 1968; Kittel, 1957; Shulman, 1970; Tallmadge, 1968; Wittrock, 1963). While studies have come to contradictory conclusions about the superiority of a discovery-type or a rule-type instructional method, there appears to be a consensus on the fundamental difference between learning by discovery and learning by rule. Subjects learning by discovery proceed by solving problems and generalizing with very little initial information. The task of the rule learner is to interpret initial information and apply it to problems. Other differences between the methods are probably not as essential.

A simple hypothesis suggests that skills involved in solving problems and generalizing are more important to success in learning by discovery than in learning by rule. This idea leads to the expectation of an ATI such that the skills of Ss learning by discovery should be strongly related to their performance while the skills of Ss learning by rule should be less strongly related to performance.

Available evidence appears to discredit this hypothesis. Tallmadge (1968) and Corman (1957) found no reliable ATI for groups of varying ability learning by a discovery-type or a rule-type method. These studies used scores on tests of general ability as measures of aptitude. Recently Bracht (1970) surveyed ATI literature and reported that a disordinal ATI is more likely to be found if the tests of ability are specific to the learning task. Thus, the lack of evidence may be due to the use of tests of general ability. Moreover, an ATI found with a general aptitude would yield very little information about the processes of learning. The first experiment was performed in an attempt to achieve reliable ATIs in the expected direction, as well as to analyze the processes involved in learning by discovery and learning by rule.

Method

Materials -- Subjects were taught how to solve problems involving binomial probability by one of two different programmed texts. The texts were constructed by parsing an instructional binomial problem into a hierarchy of components. This instructional problem required finding the probability of three successes in five trials of rolling a die. Subjects advanced through the text by solving multiple choice problems concerning each component of the problem. The sequence is presented schematically in Fig. 1 where components are represented by their symbols in the formula. A correct answer allowed S to bypass lower level instruction on that particular component (Campbell, 1963), while an incorrect answer sent S into a remedial loop. Once the entire instructional problem was solved, S had to successively solve three criterion problems that changed the values of the instructional problem.

Subjects learning by rule were given the binomial formula and relevant definitions on the first page of the text. Thereafter, all questions and instruction were phrased in terms of the formula. Subjects learning by discovery were asked the same questions at each stage of the hierarchy as Ss learning by rule. However, the questions for the discovery group were phrased in ordinary English, as nontechnically as possible. For example, Ss learning by rule were asked to find the value of $p^r * q^{n-r}$ at the same point in the instructional sequence that Ss learning by

Insert Figure 1 about here

discovery were asked to find the probability of a particular sequence of rolls. Definitions and notation for the variables were introduced to discovery Ss only after they had solved various parts of the instructional problem. Using the notation, Ss generalized their solutions to obtain parts of the formula. Discovery Ss never saw the entire binomial formula at once. Sequencing in the discovery and rule texts was identical.

Ability tests -- Tests of three abilities specific to binomial probability were administered. A test of probabilistic concepts consisted of 14 multiple choice questions concerning identification of the probabilities of single events, joint events, the nonoccurrence of events, the occurrence of either of two events, and the occurrence of simple sequences of events. A second test measured skill in the arithmetic operations necessary for calculating binomial probabilities. Eight problems were given involving computation of factorials, addition of fractions, and exponentiation of fractions. The third test was adapted from Leskow & Smock (1970). Subjects were asked to write out as many of the permutations of the digits 1234 as they could according to a plan that would exhaust all possibilities without repeating any. Scores were based on how closely S approximated one of two strategies: (1) holding initial digits constant and changing digits on the right, or (2) rotating the preceding permutation. The relevance of the first two tests to binomial probability is obvious. With regard to the permutations test, Piaget & Inhelder (1951) have hypothesized that a prerequisite for understanding probability is the ability to deal systematically with a set of possibilities. In discovering probabilistic concepts, the ability to count the elements of an outcome space seems especially important. To obtain measures of general ability, Ss were asked to report their scores on the Mathematical Scholastic Aptitude Test (MSAT).

Procedure -- Subjects were given the pretests and then the programmed texts were handed out at random. When S completed the programmed booklet he was given a 5-min break before beginning the posttest. The posttest consisted of ten binomial questions involving different situations.

Subjects -- A total of 57 Ss (male and female) from the University of Michigan paid subject pool participated in the experiment, 29 in the discovery group and 28 in the rule group. Up to five Ss served in each experimental session.

Measures of Learning -- For each S three measures of learning were obtained: the number of errors made in answering the multiple choice problems in the programmed text, the amount of time taken to complete the instructional sequence correctly, and the proportion of errors made on the posttest.

Results

Scores on the permutations test did not account for a significant portion of variance for any of the three measures of learning. This test was excluded from further analyses. For the remaining three abilities, Ss were divided into three groups approximately equal in size on the basis of each test score.

Of the 57 Ss 43 provided their MSAT scores. The range was 419 to 774. Low scoring (≤ 599; $N_D = 5$, $N_R = 8$), Intermediate (600 to 699; $N_D = 8$, $N_R = 8$), and High scoring (≥ 700; $N_D = 9$, $N_R = 5$) were formed. The first column of Fig. 2 shows the relationship between MSAT scores and the three measures of learning.

Insert Figure 2 about here

Scores for the 57 Ss on the 14 item test of probabilistic concepts yielded a range of 5 to 14 correct. Low scoring (≤ 10 correct; $N_D = 10$, $N_R = 6$), Intermediate (11 or 12 correct; $N_D = 8$, $N_R = 10$), and High scoring (13 or 14 correct; $N_D = 11$, $N_R = 12$) groups were formed. The middle column of Fig. 2 shows the results of the concepts grouping for all Ss.

Arithmetic operations scores ranged from 0 to 8. The sample was divided into Low scoring (≤ 4 correct; $N_D = 6$, $N_R = 8$), Intermediate (5 to 7 correct; $N_D = 11$, $N_R = 7$), and High scoring (8 correct; $N_D = 12$, $N_R = 13$) groups. The third column of Fig. 2 shows the results when skill with arithmetic operations was used as the ability criterion.

Table 1 gives the results of analyses of variance for the various combinations of ability criteria and measures of learning.

Insert Table 1 about here

Discussion

Several sets of findings are of psychological interest. First, consider overall differences due to instructional method. Subjects committed more errors in learning by discovery than in learning by rule. This difference is a straightforward result of the difference in methods, since the discovery method required Ss to first solve problems then infer principles from the problems. However, there was not a reliable difference between the two methods in time spent in learning. This finding suggests that there was not a substantial difference in the overall difficulty of the two teaching programs. The lack of a main effect due to method on the posttest suggests that there was no reliable difference in the effectiveness of instruction.

The differences among ability groups for all analyses were highly significant ($p < .01$). In every case, the groups scoring higher on the test of ability performed better on the measures of learning. Thus the tests of concepts and arithmetic operations as well as the MSAT measured characteristics relevant to the learning task.

The main point of the experiment was to test the hypothesis that skills involved in solving problems and generalizing are more important to success in learning by discovery than in learning by rule. Reliable ATIs were obtained in seven of the nine analyses, all in the expected direction. Thus the hypothesis was supported.

Specifically, from the graphs of errors in learning in Fig. 2, it is apparent that all three groups of Ss learning by rule made few errors, but groups of Ss learning by discovery were systematically ordered. The abler discovery Ss made fewest errors while the intermediate and low ability groups made progressively more errors. The same general pattern of results was obtained in analyses of time spent in learning.

Finally, consider the ATI on the posttest. Consistent with Corman (1957) and Tallmadge (1968), there was no evidence of an interaction between instructional method and general ability as measured by the MSAT. However, interactions were found between the methods used and the tests that measured abilities specifically involved in the learning task. The effect was at least marginally significant for both the test of concepts and the arithmetic test.

Knowledge of probabilistic concepts and arithmetic operations was more important to success in learning by this version of discovery than this version of rule. To that extent there is some clue as to the difference between the process of learning by discovery and the process of learning by rule. If acquisition of concepts by discovery involves more problem solving and generalizing activity than does learning by rule, it would be expected that the learning outcomes produced by the two methods might differ. Since the set of problems on the posttest was not generated in any systematic fashion, little can be said concerning the characteristics of the cognitive structure produced by each method of instruction.

A second experiment was performed to replicate the obtained ATIs and to extend understanding of what is acquired under each type of instruction by means of a systematic transfer analysis.

Experiment II

Katona (19[illegible]) found that meaningful learning allows Ss to solve problems in a variety of circumstances. If Ss discovered the principle of solving a set of problems, they performed better on tests of long-term retention and transfer than Ss who had memorized and practiced a rule for solving the problems. On the other hand, when tested immediately on problems very similar to the instructional materials, Ss who had learned by memorizing and drill performed better.

Other reported differences in retention and transfer between Ss learning by discovery or learning by rule have been inconsistent (e.g., Kittel, 1957; Guthrie, 1968; Wittrock, 1963). The diversity of results is probably due in part to the diversity of instructional materials and instructional methods.

In one study that used instructional materials and methods similar to those in the present study, Gagné and Brown (1961) gave three groups of Ss programmed instruction in the summation of algebraic series. The groups of interest were the rule-example group and the guided discovery groups which roughly correspond to the rule and discovery groups in the present study. While all three instructional methods produced savings in time spent in relearning (a measure of retention), the guided discovery group showed the highest proficiency in solving problems on a posttest (a measure of transfer).

Results of Experiment I indicated that there was no overall difference between the discovery and rule groups in number of problems solved on the posttest. Since a rather haphazard selection of problems was used, the discovery method might have produced better performance on some types of problems with the rule method producing better performance on other types of problems.

How might instructional method affect performance on various types of problems? The answer depends on the characteristics of the cognitive structure produced by each instructional method. One hypothesis is that the problem solving and generalizing activity required of Ss learning by discovery produces greater integration of new information into existing cognitive structure. Because Ss learning by discovery think about and solve problems before being given an algorithm, they understand the material in a more meaningful way (Katona, 1940) than Ss learning by rule. Subjects learning by discovery thus acquire new structural links between concepts already known, rather than first representing concepts by notation and then memorizing relations among coded variables.

If this hypothesis were true, then the difference in performance between fairly direct problems and problems requiring interpretation (in the sense of relating what was known previously to the principle recently learned) should be greater for Ss learning by rule than for Ss learning by discovery. Specifically, on posttest problems that are posed in terms of components of the formula, performance of Ss learning by rule should be relatively better than on word problems because word problems require more interpretation. Moreover, on problems on the posttest that can be solved by directly applying the rule, Ss learning by rule should perform relatively better than on problems that must first be transformed to apply the rule, or that cannot be solved by using the rule. If the structure acquired by Ss learning by discovery is well integrated then the performance of those Ss on a posttest should be less affected by changes in the amount of interpretation necessary.

Method

Materials -- Subjects were taught how to solve problems involving joint probability (e.g., finding the probability of a particular sequence of successes and failures) by means of programmed instruction similar to the first half of the texts used in Experiment I. The instructional procedures differed from those in the first experiment in several important ways. First, a Computer Assisted Instruction (CAI) system was used instead of a programmed text. Subjects sat in booths equipped with keyboards and display screens and responded to questions by typing in answers. Second, Ss had to calculate and enter numerical answers rather than choose among a set of possible responses. Third, at all times Ss had several options available. Subjects could always return to a frame that summarized the instructional problem; they could at any time get out of an instructional loop and attempt to solve the instructional problem; they could use a programmed arithmetic calculator for any difficult computations. Additionally, Ss learning by rule could return to a frame defining all the variables at any time. Finally, Ss learning by discovery were not exposed to the formula or definitions until the second day of the experiment.

Ability tests -- Tests were again given in conceptual, arithmetic, and permutation skills, but each test was modified somewhat from the first experiment. The test of probabilistic concepts consisted of eight questions concerning identification of the probability of single events, occurrence of either of two events, occurrence of joint events, and nonoccurrence of events. The test of arithmetic operations consisted of eight problems involving addition, subtraction, multiplication, and exponentiation of fractions. The permutation task was changed so that after S typed in a permutation, his display screen was erased, leaving only the last acceptable permutation he wrote. This procedure is more similar to that used by Leskow & Smock (1970). Permutations were scored for the strategy of holding digits constant from the left. MSAT scores were again obtained as measures of general mathematical ability.

Procedure -- On the first day of the experiment, Ss were randomly assigned to the discovery or rule group. They then received instruction in the use of the CAI equipment, and were given the ability tests followed by the instructional problem which concerned finding the probability of a particular sequence of successes and failures in rolling a die. Subjects returned 24 hours later and again had to solve the instructional problem. Scores on solving the instructional problem were used to measure retention. Following the instructional problem, all Ss had to write out the formula for joint probability, $p^r \cdot q^{n-r}$, once correctly. For Ss learning by discovery, this task required inferring the formula from their solution of the instructional problem. For Ss learning by rule, the task simply required giving the formula from memory as it had already been presented. Once Ss wrote out the formula correctly, they went on to the set of criterion problems. The posttest immediately followed the last criterion problem.

Transfer Design -- The posttest consisted of 18 problems, three of each of six types in a 2x3 design. The first factor was problem-context. Half the problems were word problems, half were posed in terms of the components of the formula. The second factor was problem-type and involved the amount of transformation necessary before the joint probability formula could be applied. Familiar problems were similar to the instructional and criterion problems in that all values necessary to solving the joint probability formula were explicitly stated and the formula could be directly applied to obtain a solution. Transformed problems did not state all values of the formula explicitly. Instead, the S was required to obtain some of them by simple calculation. The third type of problem was called a Luchins problem (Luchins, 1942). These problems had very direct solutions, but were not solvable by direct application of the rule learned. An example of each of the six types of problems is given in Table 2. The problems were randomized at the start of each session.

Insert Table 2 about here

Subjects --- A total of 72 Ss (male and female) from the University of Michigan paid subject pool participated in the experiment, 36 in each instructional group. The CAI system was set up to handle up to five Ss in a single session.

Measures of Learning --- For each S separate scores were obtained for errors made on questions in the programmed instruction and time spent in learning on each part of the instructional sequence. These scores were later summed to yield overall measures of errors and time in learning. For problems on the posttest, the overall proportion of errors made and the time spent in solving each problem were obtained for each S.

Results

Analysis of the relearning concerned comparing the errors and time to solve the instructional problem on the first and second day. Table 3 shows that Ss learning either by discovery or by rule solved the instructional problem on the second day in less time and with fewer errors than on the first day. Since so few Ss made any errors at all on the second presenta-

---

Insert Table 3 about here

---

tion of the instructional problem, the partial errors and time scores were not analyzed for effects of ability. Instead, scores on the instructional problem for the first and second days were combined with errors and time taken to give the formula and solve the criterion problems. These summed scores of time and errors were used in all further analyses of learning.

Scores on the test of arithmetic operations were not strongly related to any of the measures of learning. The test was excluded from further analyses. On the basis of each of the remaining three abilities, Ss were divided into three groups of approximately equal size.

Of the 72 Ss, 65 provided their MSAT score. The range was 450 to 800. Low scoring ($\leq$ 599; N_D = 10, N_R = 10), Intermediate (600 to 699; N_D = 12, N_R = 15), and High scoring ($\geq$ 700; N_D = 11, N_R = 7) groups were formed. The first column of Fig. 3 shows the relationship between MSAT scores and three measures of learning (overall errors, overall time in learning, proportion of errors on the posttest).

Scores on the test of probabilistic concepts yielded a range of 0 to 8 correct. Subjects were grouped into Low (0 to 5; N_D = 11, N_R = 9), Intermediate (6 or 7; N_D = 17, N_R = 19), and High scoring groups (8 correct; N_D = 8, N_R = 8). The middle column of Fig. 3 shows the results of grouping by scores on the test of concepts.

Insert Figure 3 about here

Scoring for the strategy of generating permutations by the number of digits held constant from the left gave a range of 1 to 32, the maximum score possible. Groups of Low ($\leq$ 11; N_D = 12, N_R = 14), Intermediate (11 to 29; N_D = 12, N_R = 10), and High (30 to 32; N_D = 12, N_R = 12) ability were formed. Results are presented in the last column of Fig. 3. Table 4 summarized the analyses of variance for all combinations of ability, instructional method and measure of learning.

Insert Table 4 about here

Performance on the different kinds of posttest problems of Ss in the two conditions is graphed in Fig. 4. Data from the posttest were analyzed by means of a 2X3X2X3 analysis of variance for each ability grouping.

Instructional method and aptitude level were between subject variables, and those results are incorporated in Fig. 3 and Table 4. Problem-context and

Insert Figure 4 about here

problem-type were within-subject variables. As analyses of the posttest data for all three abilities followed the same general pattern, a weighting system was devised so that each score (concepts, permutations, MSAT) contributed about equally to the variance of a weighted abilities score. Weighted Score = Concepts Score + $\frac{\text{Permutation Score}}{6}$ + $\frac{\text{MSAT}}{44}$. The full analysis based on the weighted abilities score is given in Table 5.

Insert Table 5 about here

Discussion

One goal of studying aptitude and instructional vaiiables is to be able to assign Ss of varying ability to optimal instructional methods. The present results suggest that Ss lacking in skills necessary to solve problems may learn more efficiently when instructed by techniques requiring interpretation and application of a rule. By every measure, Ss low in relevant abilities performed better when instructed by the rule method. That the rule method used in this study was not inherently better can be inferred from two results found in Experiment I and replicated in Experiment II. First, while Ss learning by discovery did generally make more errors on the teaching program, they still managed to learn the material in about the same amount of time as Ss learning by rule. Results in Table 3 indicate that Ss learning by discovery did not made more errors or take as much time as Ss learning by rule in solving the instructional problem. The extra time and errors were incurred when discovery Ss had to infer the formula and their solutions and apply it to the criterion problems. Second, there was little difference between instructional groups in overall performance on the posttest. The apparent method main effect in the analyses in Tables 4 and 5 was largely due to the simple effect of method for low-ability Ss.

A second goal of the present study was to describe the differences in the process of acquiring cognitive structure by discovery and rule. The fact that real differences exist was supported again in Experiment II where reliable ATIs were obtained in six of the nine tests, all in the expected direction. In Experiment I the discovery method required the availability of relevant probabilistic concepts and computational skills to a greater degree than the rule method. In Experiment II where Ss were given arithmetic calculators, computational skill was unrelated to performance, but the discovery method required conceptual ability and the ability to solve problems in a systematic way to a greater degree than the rule method.

Analysis of the differences in the process of acquiring cognitive structure might begin by identifying the component processes involved in learning under each method. First, consider the rule method. To solve parts of the instructional problem, a subject might carry out the following steps, not necessarily in a serial fashion.

1. Read the problem text.

2. Select information from the text pertaining to the values of relevant variables, and co-ordinate this information to the coded representations of variables in memory. For example, from the phrase, "the chances of success were 1/4," Ss could extract information in the form, "p = .25".

3. Select a rule or formula for using the variables whose values have been taken from the text. This might be looked up in available information, or retrieved from memory.

4. Perform any transformations needed to make the rule applicable to the information.

5. Calculate the answer.

Since the learning by rule did not greatly involve conceptual and other skills, individual differences in these skills were not associated with differences in performance. On the other hand, a measure of working memory, for example the ability to memorize, transform and apply formulas, might be related to success in learning by rule.

Now consider the discovery method. In the discovery method, Ss had to solve the instructional problem without first being given an algorithm. A discovery S might carry out the following steps:

1. Read the problem text.

2. Interpret the information in the problem in relation to concepts that are understood. The discovery method did not provide a well-specified list of variables as did the rule method. Therefore, interpretation of information in the discovery method probably had more

of the properties of understanding a sentence than in the rule method, and less of the character of filling in values of variables in a list.

3. Search for or systematically generate relationships among concepts used in the problem, particularly relationships that seem to move in the direction of relating the given information with the unknown. This is the kind of process that has been investigated in classical studies of problem solving such as those of Duncker (1945), Polya (1955), and Wertheimer (1959). Subjects might find relationships that involved their understanding of the concepts in the problem, or they might apply a more general relational structure that fit the needs of the problem, or they might find a set of concepts in memory whose relationships seem to provide an analogy to the situation in the problem.

4. Carry out any calculations needed to obtain the answer. This process may well entail a great deal of computational ability, since no algorithm is present to relate specified variables and operations in a compact way.

Since the process of learning by discovery requires conceptual, systematizing and other skills, individual differences in these skills led to similar differences in performance.

Given these distinctions in the process of acquisition, it follows that there are differences in the learning outcomes of the two instructional groups. The results pertinent to this question involve the interactions of method and the two transfer variables appearing in Table 5 and graphed in Fig. 4. Both two-way interactions involving instructional method and transfer were at least marginally significant. While the overall performance of Ss learning by discovery is depressed because of the low ability group, Ss learning by rule showed a much greater decrement in performance on problems requiring more interpretation. The difference between percentage of formula and word problems solved was 13% for the rule group compared to 3% for the discovery group. Differences between Familiar and Luchins problems solved correctly were 22% for the rule group and 9% for the discovery group. These trends were present at all ability levels, although the curves for the two instructional methods crossed only in the high and intermediate ability groups. The average time taken to solve the six types of problems, given a correct solution, was also computed for each instructional group. These results are difficult to analyze because of missing data, but in general show the same method-transfer interactions.

These data indicate that the result of learning by discovery is a well integrated cognitive structure. Subjects can solve problems that require relating what they knew previously to the principle learned about as well as problems that require direct application of the principle. This feature of cognitive structure has been termed "external connectedness" and was found to be characteristic of Ss who learned about binomial probability under instruction emphasizing general concepts rather than a formula (Mayer & Greeno, in press). Thus there is some support for the claim (Gagné, 1965) that meaningful conceptual learning and the discovery and generalization of a principle result in about the same outcome.

The result of learning by rule is primarily the addition of new components to cognitive structure rather than the reorganization of existing components. These new components include a list of defined variables and the sequence of operations relating them. The new components may in fact have a great degree of "internal connectedness" as shown by the advantage of Ss learning by rule on Familiar problems and problems posed in the context of the formula. However, the fact that the advantage is lost when the problems require more interpretation shows that the new structural components added by rule Ss were not well integrated into existing cognitive structure. A test of long-term retention should, if this explanation is correct, show that the discovery Ss retained more information. The test of relearning after 24 hours used in the present study merely demonstrated that neither group had forgotten much instruction during that time.

A final set of conclusions concern procedures involved in studying aptitude and instructional variables. With regard to aptitude tests, a choice was obviously made in the present study for simplicity over psychometric elegance. One valid criticism is that the unreliability of the measuring instruments may have influenced the results. It is not known, for example, whether the failure of the test of arithmetic operations in the second experiment was due to allowing Ss to use calculators or the unreliability of the test. However, the degree of replication that was found between the two experiments regarding the concepts test makes this possibility less likely. The usefulness of a general ability criterion in studies of ATI is still in question. The fact that the general ability measure produced a reliable ATI on the posttest in the second but not in the first experiment suggests that its utility may be linked to the instructional material. In any case there is a tradeoff between the reliability offered by established tests of general ability, and the information concerning processes of acquisition afforded by tests specially constructed for experimental materials and instructional methods.

An unexpected result was the significant two-way interaction of ability and problem-type, and the three-way interaction of ability, problem-type and problem-context found in the analysis using the weighted average of ability test scores. Graphing these data revealed that the weighted score was most strongly related to performance on Luchins problems, particularly when posed in a formula context. Thus the weighted average of abilities was a particularly strong measure of how easily Ss could manipulate the newly learned components of the formula independently of the rule usually relating them.

References

Ausubel, D. P., In defense of verbal learning. Educational Theory, 1961, 11, 15-25.

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.

Campbell, V. N., Bypassing as a way of adapting self-instruction programs to individual differences. Journal of Educational Psychology, 1963, 54, 337-345.

Corman, B. R., The effect of varying amounts and kinds of information as guidance in problem solving. Psychological Monographs, 1957, 71, no. 2 (Whole no. 431).

Cronbach, L. J., How can instruction be adapted to individual differences? In Gagné, R. M. (ed.) Learning and Individual Differences. Columbus: Merrill, 1967.

Duncker, K. On problem solving. Psychological Monographs, 1945, 58, no. 5 (Whole no. 270).

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

Gagné, R. M., and Brown, L. T., Some factors in the programming of conceptual learning. Journal of Experimental Psychology, 1961, 62, 313-321.

Guthrie, J. T., Expository instruction versus a discovery method. Journal of Educational Psychology, 1967, 58, 45-49.

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

Kittel, J. E., An experimental study of the effect of external direction during learning on transfer and retention of principles. Journal of Educational Psychology, 1957, 48, 391-405.

Leskow, S., and Smock, C. D. Developmental changes in problem solving strategies: Permutation. Developmental Psychology, 1970, 2, 412-422.

Luchins, A. S., Mechanization in problem solving: the effect of Einstellung. Psychological Monographs, 1942, 54, no. 6 (Whole no. 248).

Mayer, R., and Greeno, J. G. Structural differences between learning outcomes produced by different instructional procedures. Journal of Educational Psychology, in press.

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

Piaget, J. and Inhelder, B. La genese de l'idee de hasard chez l'enfant. Paris: Presses Universitaires de France, 1951.

Polya, G. Mathematical Discovery. New York: John Wiley and Sons, 1965.

Shulman, L. S. Psychology and mathematics education. In E. G. Begle (Ed.) Mathematics education. The sixty-ninth yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1970, pp. 23-71.

Tallmadge, G. K. Relationships between training methods and learner characteristics. Journal of Educational Psychology, 1968, 59, 32-3[illegible].

Wertheimer, M. Productive thinking. New York: Harper & Row, 1959.

Wittrock, M. C. Verbal stimuli in concept formation: learning by discovery. Journal of Educational Psychology, 1963, 54, 183-190.

Table 1

Test Statistics in Analyses of Variance for Experiment I

Measure of Learning	Test of Ability	Ability Main Effect	Method Main Effect	Interaction Effect (ATI)
Errors in Learning	Arithmetic	$F(2,51)=9.46$***	$F(1,51)=21.98$***	$F(2,51)=11.89$***
	Concepts	$F(2,51)=6.37$***	$F(1,51)=14.77$***	$F(2,51)=1.33$
	MSAT	$F(2,37)=8.27$***	$F(1,37)=10.57$***	$F(2,37)=6.65$***
Time in Learning	Arithmetic	$F(2,51)=7.87$***	$F(1,51)=1.59$	$F(2,51)=3.79$**
	Concept	$F(2,51)=7.18$***	$F(1,51)=1.59$	$F(2,51)=4.37$***
	MSAT	$F(2,37)=19.99$***	$F(2,37) < 1.00$	$F(2,37)=5.44$***
Errors on Posttest	Arithmetic	$F(2,51)=6.97$***	$F(1,51) < 1.00$	$F(2,51)=3.12$***
	Concept	$F(2,51)=6.89$***	$F(1,51) < 1.00$	$F(2,51)=3.72$**
	MSAT	$F(2,37)=5.56$***	$F(1,37) < 1.00$	$F(2,37) < 1.00$

***$p < .01$

**$p < .05$

*$.10 > p > .05$

Table 2

Examples of the Six Types of Questions Used in Experiment II

Word Questions

Familiar: A die has five spots on one of its six sides, and other numbers on the other sides. If you roll it ten times, what is the probability of getting three fives followed by seven other numbers?

Transformed: If you bet on 2 of 38 numbers in a game of roulette, you win only if one of those numbers is rolled. If you make such a bet, what is the probability of winning on the first two rolls and losing on the next three?

Luchins: You play a game five times in which the probability of winning each time is .17, and the probability of winning three games out of five is .32. What is the total number of successes plus the total number of failures?

Formula Questions

Familiar: R=2, N-R=4, P=1/5, Q=4/5. What is the joint probability?

Transformed: N=7, R=2, P=.31. What is the joint probability?

Luchins: Joint Probability = 15/128, N=5, P=.25, Q=.75. What is the value of R + (N-R)?

Table 3

Comparison of Mean Number of Errors and Time to Solve Instructional Problem on First and Second Day

Group	Measure	First Day	Second Day	F
Discovery	Errors	3.6	0.6	14.92***
Rule	Errors	3.1	0.3	18.54***
Discovery	Time (min)	6.9	1.8	25.51***
Rule	Time (min)	11.6	2.7	61.28***

Table 4

Tests Statistics in Analyses of Variance For Experiment II

Measure of Learning	Test of Ability	Ability Main Effect	Method Main Effect	Interaction Effect (ATI)
Errors in Learning	Permut.	F(2,66)=10.90***	F(1,66)=8.06***	F(2,66)=1.56
	Concept	F(2,66)=15.26***	F(1,66)=9.18***	F(2,66)=3.54**
	MSAT	F(2,59)=7.80***	F(1.59)=6.24**	F(2,59)=6.26***
Time in Learning	Permut.	F(2,66)=8.57***	F(1,66)=1.17	F(2,66) < 1.00
	Concept	F(2,66)=7.69***	F(1,66)=1.15	F(2,66) < 1.00
	MSAT	F(2,59)=5.02***	F(1,59)=1.22	F(2,59)=3.39**
Errors on Posttest	Permut.	F(2,66)=9.83***	F(1,66)=3.76*	F(2,66)=3.23**
	Concept	F(2,66)=17.76***	F(1,66)=4.47**	F(2,66)=4.04**
	MSAT	F(2,59)=8.90***	F(1,59)=1.44	F(2,59)=3.26**

***p < .01

**p < .05

*.10 > p > .05

Table 5

Analysis of Posttest Scores for Weighted Abilities Grouping

Source	SS	df	MS	F
A: Ability	89.43	2	44.72	31.94***
B: Method	7.78	1	7.78	5.56**
A x B	8.92	2	4.46	3.19**
Error (a)	92.30	66	1.40	
C: Problem-Context	5.79	1	5.79	9.98***
D: Problem-Type	19.09	2	9.54	16.45***
A x C	1.03	2	.52	.90
A x D	7.98	4	2.00	3.45**
B x C	3.34	1	3.34	5.76**
B x D	3.34	2	1.67	2.88*
C x D	7.03	2	3.52	6.07***
A x B x C	1.27	2	.64	1.10
A x B x D	.00	4	.00	.00
A x C x D	13.75	4	3.44	5.93***
B x C x D	2.57	2	1.28	2.21
A x B x C x D	3.51	4	.88	1.52
Error (b)	191.97	330	.58	
Total	459.10	431		

***$p < .01$

**$p < .05$

*$.10 > p > .05$

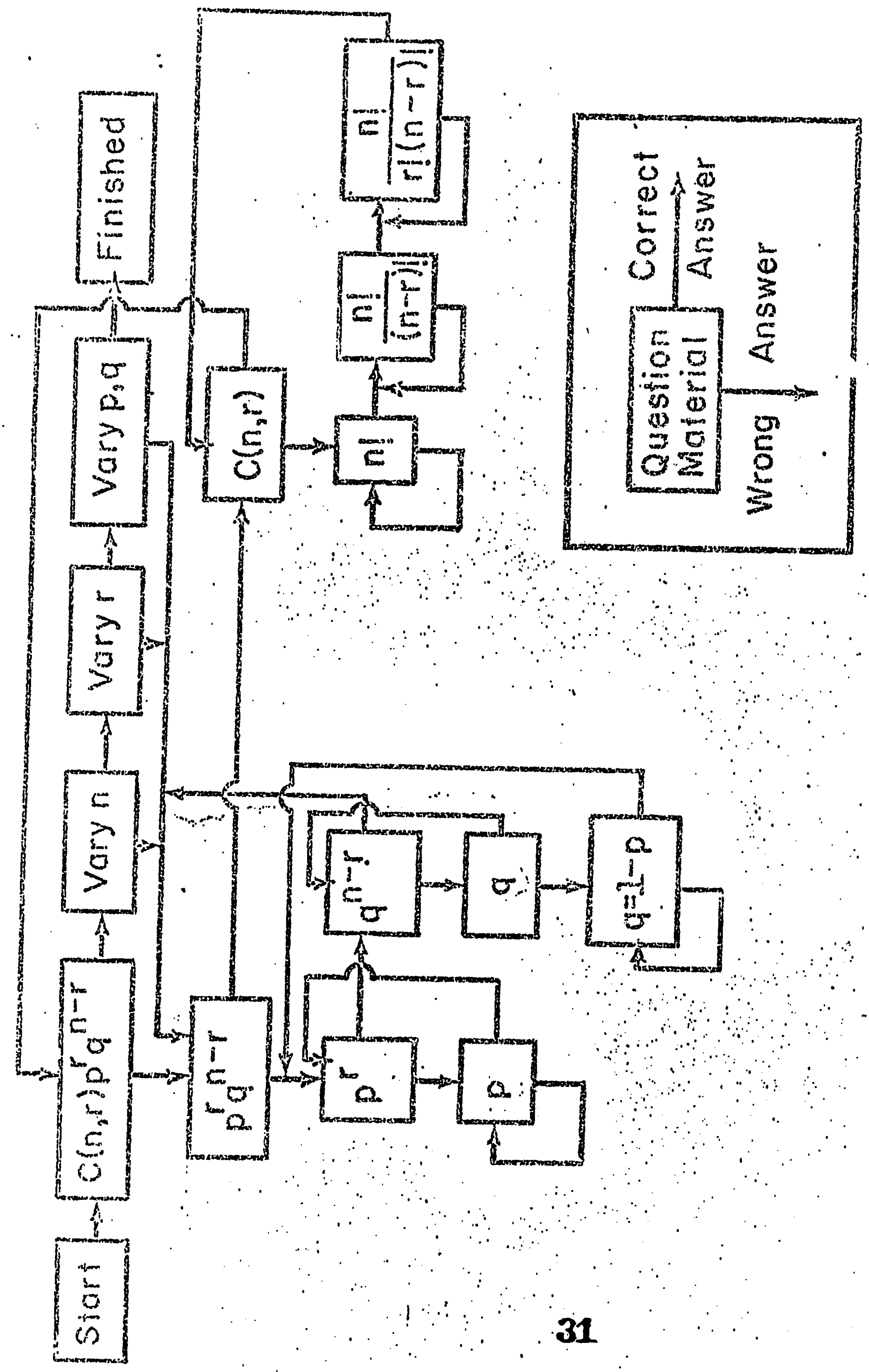
Figure Captions

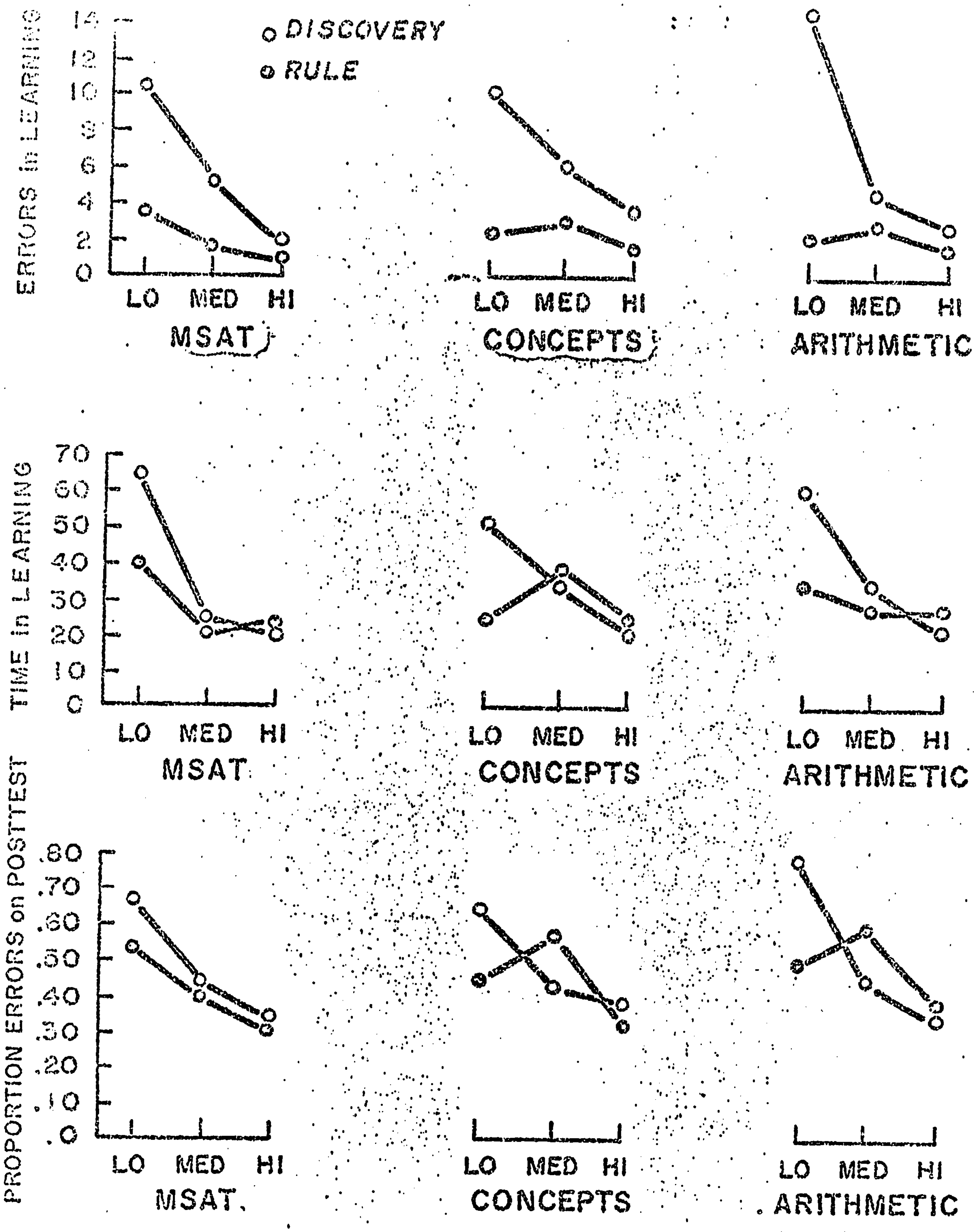
Fig. 1 Schematic representation of instructional sequence.

Fig. 2 Measures of learning as functions of ability grouping in Experiment I.

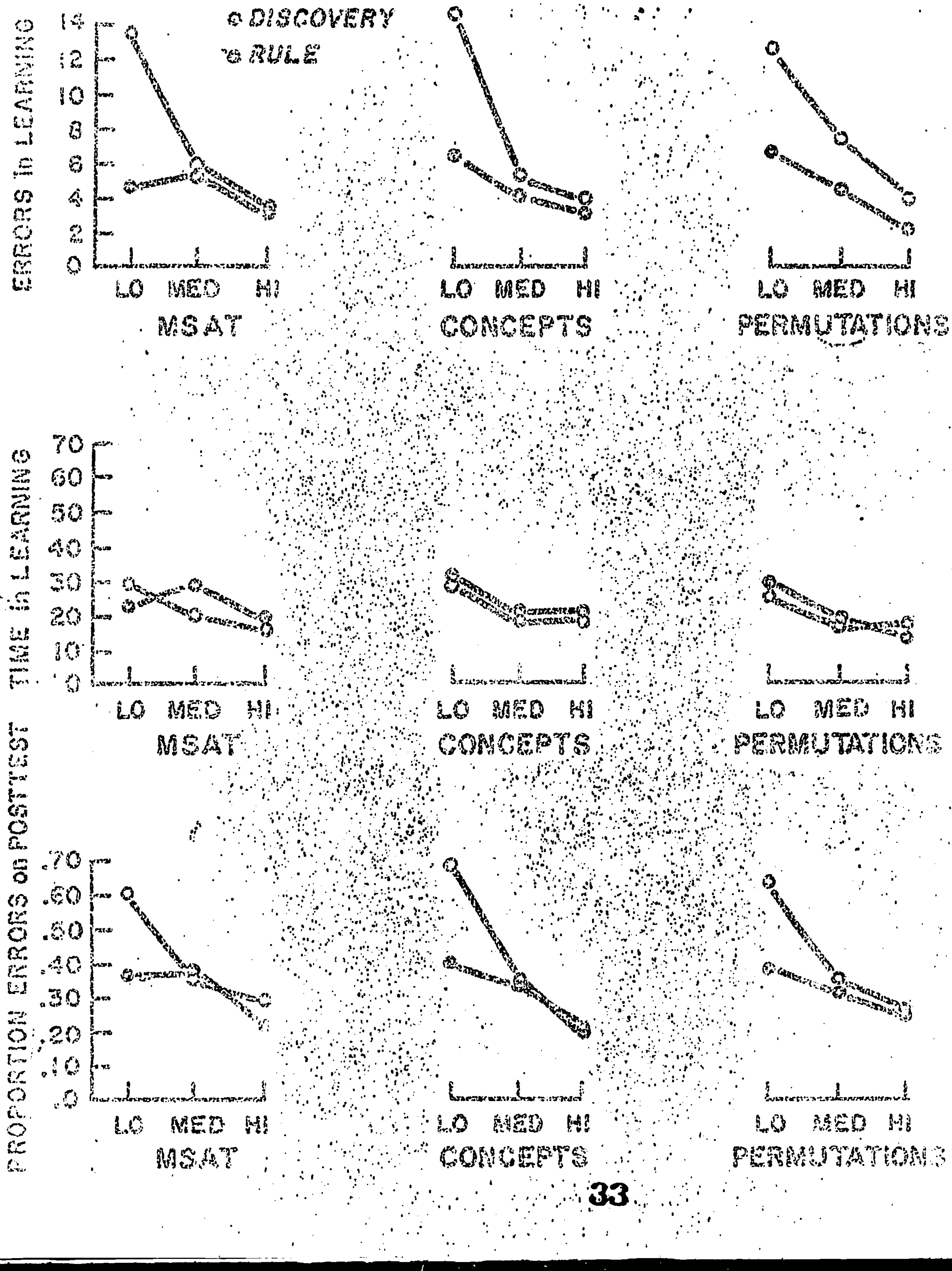
Fig. 3 Measures of learning as functions of ability grouping in Experiment II.

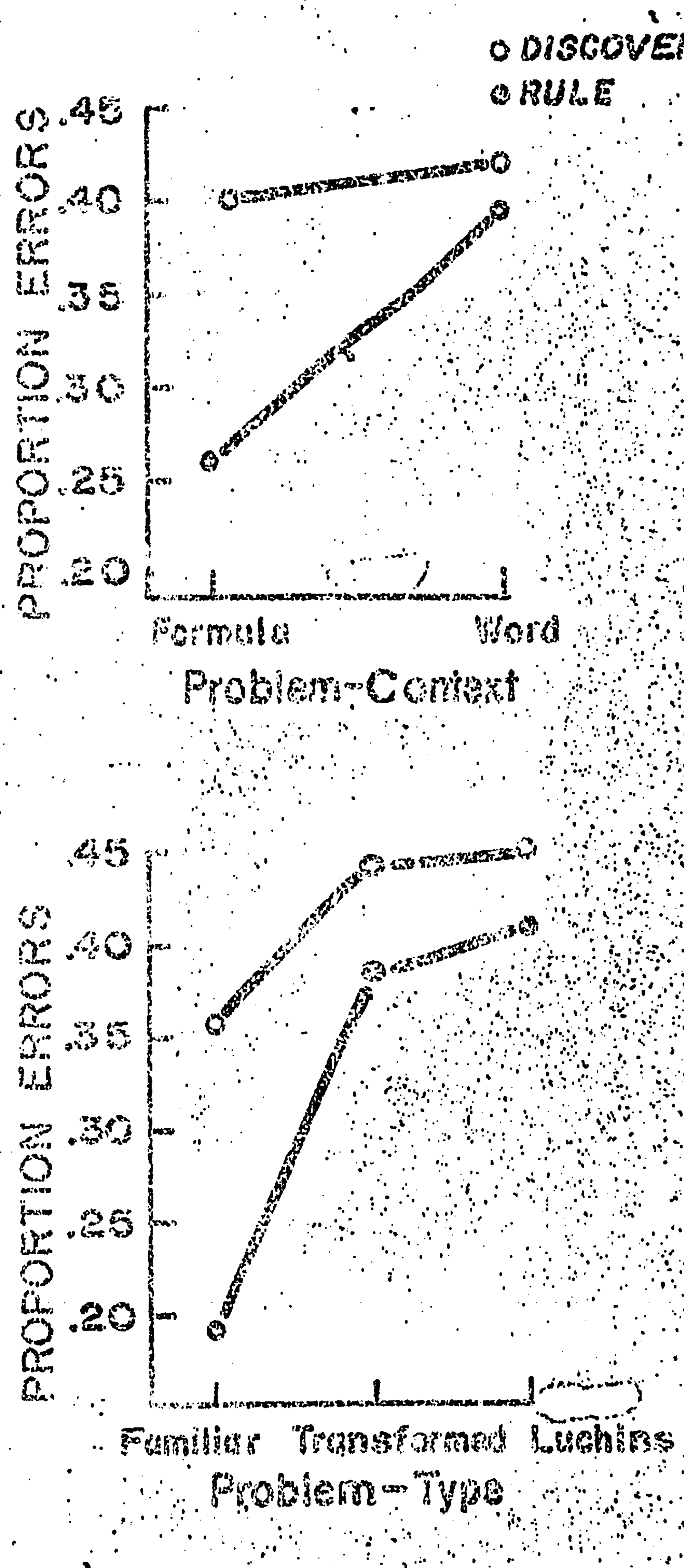
Fig. 4 Plots of method × context interaction (top graph) and method × test item type interaction (lower graph).


Start
C(n,r)p^r q^{n-r}
Vary n
Vary r
Vary p,q
Finished
p^r q^{n-r}
C(n,r)
n!
n!/(n-r)!
n!/(r!(n-r)!)
p^r
p
q^{n-r}
q
q=1-p
Question Material
Correct Answer
Wrong Answer



DISCOVERY
RULE
ERRORS in LEARNING
14
12
10
8
6
4
2
0
LO MED HI
MSAT
CONCEPTS
ARITHMETIC
TIME in LEARNING
70
60
50
40
30
20
10
0
PROPORTION ERRORS on POSTTEST
.80
.70
.60
.50
.40
.30
.20
.10
.0


DISCOVERY
RULE
ERRORS in LEARNING
TIME in LEARNING
PROPORTION ERRORS on POSTTEST
LO MED HI
MSAT
CONCEPTS
PERMUTATIONS



o DISCOVERY
o RULE
PROPORTION ERRORS
.45
.40
.35
.30
.25
.20
Formula
Word
Problem-Context
PROPORTION ERRORS
.45
.40
.35
.30
.25
.20
Familiar
Transformed
Luchins
Problem-Type

Footnote

1. This research was supported in part by the U. S. Office of Education Grant No. OEG-0-9-320447-4194. The research was carried out during the first author's tenure as a National Science Foundation Graduate Fellow. We acknowledge the assistance of Christine Schaffer in conducting the experiments.