EFFECTS OF COGNITIVE STRAIN AND VERBALIZATION ON LEARNING LINEAR AND NONLINEAR RELATIONS

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
STUART O. HALLGREN
1974

ABSTRACT

EFFECTS OF COGNITIVE STRAIN AND VERBALIZATION ON LEARNING LINEAR AND NONLINEAR RELATIONS

By

Stuart O. Hallgren

Sixty introductory psychology students learned either a linear or a nonlinear relation. Ss in the external memory (EM) group had the information from all previous trials available to them. The short-term memory (STM) group only had direct access to the information from the current trial. A selection paradigm was used in order to study the strategies and processes involved in learning the relations. The effect of verbalization during the task was examined by requiring Ss either to think aloud, state reasons for predictions, or remain silent. The experimental design included a 2(rules) X 2(memory) X 3(verbalization) X 2(sex) X 4(blocks of 16 trials) factorial design with repeated measures on the last factor. The number of males and females was the same proportion for each cell.

Three basic dependent variables were used to analyze the results: two learning measures and a focusing

index for determining the use of a focusing strategy. The external memory aided Ss in learning the relations. However, all three dependent measures indicated that the type of rule interacted with the memory conditions as a function of trials. The groups learning the linear rule were most affected by the memory conditions. The linear rule-STM group performed the worst. The nonlinear rule-STM group was able to overcome the memory limitations in learning the rule and compared favorably to the nonlinear rule-EM group by the end of the trials. Contrary to previous results, the linear rule was not easier to learn than the nonlinear rule.

The verbalization conditions interacted with the sex of the subject as a function of trials. Females performed best when thinking aloud while males performed best when learning the rule silently. Also, females had more difficulty in learning the nonlinear relation than the linear relation.

The results of the study indicated that the learning of mathematical relations is more efficient if external memory is used. Sex differences raise methodological questions concerning the effects of verbalization on the process under study, as well as raising the question of whether sex differences exist in learning functional relations.

EFFECTS OF COGNITIVE STRAIN AND

VERBALIZATION ON LEARNING LINEAR

AND NONLINEAR RELATIONS

By

Stuart Oliver Hallgren

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1974

To Kathy

ACKNOWLEDGMENTS

I wish to express my appreciation to the members of my committee: Drs. Raymond Frankmann, Lester Hyman, Donald Johnson, and Gordon Wood. I would like to acknowledge the chairman of the committee, Dr. Johnson, for his suggestions, encouragement and interest.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Research on human judgment seems to have proceeded from the explicit or implicit assumption that there is, somewhere in the judge, a single process of judgment. The characteristics of this process are extracted from an analysis of the relationships between the information available to the judge at the time of the judgment. Thus Anderson (1970) tries to determine if judgment is best characterized by additive or averaging models, and Hoffman (1968) and his associates try to discover if the process is a linear, additive one or configural.

Brehmer (1969) has offered an alternative to this assumption and suggests that the emphasis should be placed on the task rather than the S. If judgments reflect the ways in which Ss have learned to utilize information, then judgments made with respect to a particular task should reflect the characteristics of the task as well as the cognitive processes of the subject.

Using the learning experiment approach, Hammond and Summers (1965) were able to show that humans can learn to utilize nonlinear as well as linear relations for making inferences from a set of cues to a criterion variable.

Their results indicated that the implication from earlier studies of the process of judgment--the implication that humans are limited to a linear use of information--is not generally valid; rather their results showed that humans can learn to utilize nonlinear relations if there are such relations in their task.

In learning to integrate several sources of information to make a judgment, the S is faced with both a memory task and a problem of inference. Bruner, Goodnow, and Austin (1956) suggest that memorizing and organizing information for inferences causes cognitive strain--the load of information processing. Cognitive strain differs according to the difficulty of the inference and the difficulty in remembering previous instances. Learning nonlinear compositional rules is more difficult than learning a linear rule (Hammond and Summers, 1965). Brehmer (1969) has suggested one reason why nonlinear relations are learned less efficiently than linear relations: ". . . a statistician, equipped with all the tools of his trade, would require more information (more trials) to determine adequately the nonlinear aspects of a task than the linear aspects. . . ." This task characteristic of nonlinear relations places a heavy demand on memory of previous trials.

One method of reducing cognitive strain caused by the demands of memory is to make the results of previous trials available to the concept learner. It is a generally accepted empirical conclusion in concept identification (CI) that the availability of previously seen stimulus patterns and their categories facilitates learning (Cahill and Hovland, 1960). The more complex and difficult the conceptual rule, the more pronounced the availability effects (Bourne, Goldstein, and Link, 1964). Because the linear and nonlinear relations used in multiple-cue probability learning (MPL) differ in difficulty, it might similarly be expected that the importance of stimulus availability will increase with rule difficulty.

The subjects in the present experiment had two different types of access to previous information. The external memory (EM) group had the information from all previous trials available to them. In contrast, Ss in the short-term memory (STM) group only had direct access to the information from the current trial (the cue numbers they selected, their estimation, and the correct answer).

It was expected that individual differences exist in utilizing the information provided for previous trials. MPL studies provide the S with numerical information representing the different cues and the subject is then required to make a judgment, also numerical. The linear and nonlinear relations among the cues can be expressed mathematically. Since the task involved abstract mathematical relations, it was expected that mathematical aptitude would be a useful individual difference in predicting performance. Mathematically adept Ss should be better at making inferences that depend on recognizing mathematical relationships.

It was also expected that individual differences exist in performance on a complex learning task. Male students are usually more successful than female students in difficult tasks demanding concentrated effort (Maier and Burke, 1967). Therefore, it was expected that males would do better at the learning task than the females. Equal proportions of males and females were assigned to each condition.

Bruner and his associates have shown that cognitive strain also imposes limitations on the effectiveness of learning strategies. When Ss were required to solve the third of three CI tasks "in their head" with the stimulus array not in view, a focusing strategy was much better than scanning. Focusing yields the relevant attributes of a concept by a process of elimination based on the comparison of each successively encountered instance with a positive instance which is chosen as a focus. The focusing strategy is similar to scientific inquiry where one variable is varied while all other variables are held constant. Focusers were much better than scanners on the conventional selection task, and they performed at the same level as before on the new task, while the scanners showed a

significant increase in the number of trials to solution. Focusers are generally more efficient in CI tasks (Bourne, 1963; Laughlin and Jordan, 1967). Laughlin (1966) found that Ss have a greater tendency to adopt a focusing type of strategy in more complex problems, where the cognitive strain is greater.

The use of learning strategies has generally been ignored in MPL studies. As Bourne, Ekstrand, and Dominowski (1970), pp. 260, 261) have observed about MPL studies: "Unfortunately, to date, there have been relatively few attempts to describe empirically the properties of performance in quantitative concept tasks . . . the experimenters were able to describe the subjects growing reliance on the relevant dimensions with multiple correlational techniques, but were unable to specify in detail any strategies they used in achieving their solution." The present experiment used a selection paradigm (previous MPL studies used the reception paradigm) so that strategies could be studied. In particular the use and effectiveness of a focusing strategy was emphasized.

One of the most direct ways to study strategies and hypotheses used by a subject in an inference task is to ask him to verbalize as he learns the task. Unfortunately, little attention has been paid to the possible effects of verbalization on the task under study. Most of the research on the "effects of verbalization" is concerned with the effect of verbal instruction or with the effects of requiring _S_s to give reasons for what they do or to state general rules abstracted from the task. But the lack of appropriate data does not disprove the usefulness of verbal protocols. Newell and Simon (1972) have based a theoretical system of problem solving upon the verbal protocols of _S_s who were asked to "think aloud." However, while accepting verbalization as useful evidence, Newell and Simon recognize the need to assess the effects of verbalization on strategies and methods used in complex t tasks: "Because of the crucial importance of thinking-aloud behavior for our understanding of problem solving, the latter question deserves further study [p. 475]."

What the _S_ is instructed to verbalize during the task seems to determine the effects of the verbalization. Gagné and Smith (1962) required some _S_s to state a reason for each move when solving the pyramid puzzle. On a transfer problem during which no _S_ was required to verbalize, those who had previously stated reasons performed much better, both in terms of fewer unnecessary moves and faster solution times. At the end of the experiment, all _S_s were asked to state a rule about how such problems should be solved, and their answers were judged for adequacy. _S_s who had been required to verbalize during training generally gave better answers than those who had not verbalized.

Gagné and Smith suggested that Ss who had been required to give reasons for moves were more likely to analyze the problem and try to find "good reasons," and, consequently, were more likely to discover the general principles which could be used for maximally efficient performance. Requiring Ss to verbalize reasons for moves seemed to change the manner in which they worked on the problem. Recently Wilder and Harvey (1971) replicated the experiment with three groups: a control group with no special instructions, a group which was told to verbalize reasons as they solved the problem, and a third group which was told to verbalize reasons covertly. The covert and overt verbalization aided subjects in making fewer overall moves to solve the problem, and the time of solution was the same for all three groups. Also, the effect of verbalization interacted with the difficulty of the problem: the more difficult and complex the problem, the more the verbalization aided the Ss.

Newell and Simon (1972) compared the behavior of five Ss who solved logical problems while thinking aloud with twenty-four Ss who wrote their attempts at a solution on a blackboard. The distribution of the number of steps taken to solve the problem was judged similar for the two groups. Ss working under the thinking-aloud conditions generated much the same logical expressions as did the silent group. Only Ss who generated a certain class of logical expressions solved the problem. Newell and Simon therefore concluded that thinking-aloud instructions did not modify the directions of search for solution. Whether S was vocalizing his thoughts or not had no detectable effect on the paths to solution that were attempted.

Obviously, the issue concerning the effect of verbalization has not been settled. Newell and Simon argue that "thinking-aloud" is basically different from analytical verbalization which requires the subject to analyze his own behavior. The present experiment used three different groups to assess the effects of verbalization: (a) a control group which was silent during the learning task; (b) an "analysis of task group" which was required to state reasons for choosing the cue values during the selection task and also reasons for giving their estimations of the criterion value; and (c) a "thinking-aloud group" which was required to verbalize their thoughts as they attained the quantitative concept. The question under consideration was whether verbalization affects the efficiency or the strategies used in learning the inference task.

METHOD

Subjects

*S*s were 60 (36 male, 24 female) introductory psychology students at Michigan State University. The *S*s were volunteers who received credit for participation. Each *S* was randomly assigned to a treatment condition and the proportion of males and females was the same for each condition.

Apparatus

A selection display contained sixty-four stimulus instances. Three cues were present for each stimulus instance, and the numbers used as the cue values ranged from (1,1,1) to (4,4,4). Each of the three cues was represented by a different color and the numbers were on rectangles 5 cm. wide and 2.5 cm. high. The rectangles had a metal pin through the center so that the cue values could be rotated out of view after *S* had used the cue values. The stimuli were arranged in an eight-by-eight display. The frame of the display was 58 cm. high and 46 cm. wide.

The array was semi-ordered. The value of the first cue divided the board into quartiles. Within each

quartile, the cues were ordered according to the second cue value, and then the third cue value. The ordering was then broken by exchanging several adjacent stimulus instances.

The S interacted with a PDP8/I computer using a teletype. The computer typed the trial number and then typed requests for each of the three cue values. The computer then requested that the subject estimate a criterion value. After this value was typed in, the computer calculated the correct answer and typed the results. In the external memory condition the S had access to all the teletype paper with the previous results. In the short-term memory condition the previous results were covered by a paper shield which rested on top of the carriage.

Learning Tasks

The two tasks required Ss to infer the value of a criterion variable from the values of three cue variables. The tasks differed with respect to the relation between the cue values and the criterion values. One task involved a linear relation: Criterion = 3(Cue 1) + 2(Cue 2) - (Cue 3). The other task involved a nonlinear relation: Criterion = 3(Cue 1) + | 2(Cue 2) - 2(Cue 3)|. Although the answer involved the difference between two cues, it depended on which of the two cues had the greater value.

The range of criterion values for the two tasks was similar.

Procedure

Ss were run individually. Each S was handed a two-page booklet of instructions (see Appendix A) and was asked to read along as E read aloud. They were encouraged to ask questions. The Ss in the thinking-aloud or analytical conditions received instructions about verbalizing during the session and were told that they were being recorded. E showed S the proper sequence of events in making their predictions. These steps included: (a) announcing aloud the cue values and turning over the appropriate rectangle on the display board; (b) entering the cue values into the computer by typing the numbers on the teletype. Hitting the space bar entered the number; (c) entering the prediction and announcing it aloud; and (d) receiving the correct answer from the computer on the teletype. After S had gone through a couple of predictions, E showed them how to correct an entry by hitting the "rub out" key; E stayed in the room until each S was following the procedure correctly and until each S was verbalizing, if necessary.

RESULTS

Four different dependent variables were used in analyzing the data. The dependent variables included two measures of learning for each _S_: (a) the correlation (using Pearson _r_) between the _S_'s estimation and the criterion value, called the achievement score, and (b) the absolute difference between the estimation and the criterion value, called the difference score. A focusing index, assessing the use of a focusing strategy, was obtained by comparing the number of trials where only one cue value was changed with the total number of trials considered. The dependent variables for learning and strategy were analyzed by a 2(rules) X 2(memory) X 3(verbalization) X 2(sex) X 4(blocks of trials) analysis of variance (ANOVA) with repeated measures on the last factor. The total time for completing the learning trials was analyzed in a 2(rules) X 2(memory) X 3(verbalization) X 2(sex) ANOVA.

Achievement Score

Table B1 (see Appendix B) shows the results of the analysis of variance for the achievement score. There was a significant difference between memory conditions with

respect to overall level of achievement, $F(1,36) = 8.08$, $p < .01$. The external memory condition was superior to the short-term memory condition. The blocks effect was significant, $F(3,108) = 20.97$, $p < .01$. There was a rule, memory, and blocks interaction, $F(3,108) = 11.55$, $p < .01$ (see Figure 1). The memory conditions had a greater effect on learning the linear rule then on learning the nonlinear rule. The nonlinear groups converged after four blocks of trials while the linear groups maintained a clear and consistent difference between the external and short-term memory conditions. The rule, sex, and blocks interaction was also significant, $F(3,108) = 3.72$, $p < .05$ (see Figure 2); females found the nonlinear rule more difficult to learn than did their male counterparts. After four blocks of trials, the difference was decreasing.

Difference Score

Table B2 (see Appendix B) shows the results of the analysis of variance for the difference score. There was a significant rule effect, $F(1,36) = 4.85$, $p < .05$, and memory effect, $F(1,36) = 6.19$, $p < .05$, with respect to the overall level of learning. The linear rule was more difficult to learn than the nonlinear rule. Learning was facilitated by the use of an external memory in comparison to the short-term memory condition. The block effect was also significant, $F(3,108) = 27.13$, $p < .01$. There was a rule, memory, and block interaction,

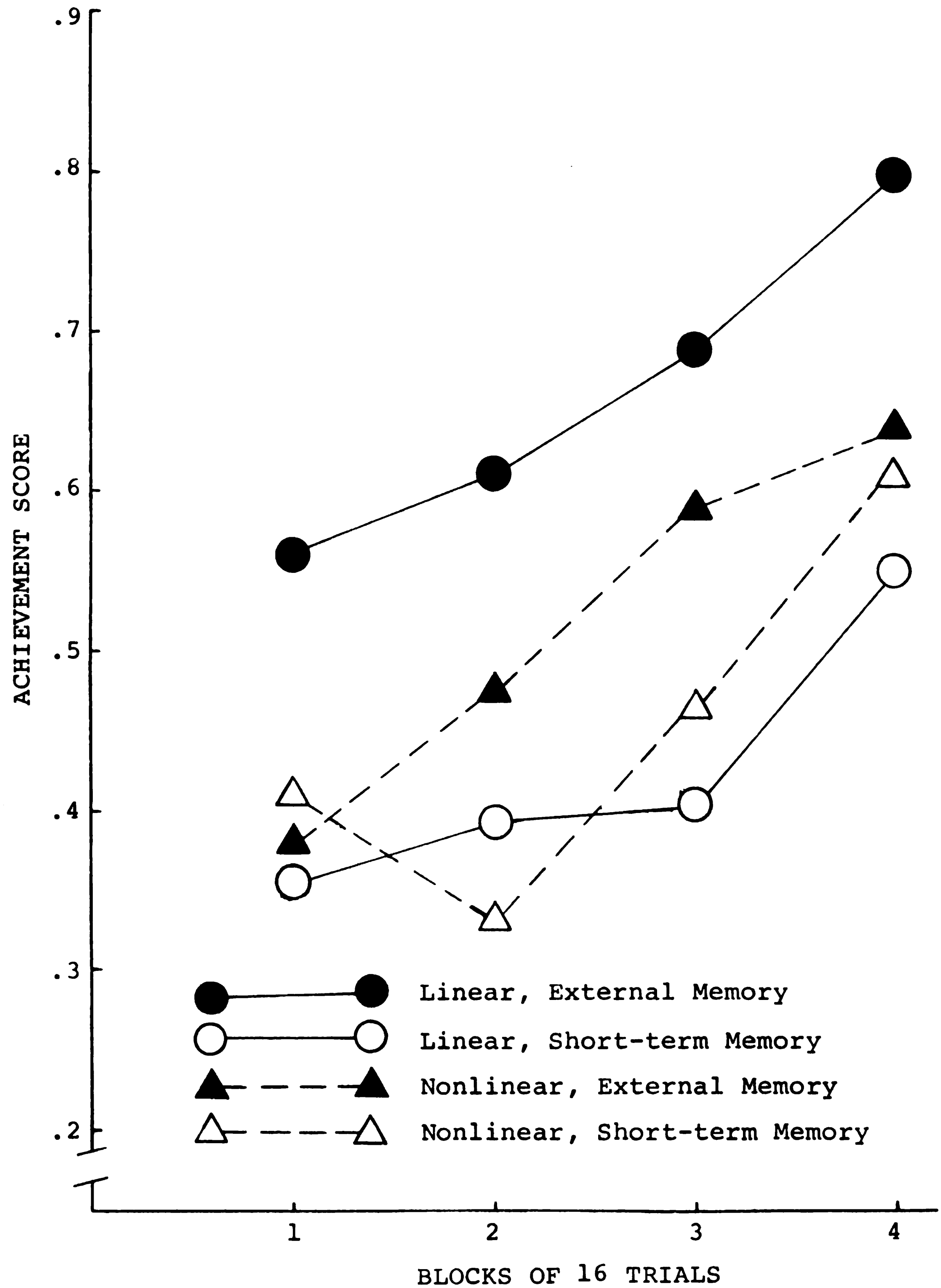

Figure 1. Mean achievement score as a function of rule, memory, and blocks of trials.

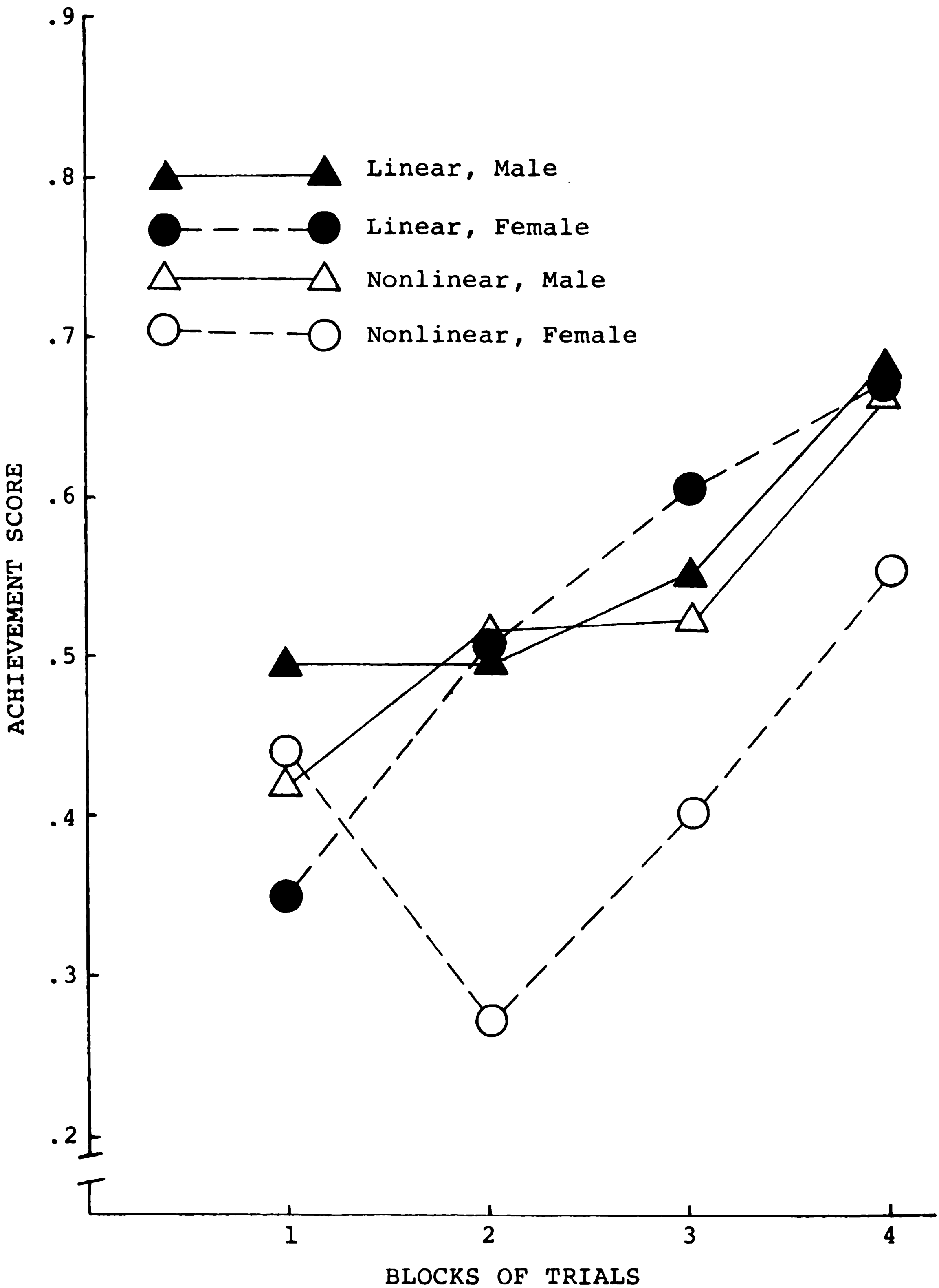

Figure 2. Mean achievement score as a function of rule, sex, and blocks of trials.

$F(3,108) = 4.51$, $p < .01$ (see Figure 3). The short-term memory condition increased the difference between the groups learning the linear and nonlinear rules. However, the external memory condition decreased the difference between the two learning tasks. The verbalization, sex, and block interaction was also significant, $F(6,108) = 4.34$, $p < .01$, as shown in Figure 4. The analytical condition had an interfering effect on females as trials progressed. The females learned best under the thinking-aloud condition and were superior to males, while the reverse was true for the silent condition.

Focusing Score

The results of the analysis of variance for the focusing score are shown in Table B3 (see Appendix B). The significant verbalization and sex interaction is shown in Table 1, $F(2,36) = 4.22$, $p < .05$.

TABLE 1. Mean Focusing Scores for the Verbalization and Sex Interaction.

	Male	Female
Aloud	.386	.485
Analytical	.459	.452
Silent	.599	.362

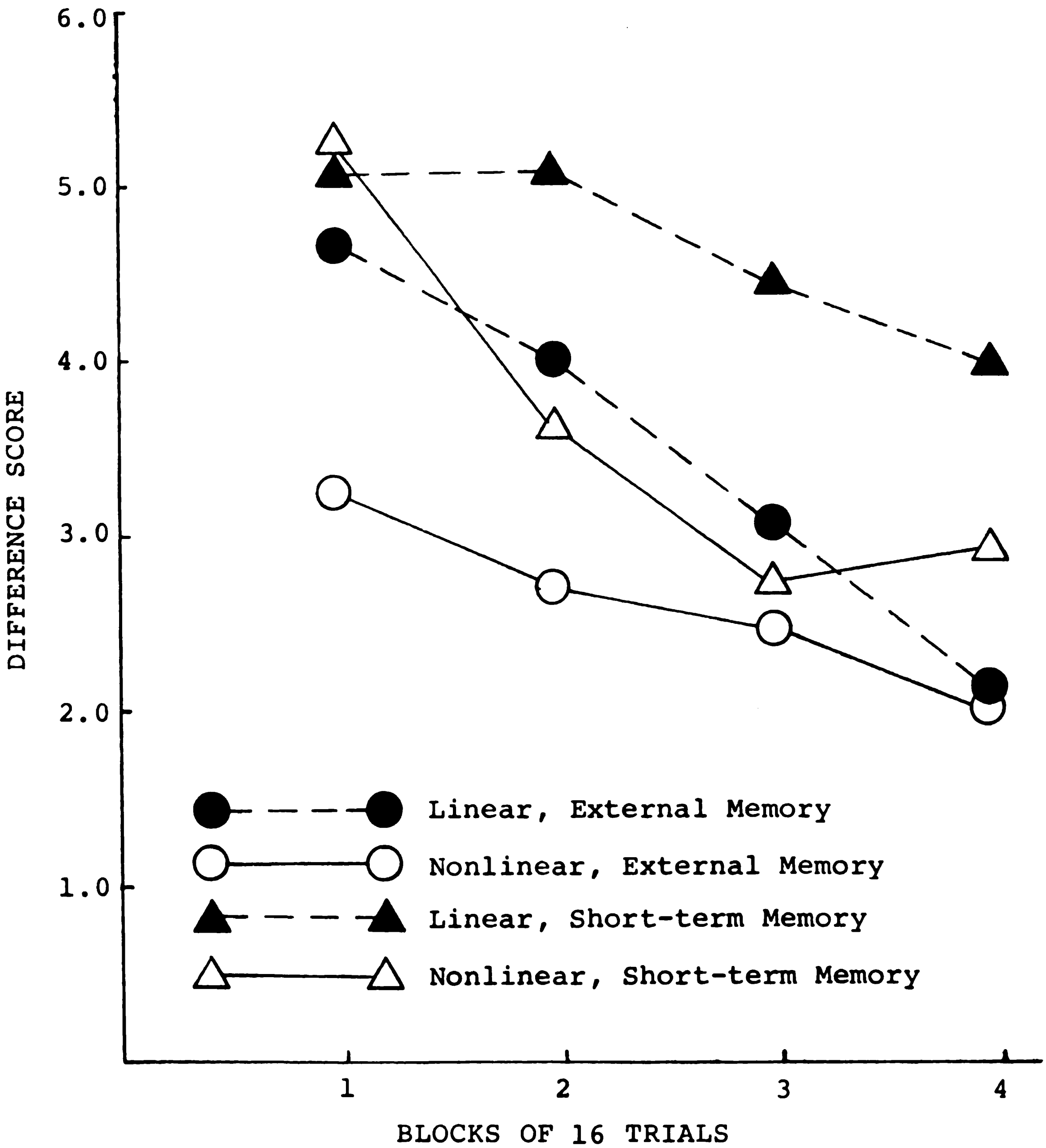

Figure 3. Mean difference score as a function of rule, memory, and blocks of trials.

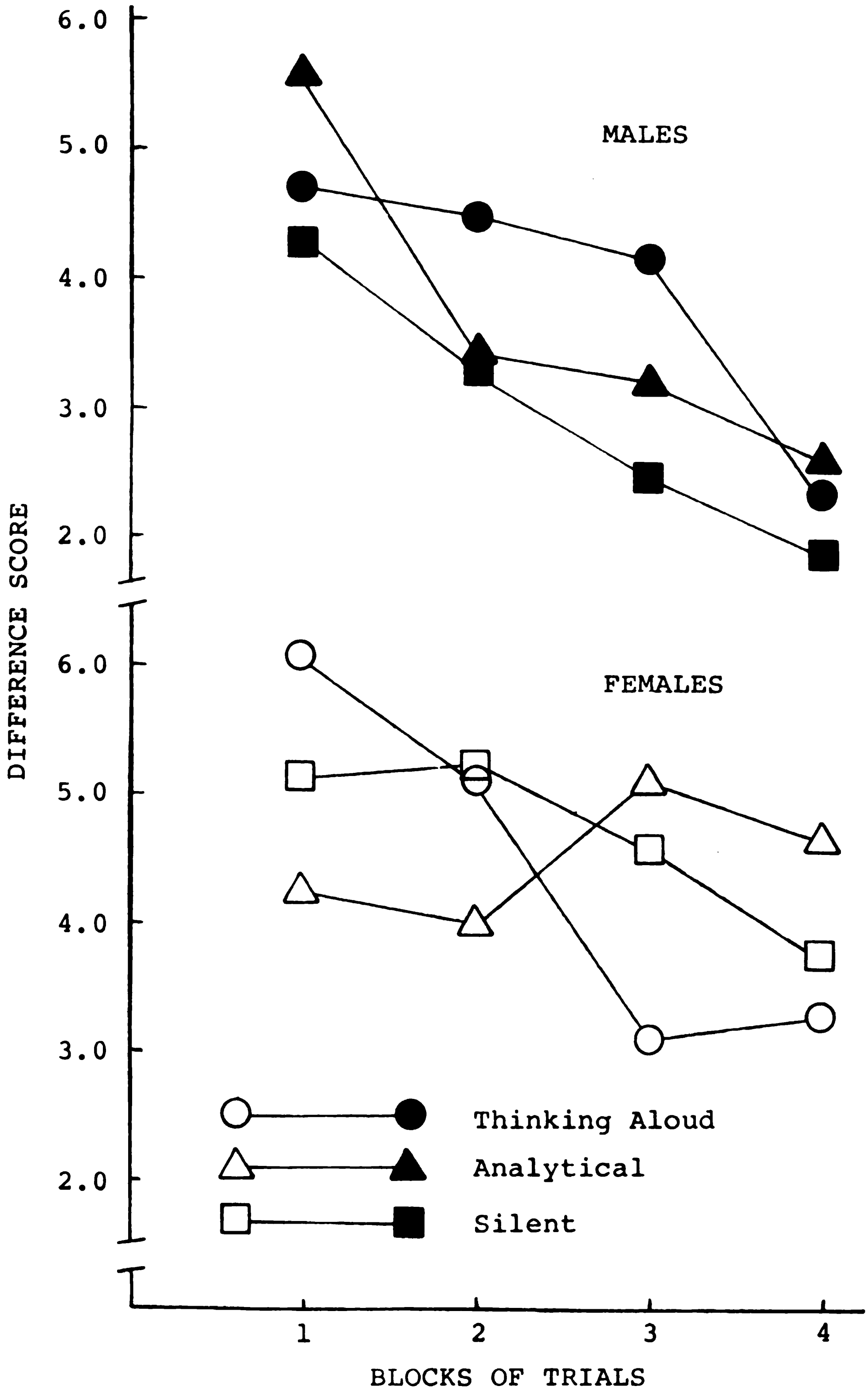

Figure 4. Mean difference score as a function of verbalization, sex, and blocks of trials.

The females used a focusing strategy more often than males in the thinking-aloud condition while the opposite was true for the silent condition. Use of the focusing strategy in the analytical condition is similar for males and females. The significant rule and blocks interaction is shown in Table 2, $F(3,108) = 3.11$, $p < .05$.

TABLE 2. Mean Focusing Scores for the Rule and Block Interaction.

	Block 1	Block 2	Block 3	Block 4
Linear rule	.365	.454	.469	.519
Nonlinear rule	.438	.500	.494	.457

The group learning the nonlinear rule used a focusing strategy more often than the group learning the linear rule for the first three blocks of trials. However, the opposite was true for the last block of trials. The rule, memory, and blocks interaction was significant, $F(3,108) = 2.95$, $p < .05$ (see Figure 5). For the external memory condition, both rules were learned as the dependence on a focusing strategy increased. The linear group using short-term memory also showed a growing dependence on focusing. The nonlinear group in the short-term memory condition, however, initially showed frequent use of a focusing strategy. Their use of the

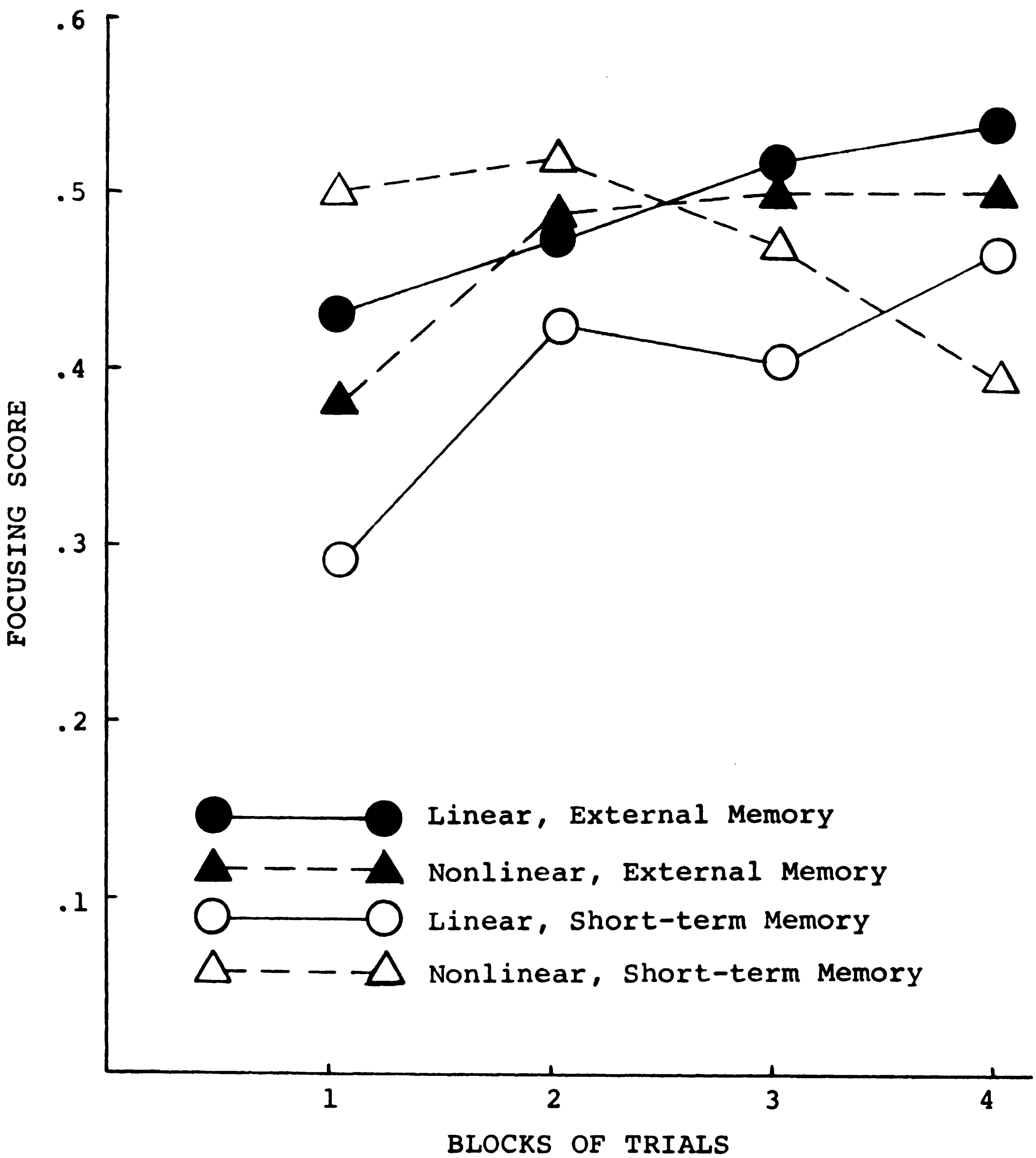

Figure 5. Mean focusing score as a function of rules, memory, and blocks of trials.

focusing strategy declined after the second block of trials.

Total Time

The total time for completing the learning trials was analyzed and the results of the analysis of variance are presented in Table B4 (see Appendix B). The verbalization effect was significant, $F(2,36) = 3.27$, $p < .05$. However, when a comparison of means using Tukey's HSD test was made, no significant differences were detected at $p < .01$; the difference between the analytical and the aloud groups was significant ($p < .05$). The memory and sex interaction was significant (see Table 3) with males taking more time on the short-term memory condition than on external memory condition. The opposite was true for the females.

TABLE 3. Mean Time for Memory X Sex Interaction

	Male	Female
External memory	56.06	62.50
Short-term memory	60.05	57.00

Correlations

Test scores from the Michigan State University Entrance Exams were used as measures of aptitude. Four tests measured reading comprehension and arithmetic skills.

The reading test score (R) provided a measure of verbal ability while the total math (TM) score measured quantitative ability. The arithmetic (A) score identified basic deficiencies in simple arithmetic. The algebra (M) score was specific to past work in algebra. Each aptitude score was used in conjunction with the four dependent measures used in the study: achievement score (Ach), difference score (Diff), focusing score (Focus), and total time. Table 4 indicates the relationships among the measures. (N = 60 for correlations involving only the dependent measures from the present study. For any correlation involving an aptitude score, N = 51). The correlations indicated that none of the aptitude scores were good predictors for the learning task. Significant correlations existed between the aptitude tests and between the dependent measures.

TABLE 4. Correlations Between Tests and Dependent Measures.

	TM	M	A	R	Time	Diff	Ach	Focus
TM	1.00							
M	.98**	1.00						
A	.88**	.77**	1.00					
R	.54**	.52**	.52**	1.00				
Time	.12	- .11	- .11	- .20	1.00			
Diff	- .20	- .22	- .21	- .20	.25*	1.00		
Ach	.09	.08	.15	.01	- .17	- .65**	1.00	
Focus	.22	.24	.13	.18	.23	- .23	.55**	1.00

*$p < .05$.

**$p < .01$.

DISCUSSION

The major purposes of the present study were (1) to determine the effects of availability of previous information on learning linear and nonlinear relations, and (2) to study the effect of verbalization on the efficiency and strategies used in learning the relations. Briefly, the results indicated that (1) while the evidence was inconclusive, use of external memory had the greatest effect on learning the linear relation, (2) contrary to previous results, the linear rule was not easier to learn than the nonlinear rule, and (3) females performed best when thinking aloud while males performed best while learning the rule silently.

The availability of an external memory aided the Ss in learning the relations. Similar results have been found in concept identification studies (Cahill and Hovland, 1960; Hunt, 1961). Newell and Simon (1972) have suggested that the use of an external memory reduces the load of information processing on short-term memory. The learner is not involved in rehearsal to retain the information from each trial, nor does he need to search his memory for the results of previous trials. The concept identification studies showed that Ss offered fewer hypotheses that were inconsistent with previous results when an external memory was available.

Previous studies on learning linear and/or non-linear rules have indicated that the learning process is typically inefficient and slow (Smedslund, 1955; Brehmer, 1969). In fact, the studies often run each S for time periods of two to four hours. If the researcher is willing to separate memory and inference processes, the use of an external memory should increase the efficiency of learning, thereby decreasing the amount of time necessary to run each S.

Three of the dependent measures (focusing score, achievement score, difference score) indicated that the type of rule interacted with the memory conditions as a function of trials. The interactions did not exactly duplicate one another, so it is difficult to make one overall description for the interaction. Researchers have disagreed over the use of the difference score as a measure of learning. Lee (1971) has objected to the heavy reliance on correlational analysis in multiple-cue probability learning. If S is to receive a payoff according to the accuracy of his prediction, then a difference score would be a better measure. Correlations are not good indicators of accuracy; of two Ss, the one with the lower correlation could be the most accurate. However, Uhl (1963) has claimed that the difference score is sensitive to three sources of variation which make its interpretation ambiguous: (a) the validity of the S's subjective weighting of the stimulus, (b) the variance of the S's distribution of responses, and (c) the mean of the S's responses. The achievement score is not affected by systematic over- or underestimation of the criterion. The correlational analysis indicated a highly significant relationship between the two learning scores. Most generally, however, researchers have ignored the use of the difference score in favor of the achievement score. Despite the controversy over the relative validities of the learning measures, some similarities between the measures did exist.

According to both learning measures, the groups learning the linear rule were most effected by the memory conditions. At the end of learning trials, the linear groups were widely separated, while the nonlinear groups tended to converge. While external memory aided the Ss in learning both relations, the group learning the nonlinear relation, when limited to the short-term memory, was able to overcome the memory limitations in learning the rule. The group learning the linear rule, while using short-term memory, performed worst according to the learning measures.

Concept identification studies (Bourne, Ekstrand, and Montgomery, 1969) have shown that the use of a focusing strategy increased as the memory requirements increased. The present study indicates just the opposite. The use of external memory resulted in more frequent use of the focusing strategy, while the most difficult condition, learning the linear rule using short-term memory, used a focusing strategy least of all. Instead of the focusing strategy developing as a compensatory aid in reducing cognitive strain, the use of a focusing strategy was more indicative of a well structured, knowledgeable attack on the problem.

The use of external memory facilitated the inference of the rule, and had the greatest effect on the more difficult rule. However, the linear rule was not easier to learn than the nonlinear relation. Previous studies have consistently shown that nonlinear rules are more difficult than linear rules (Slovic and Lichtenstein, 1971). With the results of the present study, it is difficult to evaluate Brehmer's hypothesis that nonlinear rules are more difficult to learn than linear relations because nonlinear rules require the integration of more information. The comparison of linear and nonlinear rules in the present study wasn't representative of the difference in difficulty.

The present experiment used a nonlinear rule involving a configural relationship which has been modeled by using an absolute difference. However, the most heavily weighted cue in the relationship is strictly linear. Thus the rule can be thought of as a mixture of linear and nonlinear relations. Brehmer (1969) found that a rule involving absolute difference was more difficult than an additive rule. The rules used by Brehmer involved the use of two cues and the values for the cues and the criterion varied along a much wider range of values. Also, past research has involved the use of a reception paradigm rather than the selection paradigm in the present study. Any of these variables, or a combination of them, could explain why the nonlinear rule used in the present study wasn't more difficult than the linear rule.

The sex of the S interacted with each of the other variables considered in the experiment. Both the difference and the focusing score indicated that the sex of the S interacted with the verbalization conditions. Males and females differed in the optimum condition for learning the relations. Males performed best when they remained silent during the experiment and they performed better than the silent females. This relationship was completely reversed for the thinking-aloud condition.

Both Gagné and Smith (1962) and Wilder and Harvey (1971) found that when male Ss were required to give reasons for each move they made in solving a problem, their performance was improved. The present study did not find that requiring male Ss to reason about their choice of cue values or their predictions aided performance. Males performed best when silent. Newell and Simon (1972) suggested that collecting verbal protocols had no effect on problem solving involving a cryptarithmetic problem. They used only male Ss. The present study suggests that males perform worst when they are required to think aloud. Thinking aloud had an interfering effect.

The difference between the performance of males and females could be explained by sex differences in abilities relating to verbalizing and mathematical reasoning. Females are superior to males in verbal fluency. From early infancy to adulthood, females express themselves more readily and skillfully than males (Tyler, 1965). On the other hand, males are superior on tests of mathematical reasoning (but not on tests that require simple computations).

If one assumes that the type of verbalization affected the representation of the task, then the differences in abilities of fluency and mathematical reasoning could explain the interaction. The males performed best when working with numbers and abstract, mathematical representations. Asking males to verbalize could change their representation of the problem or interfere with the

mathematical task. The focusing score indicates that males organized their search for information most successfully when they were silent. With females, verbalizing wasn't interfering, it was facilitative. Giving the task a verbal context could have helped them organize the task. When females verbalized, they used a focusing strategy more often than when they were silent.

Females perform worse at problem solving when they perceive the task as masculine (Milton, 1959). It is possible that a situation involving mathematical relationships and interaction with a computer was regarded as a masculine task, whereas thinking aloud was regarded as feminine.

The interaction between verbalization and sex raises a methodological issue for researchers collecting verbal protocols. While the present study does not agree with other studies on what the effects of verbalization are, it does agree with Gagné and Smith that verbalization does have an effect. Verbal protocols can present evidence for the processes involved in problem solving, concept learning, etc. But control groups are necessary safeguards to check on the effect of verbalization on the processes under study.

Sex differences have generally been ignored in MPL studies. Todd and Hammond (1965) did not find any sex differences in learning linear relations. In the present study, the females had more difficulty in learning the nonlinear relation than the linear relation. In fact, their performance decreased dramatically during the second block of trials. This finding raises the question whether sex differences exist in learning functional relations. Future research should be guided by the possibility that females find nonlinear relations more difficult to learn than do males.

LIST OF REFERENCES

LIST OF REFERENCES

Anderson, N. H. Functional measurement and psychophysical judgment. Psychological Review, 1970, 77, 153-170.

Bourne, L. E., Jr. Some factors affecting strategies used in problems of concept formation. The American Journal of Psychology, 1963, 75, 229-238.

Bourne, L. E., Ekstrand, B. & Montgomery, B. Concept learning as a function of the conceptual rule and the availability of positive and negative instances. Journal of Experimental Psychology, 1969, 82, 538-544.

Bourne, L. E., Jr., Goldstein, S., & Link, W. E. Concept learning as a function of availability of previously presented information. Journal of Experimental Psychology, 1964, 67, 439-448.

Brehmer, B. Cognitive dependence on additive and configural cue-criterion relations. The American Journal of Psychology, 1969, 82, 490-503.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.

Cahill, H. E. & Hovland, C. I. The role of memory in the acquisition of concepts. Journal of Experimental Psychology, 1960, 59, 137-144.

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

Hammond, K. R. & Summers, D. A. Cognitive dependence on linear and nonlinear cues. Psychological Review, 1965, 72, 215-224.

Hoffman, P. J. Cue-consistency and configurality in human judgment. In B. Kleinmuntz (Ed.), Formal representation of human judgment, 1968, 53-90.

Hunt, E. B. Memory effects in concept learning. Journal of Experimental Psychology, 1961, 62, 598-604.

Laughlin, P. R. Selection strategies in concept attainment as a function of number of relevant problem attributes. Journal of Experimental Psychology, 1966, 71, 773-777.

Laughlin, P. R. & Jordan, R. M. Selection strategies in conjunctive, disjunctive, and biconditional concept attainment. Journal of Experimental Psychology, 1967, 75, 188-193.

Lee, W. Decision theory and human behavior. New York: Wiley, 1971.

Maier, N. R. F. & Burke, R. J. Response availability as a factor in the problem solving performance of males and females. Journal of Personality and Social Psychology, 1967, 5, 304-310.

Milton, G. A. Sex differences in problem solving as a function of role appropriateness of the problem content. Psychological Reports, 1959, 5, 705-708.

Newell, A. & Simon, H. A. Human problem solving. Englewood Cliffs, N.J.: Prentice Hall, 1972.

Slovic, P. & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Smedslund, J. Multiple-probability learning. Oslo: Akademisk Forlag, 1955.

Todd, F. J., & Hammond, K. R. Differential feedback in two multiple-cue learning tasks. Behavioral Science, 1965, 10, 429-435.

Tyler, L. E. The psychology of human differences. New York: Appleton-Century-Crofts, 1965.

Uhl, C. Learning of interval concepts: I. Effects of differences in stimulus weights. *Journal of Experimental Psychology*, 1963, *66*, 264-273.

Wilder, L. & Harvey, D. J. Overt and covert verbalization in problem solving. *Speech Monographs*, 1971, *38*, 171-176.

APPENDIX A

INSTRUCTIONS

APPENDIX A

GENERAL INSTRUCTIONS

Your task in this experiment is to use the value of three different cue numbers to make a prediction of a fourth number. The display in front of you has a series of rectangles painted white and three numbers of different colors. Each color represents a source of information separate from the other colored numbers. For example, 324 should be thought of as a blue 3 and a red 2 and a green 4--not three hundred and twenty-four. The numbers when combined using addition, subtraction, absolute value, multiplication, division, or exponentiation will produce the answers you are to learn to predict.

The cue numbers differ in their importance (or weight) in producing the answer. An example of the possible relationship between two cue values would be:

cues		answer	cues		answer
2,5	=	15	1,2	=	7
2,2	=	12	1,5	=	10

What is the mathematical relationship between the two numbers on the left that would produce the answer on the right?

Answer: $2(5) + 5 = 15$ $\quad$ $1(5) + 2 = 7$

$2(5) + 2 = 12$ $\quad$ $1(5) + 5 = 10$

In this case, the first cue number was five times (a weight of five) more important than the second cue number. When a cue number is weighted, this involves multiplying the cue number by a constant value. The relationship between the cue values was addition.

Consider this problem:

cues		answer	cues		answer
3,1	=	1	4,4	=	3
3,4	=	4	2,4	=	6

What is the mathematical relationship between the two numbers on the left (including the weight of each cue) that would produce the number on the right?

Answer:

$\frac{1(3)}{3} = 1$ $\quad$ $\frac{4(3)}{4} = 3$

$\frac{4(3)}{3} = 4$ $\quad$ $\frac{4(3)}{2} = 6$

In this case, the second cue number was weighted by the three (the number in the parenthesis). The relationship between the cue numbers involved division; the second cue number was divided by the first cue number.

The number of possible combinations of different cue weights (the importance of different cues) and the different mathematical relationships between the cues makes the chance of a lucky guess extremely unlikely. You need to discover a way to seek information so you can tell how cues are weighted and how the different cue values are related. In your task, you will be working with three cues instead of two.

You will be required to go through all sixty-four combinations of cue numbers on the board. Try to pace yourself so that you make one prediction every minute; then the task will last about one hour. It is very possible you won't be able to infer the correct rule, but your predictions should get closer to the correct answer.

INSTRUCTIONS FOR THINKING-ALOUD

During the experiment, we would like you to think "aloud" as you are learning the rule. This means that you should verbalize any and all thoughts, whether the thoughts are complete or just fragments. As long as you are thinking you should be talking about the thoughts. Try to keep a steady stream of talking. Don't stop talking to think. We are very interested in what people are thinking as they learn the rule. So please, think aloud.

INSTRUCTIONS FOR ANALYTICAL CONDITION

During the experiment we want you to carefully reason and plan each action you take. So, you are required to (1) state a reason for selecting the cue values that you select for each trial and (2) state a reason why you make the prediction for your answer each trial. Be sure to state these reasons outloud before you enter the values for the cues or for the prediction into the computer. We are very interested in what reasoning processes people use as they learn the rule. So please, state your reasons outloud.

APPENDIX B

ANALYSIS OF VARIANCE TABLES

APPENDIX B

TABLE B1. Achievement Score: Analysis of Variance

Source	df	MS	F
R (Rule)	1	0.168	0.994
V (Verbalization)	2	0.040	0.237
M (Memory)	1	1.366	8.083**
S (Sex)	1	0.271	1.604
RV	2	0.279	1.651
RM	1	0.494	2.923
RS	1	0.149	0.882
VM	2	0.077	0.453
VS	1	0.351	2.077
MS	1	0.064	0.379
RVM	2	0.037	0.219
RVS	1	0.067	0.396
RMS	1	0.203	1.201
VMS	2	0.021	0.124
RVMS	2	0.424	2.506
Error	36	0.169	
B (Block)	3	0.650	20.968**
RB	3	0.018	0.581
VB	6	0.063	2.030
MB	3	0.028	0.903
SB	3	0.043	1.388
RVB	6	0.054	1.742
RMB	3	0.358	11.549**
RSB	3	0.115	3.720*
VMB	6	0.032	1.032
VSB	6	0.051	1.645
MSB	3	0.011	0.355
RVMB	6	0.055	1.765
RVSB	6	0.051	1.645
RMSB	3	0.061	1.968
VMSB	6	0.056	1.797
RVMSB	6	0.024	0.758
Error	108	0.031	

*p < .05.

**p < .01.

TABLE B2. Difference Score: Analysis of Variance.

Source	df	MS	F
R (Rule)	1	52.020	4.846*
V (Verbalization)	2	6.225	0.580
M (Memory)	1	66.470	6.192*
S (Sex)	1	0.470	0.044
RV	1	0.450	0.042
RM	1	0.540	0.050
RS	1	15.550	1.449
VM	2	0.637	0.059
VS	2	7.760	0.723
MS	1	0.035	0.003
RVM	1	5.829	0.543
RVS	2	11.765	1.096
RMS	1	10.100	0.941
VMS	2	5.610	0.527
RVMS	2	22.795	2.124
Error	36	10.734	
B (Block)	3	34.260	27.126**
RB	3	2.107	1.667
VB	6	2.555	2.020
MB	3	0.890	0.705
SB	3	2.080	1.647
RVB	6	2.175	1.722
RMB	3	5.700	4.513**
RSB	3	1.050	0.831
VMB	6	1.267	1.003
VSB	6	5.481	4.340**
MSB	3	0.898	0.711
RVMB	6	2.365	1.872
RVSB	6	0.779	0.617
RMSB	3	0.423	0.335
VMSB	6	1,820	1.441
RVMSB	6	0.906	0.718
Error	108	1.263	

*p < .05.

**p < .01.

TABLE B3. Focusing Score: Analysis of Variance.

Source	df	MS	F
R (Rule)	1	0.024	0.179
V (Verbalization)	2	0.105	0.784
M (Memory)	1	0.117	0.874
S (Sex)	1	0.138	1.031
RV	2	0.021	0.153
RM	1	0.142	1.061
RS	1	0.196	1.464
VM	2	0.042	0.310
VS	2	0.566	4.223*
MS	1	0.040	0.299
RVM	2	0.001	0.008
RVS	2	0.057	0.426
RMS	1	0.065	0.471
VMS	2	0.067	0.497
RVMS	2	0.004	0.261
Error	36	0.134	
B (Block)	3	0.099	6.000**
RB	3	0.051	3.109*
VB	6	0.012	0.746
MB	3	0.033	1.976
SB	3	0.045	2.697
RVB	6	0.006	0.364
RMB	3	0.049	2.948*
RSB	3	0.030	1.804
VMB	6	0.006	0.390
VSB	6	0.016	0.984
MSB	3	0.021	1.270
RVMB	6	0.011	0.659
RVSB	6	0.012	0.656
RMSB	3	0.029	1.778
VMSB	6	0.019	1.154
RVMSB	6	0.008	0.473
Error	108	0.017	

*$p < .05$.

**$p < .01$.

TABLE B4. Total Time: Analysis of Variance.

Source	df	MS	F
R (Rule)	1	38.000	0.566
V (Verbalization)	2	219.500	3.268*
M (Memory)	1	0.866	0.013
S (Sex)	1	41.000	0.610
RV	2	52.000	0.774
RM	1	1.134	0.017
RS	1	244.000	3.633
VM	2	150.070	2.234
VS	2	22.500	0.335
MS	1	325.134	4.840*
RVM	2	62.933	0.937
RVS	2	149.500	2.226
RMS	1	110.866	1.651
VMS	2	69.433	1.034
RVMS	2	215.367	3.209
Error	36	67.166	

*$p < .05$.

**$p < .01$.