EFFECTS OF VERBALIZATION, TEMPO, AND PROBLEM SIZE ON PROBLEM SOLVING

Dissertation for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
STUART O. HALLGREN
1976

This is to certify that the

thesis entitled

EFFECTS OF VERBALIZATION, TEMPO, AND
PROBLEM SIZE ON PROBLEM SOLVING

presented by

Stuart O. Hallgren

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

Donald M. Johnson
Major professor

Date 5-4-76

O-7639

G4954

ABSTRACT

EFFECTS OF VERBALIZATION, TEMPO, AND PROBLEM SIZE ON PROBLEM SOLVING

By

Stuart O. Hallgren

Eighty-four introductory psychology students solved either a five or six disk Tower of Hanoi problem. The effect of verbalization on problem solving was examined by requiring subjects to either think aloud, state reasons for moves, or remain silent. Half of the subjects solved the problem at their own speed, while the other half were required to wait at least seven seconds before making the next move. The experimental design included a 2(problem size) x 3(verbalization) x 2(tempo) completely randomized factorial design.

The results were examined in terms of the number of excess moves, the time per move, the number of stereotypic moves, and an episode analysis. The analytical group, which stated reasons for making moves, solved the problems in fewer excess moves than the silent or thinking aloud groups. The superior performance of the analytical group does not depend on a transfer paradigm nor can the effects be explained as a speed-accuracy trade-off. Analytical instructions enhanced active processing of the problem and division of the problem into subproblems and can therefore be valuable in training people in solving problems. The thinking aloud group was poorer at problem solving than the silent group. Instructions to think aloud resulted in different behavior than instructions to be analytical, contrary to the position of researchers who equated the two. Thinking aloud caused more excess moves to solution and more stereotypic behavior. Methodological implications were discussed. Verbal protocols produced by subjects thinking aloud do not provide an unequivocal description of the information processing of subjects thinking silently.

The effects of the different verbalization instructions increased with problem size. For the thinking aloud condition, the interaction was consistent with the hypothesis that a limited capacity for performing two tasks (problem solving and communicating) would be exceeded with large problems. For the analytical group, the interaction supported the hypothesis that the subjects' ability to reduce the search space by planning was more effective as the problem size increased. The tempo for making moves also interacted with the size of the problem. Slowing subjects down improved problem solving performance, but only for the larger problem.

EFFECTS OF VERBALIZATION, TEMPO, AND

PROBLEM SIZE ON PROBLEM SOLVING

By

liver
Stuart O. Hallgren

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Psychology

1976

To Kathy--whose love and understanding supported all my efforts

ACKNOWLEDGMENTS

I wish to express my appreciation to the members of my committee: Drs. Lester Hyman, Donald Johnson, Gary Olson, and Gordon Wood. I particularly want to thank the chairman of the committee, Dr. Johnson, for his interest in my ideas, his suggestions, and his assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Introduction to Process Tracing

Studies of problem solving tend to be primarily concerned with either the strategies and processes used by problem solvers or the variables which affect the difficulty of the problem (Bourne, Ekstrand, and Dominowski, 1971). Most process tracing experiments involve asking the subject to verbalize about their problem solving efforts. Despite the extensive use of process tracing experiments in the development of theory, little attention has been directed at assessing any possible effects of the verbalization requirement on problem solving.

History of Process Tracing

The method of thinking aloud was first proposed by Edourand Claparède in 1917 as an improvement over introspection or the postexperimental interview. Claparède used the method to study the generation of hypotheses during problem solving in order to obtain a running account of the problem processing. Verbal protocols have been used by Gestalt psychologists such as Duncker (1945), Maier (1931), and Scheerer (1963) to show the change in "direction" of thought as subjects solved functional fixedness problems. Even associative theorists have recommended the use of verbal protocols to test hypotheses on the nature of the implicit responses produced during problem solving (Mayzner, Tresselt, and Helbock, 1964).

As theories of information processing have become more popular, the collection and use of verbal protocols has also increased. Models of mental processes during problem solving have been validated by comparing traces produced by a computer program with the verbalization of subjects asked to think aloud. Protocols have been collected for a wide variety of tasks, including problems such as the missionaries and cannibals, cryptarithmetic, elementary logic problems, chess, binary choice sequence prediction, geometry proofs, and concept identification for the induction of various logical and sequential concepts. The collection of verbal protocols, and the subsequent production of problem behavior graphs and simulation programs, has been pioneered by Allen Newell and Herbert Simon (1972). They have based a theoretical system of problem solving upon the verbal protocols of subjects who were asked to think aloud. Thus, the use of verbal protocols as a mirror of thought processes has a long history in the psychology of problem solving. Although information processing theorists have made extensive use of verbal protocols, the use of protocols has not been restricted to one theoretical approach.

Criticism of the Methodology

The use of thinking aloud as a methodology has been criticized on a number of grounds. DeGroot (1965) used verbal protocols in order to study expert chess players. The major question, according to DeGroot, was how much thinking and reporting on thinking interferred with each other while the subject was thinking aloud. Verbalizing one's thoughts certainly adds an extra burden to the subject's task. Many subjects in DeGroot's study who were required to think aloud reported that the verbalization slowed down thinking. The requirement to think aloud could change the course of thought. Neisser (1963) has claimed any complex or multiple processing done by the subject may erroneously appear sequential when the subject thinks aloud. Or thinking aloud might cause the subject to use a sequential process where he would not otherwise.

Experimental Evidence on Verbalization

Does the requirement for a person to verbalize influence the mental processes under study? The experimental evidence on the queston is unclear. What the subject is asked to verbalize seems to determine the effects of verbalization.

Analytical Instructions

When a person was asked to state reasons for steps and choices made in solving a problem, the analytical verbalization aided problem solving. Gagné and Smith (1962) required some subjects to state reasons for their moves as they attempted to solve the Tower of Hanoi problem involving three disks, four disks, and then five disks. On a transfer problem (with six disks), during which no subject verbalized, those who had previously stated reasons solved the problem better than those who had solved the previous problems silently. At the end of the experiment, all subjects were asked to state a rule about how the problem should be solved, and their answers were judged for adequacy. Subjects who had verbalized during training gave better answers than those who had not verbalized.

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 overtly 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 excess moves to solve the problem. The verbalization per se was not necessary to improve problem solving. The instructions to be analytical, to think of reasons for moves, provided the subjects with a good search strategy.

Experimenters have required subjects to state hypotheses as they attempt to learn concepts. The requirement to state hypotheses improved the rate of learning. Byers and Davidson (1967) noted that consistent offering of hypotheses aided concept learning, and probably slowed down behavior involved in testing stimuli. The complexity or difficulty of the conceptual rule determines whether requiring hypotheses have a facilitating effect. Dominowski (1973) found that requiring hypotheses after each trial aided performance with either conjunctive or disjunctive rules while having no effect on learning a unidimensional concept.

Thinking Aloud Instructions

Newell and Simon (1972) have argued that instructions to state hypotheses or reasons have a different effect than instructions to think aloud. While using verbal protocols as useful evidence, Newell and Simon recognized that it was crucial to assess the effects of thinking aloud on the methods and strategies involved in problem solving. They compared the behavior of five subjects who solved logic problems while thinking aloud with twenty-four subjects who wrote their attempts at a solution on a blackboard. The distribution of the number of steps required to solve the problem was judged similar for the two groups. Subjects working under the thinking aloud conditions generated much the same logical expressions as did the silent group.

Newell and Simon therefore concluded that thinking aloud did not modify the directions of the search for solution. Unfortunately, the conclusion was based on a qualitative analysis. No statistical tests were used as a basis for their decision. Also, the data for the silent group was collected from a study on the effects of alcohol on problem solving (Carpenter, Moore, Snyder, and Lisanky, 1961). The largest amounts of alcohol produced about a thirty percent decrement in performance. Certainly a comparison using subjects who had consumed alcohol in the silent group might fail to indicate any decremental effects of thinking aloud.

There is prima-facie evidence that thinking aloud does change the course of thought. Dansereau and Gregg (1966) found that silent mental multiplication took the same amount of time as doing the multiplication aloud. According to their subject's introspection the processing could be done more rapidly during the silent condition than the aloud condition. However, the memorization of the digits was more difficult under the silent condition. Dansereau (1969) mentioned that his subjects reported thinking aloud actually interferred with the mental multiplication task. As practice with the problems increased, the subjects altered their verbal protocols so that they did not report single digit addition and multiplication. The subjects reported that they eliminated those steps because verbalizing them seemed to interfere with the retrieval of other information. Requiring subjects to verbalize while solving "hard" anagrams had a detrimental effect (Nance and Simmott, 1964). The dependent measure was time to solution. The study did not necessarily show that thinking aloud interferred with problem solving. It did show that verbalization slowed problem solving.

Verbalization Effects: Theory and Prediction

The basic research question centers upon the reactive effect of measurement. Whether verbalization affects problem solving raises major questions about the external validity of process tracing experiments and the theories based upon these studies. The development of the appropriate theory of problem solving depends upon the use of the appropriate method.

Instructions Determine Problem Spaces

The theoretical explanations center upon the information transmitted by the verbalization instructions. The instructions given to a subject about a problem solving task help define what Newell and Simon (1972) have called the problem space. The problem space includes the subject's basic representation of the task that helps him decide what to do in different situations and how to apply operators that change one situation into another. The basic information communicated through instructions and incorporated in the problem space include the basic elements of the problem, the initial state of the problem and the goal, the operators used to transform states of the problem, and the restrictions on the use of the operators. Differences in instructions will include additional information which can aid or hinder finding a solution and induce the subjects to use different problem solving methods.

Thinking Aloud as a Time Sharing Task

The evidence on the effects of thinking aloud is sparse and inconclusive. There is some reason to suspect that thinking aloud may actually interfere with problem solving. The possible interference would support Newell and Simon's contention that thinking aloud is different than analytical verbalization, but would refute their view that thinking aloud does not affect the processes under study. Asking a subject to think aloud is the most common verbalization instruction. The present study tested the hypothesis that thinking aloud interferes with problem solving. When required to think aloud, the problem space actually includes two concurrent tasks: (1) attempting to solve the problem and (2) communicating about those attempts. Both tasks demand processing time and effort. Because the thinking aloud group must pay attention and divert effort to the communications task, their performance on the problem solving task would suffer. The basic question involved whether or not verbalization detracts attention from the problem solving task. We often find it difficult to execute two activities together, although either alone is easy. For example, Peterson (1969) found that complex covert problem solving, such as the solution of anagrams, could be completed while the subject was counting or reciting the alphabet but the additional task did interfere with problem solving. The model of attention proposed by Kahneman (1973) assumes the capacity which can be allocated between two tasks is limited. As the demands on attention increase, level of performance declines.

Analytical Instructions Induce Planning

The evidence on analytical verbalization indicated that the instructions actually aided problem solving and this effect was important as far as optimizing problem solving. The question involved how analytical instructions change the basic problem space. Requiring a person to state reasons for each move in solving a problem results in a different type of information processing. Focusing a person's attention on producing reasons for making moves involves evaluating moves, considering different moves, and seeking relationships between sequences of moves. The analytical instructions are more directive than thinking aloud instructions. The subject knows what he is expected to verbalize and that his problem solving efforts must produce reasons for each move. Data from previous studies on analytical instructions (Gagné and Smith, 1962; Wilder and Harvey, 1971) indicated that analytical subjects take more time in choosing moves. The extra time suggested that the analytical subject was taking more time to plan and evaluate possible sequences of moves. Pilot studies have indicated that analytical subjects actually break the problem into parts and establish a series of subgoals. By planning subgoals, the analytical subjects are able to replace a single difficult problem with two or more simpler problems.

The importance of establishing subgoals in solving the Tower of Hanoi problem has been stressed by Egan and Greeno (1974). According to their theory, the most important outcome of experience in solving the Tower of Hanoi problem is establishing a system for generating subgoals. Once the system is learned, a cognitive structure exists for simply producing the answer. If the theory is correct, the superiority of the analytical groups on transfer problems would be explained by their development and use of the subgoal strategy.

Problem Size

The present study hypothesized that subjects in the various verbalization conditions solve problems differently and that the groups would differ in their effectiveness. The relative effectiveness of the groups depend upon the size of the problem. The size of the Tower of Hanoi problem is directly related to the number of disks. As the number of disks increases, both the number of excess moves to solution and time to solution increase at an exponential rate (e.g., Cook, 1937). Increasing the number of disks changes the problem in other ways besides increasing the number of problem elements. Both the length of the minimum solution and the size of the search space increase at an exponential rate as the number of disks increases. The minimum number of moves necessary for solution can be calculated by the formula 2^n-1, where n equals the number of disks in the problem. Excess moves involve the number of moves a subject made beyond the minimum number needed. The search space of a problem involves all possible distinct moves at any stage in the solution and can be represented by a tree structure. An example of a modified search space is presented in Figure 3 (page 34).

Psychological Significance of Problem Size

Psychologically, the size of a problem has important consequences in terms of the load on working memory and the difficulty in evaluating moves. Obviously, as the number of disks increase, there is more information to remember concerning the location and relative positions of the disks.

While solving the Tower of Hanoi problem, a subject will try to keep track of (1) subroutines or series of moves and their relative success (2) specific moves and previous problem states and (3) anticipated consequences of a plan. A crucial question about memory in the problem concerns when the limitations of short-term memory (STM) are exceeded, and how subjects deal with such a limitation. Most subjects can solve the three disk problem in the minimum number of moves, that is, seven moves. Performance on the four disk problem, however, indicates that most subjects are unable to plan the most effective solution. Planning a solution involving fifteen moves exceeds the limitations of STM. How does the memory limitation affect the strategies used by subjects? It seems doubtful that subjects will attempt to exceed memory limitations more than once or twice. The subjects would reject any strategy which involves planning ahead too many moves because they will simply lose track of the plan.

As the Tower of Hanoi problem increases in size, a problem solver encounters more difficulty in evaluating progress toward the goal. The fewer the steps or moves to solution, the easier it is to imagine or plan the intermediate steps. Consider the extreme case. Where does one move the first disk? Does it make a difference? With four disks or more, the subject is unable to trace all the possible consequences of the first move. The distance to the goal seems the same regardless of the move selected. Several other disks remain to be moved before a solution is reached. The idea that beginning moves are the most difficult to evaluate was supported by a study on learning to solve the Tower of Hanoi problem in the minimum number of moves. Egan and Greeno (1974) found that subjects made the most errors in learning the first part of the solution to the problem. Evaluation of moves is only effective when one can correctly assess the consequences of a certain move in relationship to future moves. The more difficult the evaluation of a move, the more arbitrary the selection of the move.

Interaction Between Verbalization and Problem Size

As the search space of a type of problem increases, the search strategy used to solve the problem becomes even more important. A search method or problem solving strategy that can reduce the size of the search space for the problem solver increases its relative advantage as the size of the search space increases. Setting subgoals is a search strategy that has special advantages for search problems, and its great efficiency has been noted by a number of theorists (Wickelgren, 1974; Nilsson, 1972). In essence, the purpose of setting subgoals is to replace a single difficult problem with two or more simpler problems.

The vast advantage of using subgoals in solving a problem is the dramatic reduction of the search space. Wickelgren (1974) cited an example where a problem requires a sequence of n actions necessary for a solution and m alternative actions at each problem state. By systematic trial and error there are m^n alternative paths (sequences of moves) to be investigated in the original problem. Now assume that the problem solver can define a subgoal state that is known to be on the correct path to the goal halfway from the beginning to the goal. Defining one subgoal divides the problem into two subproblems--first getting from the initial problem state to the intermediate subgoal, and, second, getting from the subgoal to the goal. Thus with the single subgoal, the number of action sequences to be investigated is $2m^{n/2}$ action sequences that are n/2 steps long versus m^n action sequences that are n steps long in the original problem without a subgoal.

The present study tested the hypothesis that the analytical group plan and establish subgoals. Therefore, the superiority of the analytical group should be more pronounced with more complex problems. On the other hand, if thinking aloud limits the time and effort involved in problem solving, the relative disadvantage of the competing tasks would increase as the problem size increased (Kahneman, 1974). In order to study the effects of problem complexity, subjects solved either the five disk or the six disk problems. Past research on analytical verbalization indicated that there was a significant difference between groups on both the five and six disk problem, but these studies (Gagné and Smith, 1962; Wilder and Harvey, 1971) gave the subjects experience solving the problem with fewer disks. Problem size was confounded with order of presentation and no clear interpretation of the results is possible.

Problem Solving Tempo and Verbalization

Time pressure is often involved in mental tasks. It is sometimes imposed by explicit instructions to hurry and sometimes by demand characteristics of the task. Regardless of its cause, time pressure can have a detrimental effect on problem solving. Ray (1962) found that when a group was asked to work as fast as possible on the Tower of Hanoi task, they took more moves to solve the problem than a group not under time pressure. Also, when subjects were instructed to work quickly, they were less likely to break a problem solving set (Luchins, 1942). Conversely, when a subject was instructed to analyze the problem carefully or to take time to think about the problem before he began, problem solving was more efficient and set was easier to break (Duncan, 1963; Rokeach, 1950). Requiring a subject to stop to think has been identified as one of the most generally effective problem solving strategies. The immediate production (and acceptance) of solutions interferes with the chances of cognitive reorganization or identifying alternative response patterns (Johnson, 1972).

Results from a pilot study indicated that subjects who were thinking aloud took less time between moves than subjects who solved the problem analytically. Studies using analytical verbalization found that analytical subjects took more time between moves than did silent subjects. It was hypothesized that analytical instructions increased planning and that thinking aloud decreased time and effort for planning activity. The hypothesis tested in the present experiment was that slowing down the subjects (somewhat) decreases the number of moves required to solve the Tower of Hanoi problem. Also, when the thinking aloud group and the silent group take more time between moves, the difference between the verbalization conditions decreases.

Time pressure increases the tendency to use set responses, and the hypothesis was formulated that thinking aloud also competes with the time and effort involved in problem solving. When the subject is involved with a complex problem, Kahneman (1974) claimed that a secondary task (such as thinking aloud) would narrow the focus of attention. Therefore, the present study tested the hypothesis that subjects who were required to think aloud would show more stereotypic responses in their solution of a complex problem.

The Problem

The Tower of Hanoi problem was not invented specially by psychologists in order to study problem solving. According to Gardiner (1959) the history of the problem extends back to 1883 when it was invented by the French mathematician Edouard Lucas and sold as a toy. The commercial venture was accompanied by an exotic tale to account for the origin of the problem. According to legend, there exists a "Tower of Brahma" in the Indian city of Benares. The tower consists of sixty-four disks of gold which are now in the process of being transferred by the temple priests. When the task is completed, the tower will crumble to dust and the end of the world will be at hand. Fortunately, any solution involves 2^n-1 minimum number of moves to solution (where n equals the number of disks) so that the solution to the Tower of Brahms will involve at least $2^{64}-1$ moves. Even if one move is completed each second, it will be many thousands of millions of years before the end is at hand.

Psychologists have used the Tower of Hanoi problem because it externalizes major portions of the subject's solution attempts. Because the solution requires several moves, use of the Tower of Hanoi problem allows the experimenter to trace and analyze the mediating processes involved between presentation of the problem and the production of the solution. Asking the subject to verbalize his thoughts has been the traditional method used to trace the mediating processes. Of course, the focus of the present study was to compare groups which verbalized during problem solving to a group that did not. The Tower of Hanoi problem allows the experimenter to trace intermediate steps even for a silent group.

Because the Tower of Hanoi problem requires many moves to solution, the number of excess moves can be used as the dependent variable. Problems that employ time to solution (anagrams, insight problems, etc.) as the dependent variable are inappropriate. If verbalization did increase the time to solution in solving anagrams (as in the Nance and Sinnott study, 1964), the result would only indicate that verbalization slowed problem solving, not that it altered the basic information processing in any way. The processes involved in solving the problem could be identical to those used by a silent group, and used in the same sequence, only the verbalization groups would take extra time to report the processes.

Even if the Tower of Hanoi problem has certain useful characteristics for studying the effects of verbalization, the question of generality of any results remains. There is no such thing as the "general problem" and, of course, the best answer to the question of generality can be answered only by further research. However, the limited attempts at establishing a problem typology do suggest the Tower of Hanoi problem can be classified within the large class of search problems. The Tower of Hanoi problem involved a large number of intermediate steps for solution and several alternatives are possible at each intermediate step. A large number of nontrivial problems in mathematics, science, and engineering involve search problems (Wickelgren, 1974).

There are a number of search problems that psychologists have employed in previous research, such as the missionaries and cannibals problem, cryptarithmetic problems and logic proofs. However, these problems all have limitations as vehicles for studying verbalization effects. First of all, the cryptarithmetic problems and the logic proofs do not compel or encourage the subjects to record intermediate steps. Therefore, it could be difficult to trace the progress of a silent group as it solves the problem. The experimenter predicted that a certain problem size was necessary to obtain differences between the verbalization groups. The logic proofs and the missionaries and cannibals problems do not involve large search spaces or long solutions.

Summary of Hypotheses

1. Verbalization effects. Three different groups were used: an analytical group, a thinking aloud group, and a silent group. The analytical group was expected to solve the problems more efficiently than the silent group, while the thinking aloud group would be the least efficient. Just the opposite results were expected as far as stereotypic behavior was concerned; the thinking aloud group would show the most stereotypic behavior. Strategy differences were expected between the verbalization groups.

2. Complexity effects. Performance on the five disk problem would be better than on the six disk problem.

3. Complexity x Verbalization interaction. The size of the differences between the verbalization groups would increase as the complexity of the problem increased.

4. Tempo effects. The group which was forced to take time between moves would solve the problem more efficiently than the group which made moves at its own pace.

5. Tempo x Verbalization interaction. The size of the differences between the verbalization groups would be less when the subjects were forced to take time between moves.

METHOD

Design and Subjects

The subjects were 84 volunteers from introductory psychology classes at Michigan State University who were fulfilling a class requirement. Each subject was randomly assigned to one of twelve experimental conditions. The basic design consisted of a factorial combination of three verbalization conditions (thinking aloud, analytical, or silent), two tempo conditions (self-paced or delayed), and two different levels of problem size (each subject solved either a five disk or a six disk problem).

Materials

The Tower of Hanoi (Tower of Brahma, disk transfer problem, etc.) has been described by Ewert and Lambert (1932). The problem was a commercial model, called the "Tower of Trouble" and the pegs were arranged in a linear fashion. Each disk was numbered one through six in order from the smallest to the largest. Each peg was lettered A, B, or C.

Procedure

Each subject was tested individually. Three subjects were eliminated from assignment to a group because of previous experience with the problem. All subjects received instructions which stated the goal of the problem, pointed out the labeling, specified the rules of the problem and required them to state each move outloud. The instructions emphasized that the problem should be solved in the least number of moves possible, without worrying about the time involved. In addition, the analytical group was required to state a full reason for each move as completely as possible. The thinking aloud group was asked to verbalize any and all thoughts about the problem and to produce a steady stream of talking. Subjects in the self-paced condition were told that the experimenter would press a button which briefly turned on a light each time the subject made a move. In the delayed condition, subjects were told to wait until the light went off (a seven second delay) before they could make another move. The complete instructions are presented in Appendix B. After all the instructions had been presented and any questions were answered, the experimenter explained that the session would be tape-recorded. During the session, the experimenter recorded each move announced by the subject, corrected any illegal moves, and induced the aloud and analytical subjects to verbalize.

RESULTS

Number of Excess Moves

In order to compare performance between the five and six disk problem, the number of excess moves was calculated. Excess moves are any moves beyond the minimum number of moves necessary for solution. An analysis of variance was performed with a 3(verbalization) X 2(pacing) X 2(problem size) completely randomized, factorial design (Kirk, 1963). The means and standard deviations for each group are presented in Table 1. The verbalization variable was significant, $F(2, 72) = 36.42$, $p < .01$. The analytical group required less excess moves, $F(1, 54) = 11.50$, $p < .01$, than the silent group. However, the silent group was not significantly different from the aloud group, $F(1, 54) = 1.62$, $p > .05$. As was expected, the problem size variable was significant $F(1, 72) = 220.94$, $p < .01$. The six disk problem ($\overline{X} = 94.76$, SD = 47.93) was more difficult than the five disk problem ($\overline{X} = 23.71$, SD = 18.19). The effect of the pacing variable was also significant, $F(1, 72) = 26.22$, $p < .01$, with the delayed condition requiring fewer excess moves than the self-paced condition. However, each main effect was also involved in an interaction.

Table 1. Group Means and Standard Deviations for Number of Excess Moves

	Problem Solving Tempo				
	Self-Paced		Delayed		
Group	5 disk	6 disk	5 disk	6 disk	Total
Analytical	$\overline{X}$ = 12.6 SD = 7.5	$\overline{X}$ = 67.3 SD = 30.6	$\overline{X}$ = 14.0 SD = 12.2	$\overline{X}$ = 31.0 SD = 16.0	$\overline{X}$ = 31.2 SD = 28.9
Silent	$\overline{X}$ = 28.0 SD = 16.1	$\overline{X}$ =121.4 SD = 28.4	$\overline{X}$ = 25.6 SD = 13.0	$\overline{X}$ = 81.9 SD = 18.0	$\overline{X}$ = 64.2 SD = 44.6
Aloud	$\overline{X}$ = 41.7 SD = 25.1	$\overline{X}$ =157.9 SD = 32.5	$\overline{X}$ = 20.4 SD = 10.2	$\overline{X}$ =109.1 SD = 14.0	$\overline{X}$ = 82.3 SD = 58.9
Total for Tempo	$\overline{X}$ = 71.5 SD = 58.4		$\overline{X}$ = 41.4 SD = 48.9		

The verbalization conditions interacted with the size of the problem, $F(2, 72) = 15.82$, $p < .01$. As shown in Figure 1, the difference between the verbalization groups is much larger for the six disk problem than for the five disk problem. On the five disk problem, the analytical group was better than the silent group, $F(1, 26) = 7.46$, $p < .02$. The analytical group was also superior to the thinking aloud group, $F(1, 26) = 7.08$, $p < .02$. However, the silent group did not differ significantly from the aloud group. For the six disk problem, all three verbalization groups differed significantly. The analytical group required fewer excess moves than the silent group, $F(1, 26) = 33.98$, $p < .01$. The aloud group required more excess moves than the silent group, $F(1, 26) = 6.07$, $p < .02$.

The problem size also interacted with the pacing conditions. Slowing the subjects down aided their problem solving on the six disk problem but not on the five disk problem. On the six disk problem, the delayed group ($\overline{X} = 74.00$, $SD = 36.14$) took fewer moves, $F(1, 40) = 29.14$, $p < .01$, than the self-paced group ($\overline{X} = 115.62$, $SD = 48.13$).

The results of the major analysis of variance are summarized in Table A1, Appendix A.

Stereotypic Moves

From a computer analysis of each move made by each subject, a measure of stereotypic behavior was obtained.

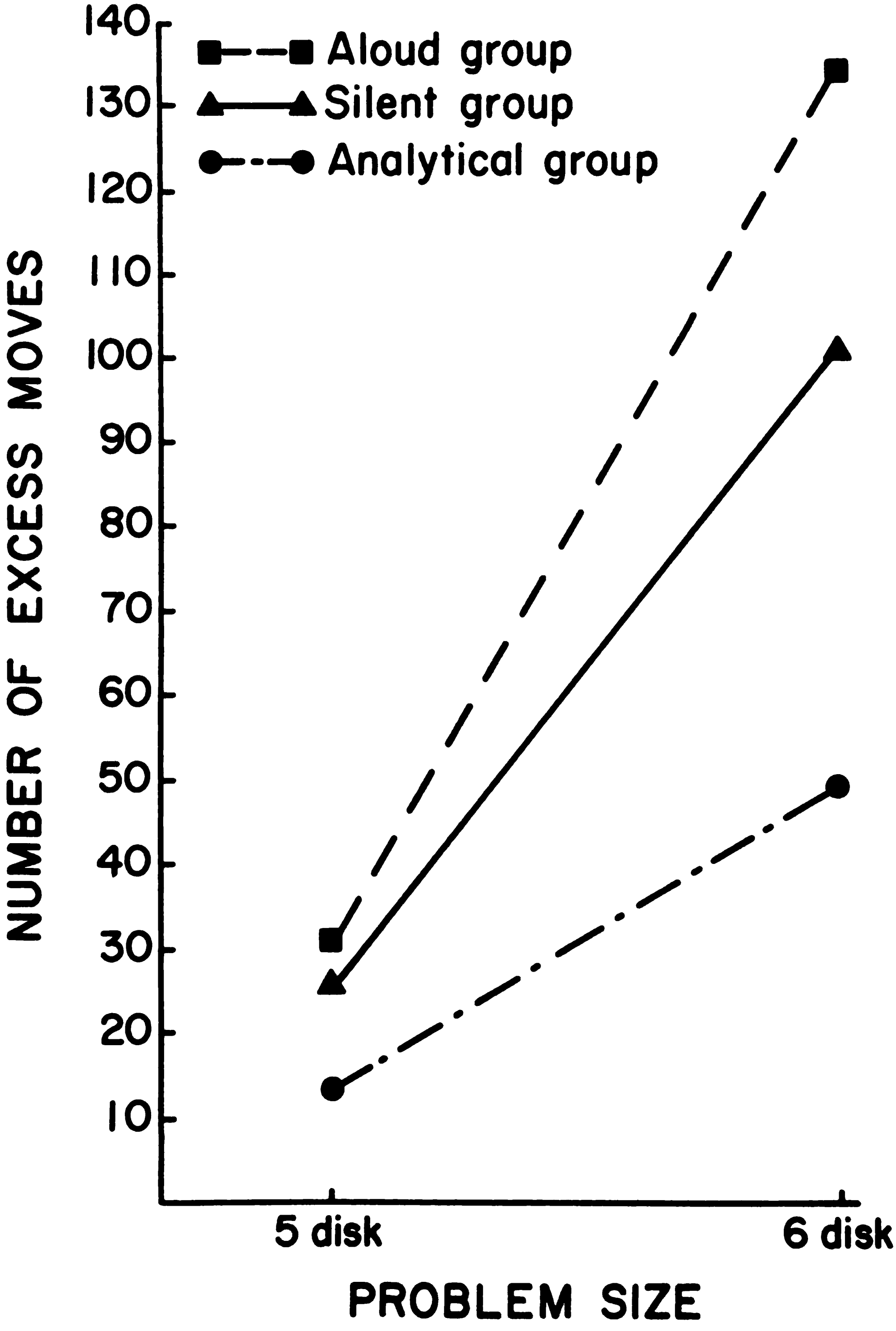

Fig. 1. The problem size by verbalization interaction for excess moves for both pacing conditions.

In order for a move to be considered as stereotypic, the move returned the problem solver to the same problem state for the third time and was part of a sequence of stereotypic moves involving two or more moves. The basic idea represented by stereotypic moves was to discover how often a problem solver returned to a previous problem state and then went through the same sequence of moves. There was a strong relationship between the number of stereotypic moves and the number of excess moves, $r = .75$. Any stereotypic move would result in an excess move; though of course not all excess moves were stereotypic moves.

Only the six disk problem resulted in stereotypic behavior. An analysis of variance was calculated for a 3(verbalization groups) X 2(pacing) factorial design. The means and standard deviations for each group are presented in Table 2. The verbalization variable was significant, $F(2, 36) = 10.88$, $p < .01$. The aloud group produced more stereotypic moves than either the analytical group or the silent group. The aloud group was significantly different than the silent group, $F(1, 26) = 10.78$, $p < .01$, and also the analytical group, $F(1, 26) = 7.34$, $p < .02$. However, the analytical and the silent groups did not differ significantly, $F(1, 26) = 0.23$, $p > .05$.

The pacing variable was also significant, $F(1, 36) = 6.90$, $p < .02$. The self-paced group ($\overline{X} = 4.71$, $SD = 4.98$)

Table 2. Group Means and Standard Deviations for Stereotypic Moves on the Six Disk Problem

Group	Problem Solving Tempo: Self-Paced	Problem Solving Tempo: Delayed	Total
Analytical	$\bar{X}$ = 2.85 SD = 3.98	$\bar{X}$ = 0.14 SD = 0.36	$\bar{X}$ = 1.52 SD = 3.13
Silent	$\bar{X}$ = 2.00 SD = 1.19	$\bar{X}$ = 1.85 SD = 1.73	$\bar{X}$ = 1.92 SD = 1.49
Aloud	$\bar{X}$ = 9.28 SD = 5.04	$\bar{X}$ = 4.57 SD = 4.40	$\bar{X}$ = 6.92 SD = 5.29
Total for Tempo	$\bar{X}$ = 4.71 SD = 4.98	$\bar{X}$ = 2.19 SD = 3.28	

produced more stereotypic moves than the delayed group ($\overline{X}$ = 2.19, SD = 3.28).

The results of the major analysis of variance are summarized in Table A2, Appendix A.

Time Per Move

An estimation of the seconds per move was obtained by dividing the total time to solution by the total number of moves. The time required to produce each move was an important factor in determining the effect of the analytical instructions. Earlier research had indicated that the analytical instructions slowed problem solving as far as time taken between moves. To study the effects of the amount of time involved between moves, some subjects were delayed while the other half of the subjects solved the problem at their own pace. A practical question, as far as the experimental manipulation was concerned, was whether the delayed condition actually slowed problem solving.

The data was analyzed using analysis of variance with a 3(verbalization groups) X 2(pacing) X 2(problem size) completely randomized factorial design (Kirk, 1963). The means and standard deviations for each group are presented in Table 3. The pacing manipulation was significant, $F(1, 72) = 191.82$, $p < .01$. The delayed group required more time per move than the self-paced group. Even though each subject was only required to wait seven seconds after

Table 3. Group Means and Standard Deviations for Time Per Move (in seconds)

	Problem Solving Tempo				
	Self-paced		Delayed		
Group	5 disk	6 disk	5 disk	6 disk	Total
Analytical	$\bar{X}$ = 9.2 SD = 2.4	$\bar{X}$ = 12.2 SD = 1.9	$\bar{X}$ = 16.6 SD = 4.3	$\bar{X}$ = 18.1 SD = 5.4	$\bar{X}$ = 14.1 SD = 5.2
Silent	$\bar{X}$ = 4.9 SD = 0.9	$\bar{X}$ = 6.6 SD = 1.7	$\bar{X}$ = 16.4 SD = 3.6	$\bar{X}$ = 17.3 SD = 3.0	$\bar{X}$ = 11.3 SD = 6.2
Aloud	$\bar{X}$ = 6.1 SD = 1.7	$\bar{X}$ = 7.7 SD = 1.6	$\bar{X}$ = 17.7 SD = 2.8	$\bar{X}$ = 19.1 SD = 3.4	$\bar{X}$ = 12.6 SD = 6.3
Total for Tempo	$\bar{X}$ = 7.78 SD = 2.99		$\bar{X}$ = 17.54 SD = 3.96		

completing a move, requiring them to pause produced almost ten additional seconds delay. Also, the verbalization variable was significant, $F(2, 72) = 5.08$, $p < .01$. The analytical instructions did slow down problem solving activity. The analytical group ($\overline{X} = 14.06$ sec., SD = 5.16 sec.) required more time per move than either the silent group or the aloud group. Only the difference between the analytical group and the silent group was significant, $F(1, 54) = 6.17$, $p < .02$. The size or the problem was also significant, $F(1, 72) = 5.63$, $p < .05$. Problem solvers working on the six disk problem took more time per move ($\overline{X} = 13.50$, SD = 5.91) than those solving the five disk problem ($\overline{X} = 11.82$, SD = 5.98 sec.).

The verbalization variable interacted with the pacing variable. The interaction is illustrated in Figure 2. The verbalization groups took significantly different times per move when the subjects were self-paced. However, when the subjects were delayed, the subjects required about the same time per move. Tukey's HSD test showed that for the self-paced condition, the analytical group required more time per move than the aloud group, $q(2, 54) = 4.45$, $p < .01$. The difference between the analytical group and the silent group was also significant, $q(2, 54) = 4.96$, $p < .01$. However, the difference between the aloud group and the silent group was not significant. None of the differences

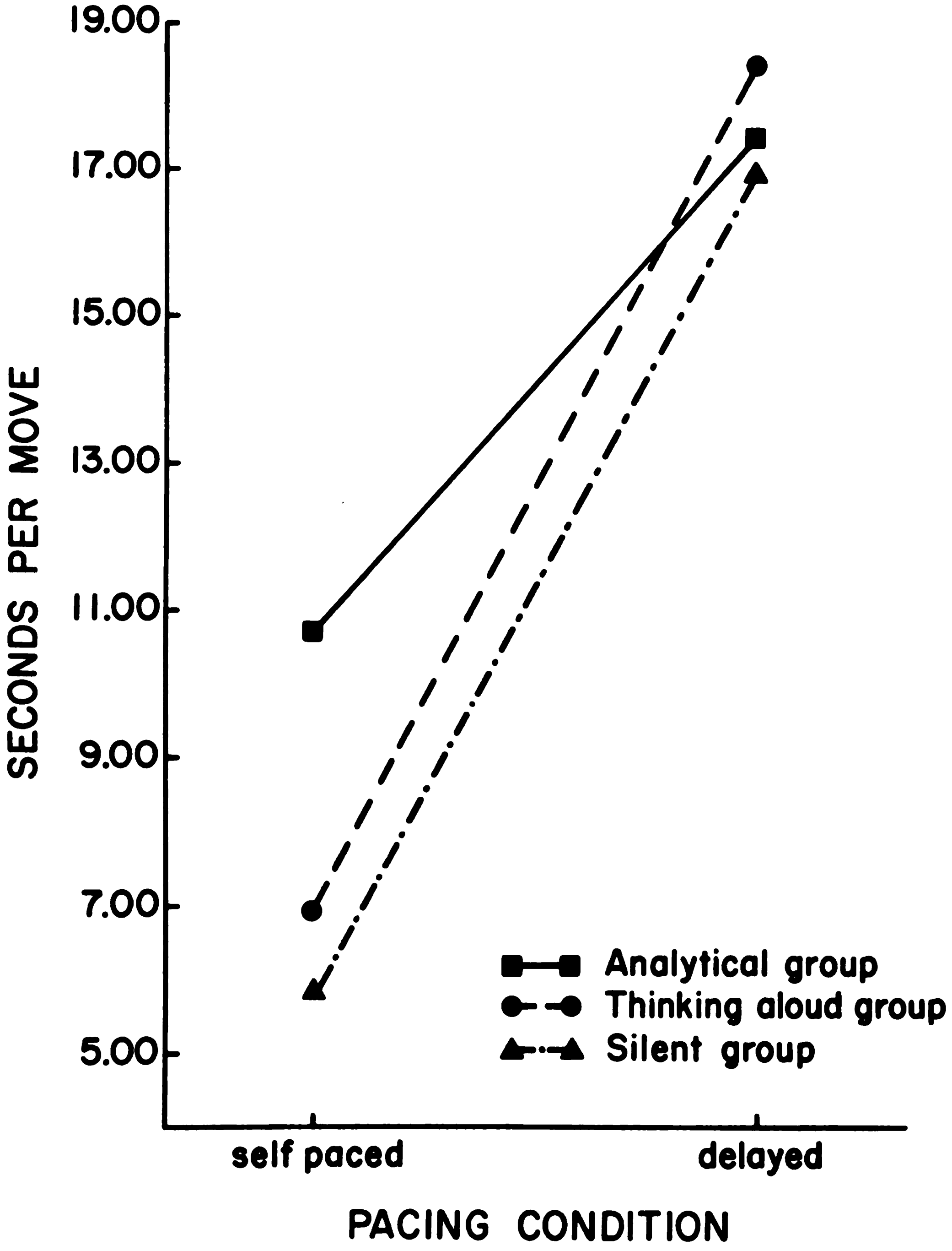

Fig. 2. Interaction between pacing and verbalization variables for time per move for the five and six disk problems.

between the verbalization groups were significant for the delayed condition.

The results of the major analysis of variance are shown in Table A3, Appendix A.

Episode Analysis

The essential aspects of the solution of the disk transfer problem involve moving the larger disks. In the six disk problem, for example, a good idea of the subject's progress toward the solution can be traced by how the subject moved the three largest disks. Each move involving one of the larger disks defines an episode. During an episode, the problem solver works on more obvious aspects of the problem--moving the three smaller disks. Most subjects have little problem with moving the three smallest disks to the desired position. On the three disk problem, many subjects will solve the problem in the minimum number of moves. However, the boundary of an episode is reached when the problem solver moves a larger disk.

Defining episodes in this manner can simplify the analysis because a large number of moves involving the smaller disks can be ignored. At the same time, differences between groups can be reflected by the sequence of episodes during their solution of the problem.

The major episodes for the six disk problem are represented in Figure 3. Each transition from one episode state

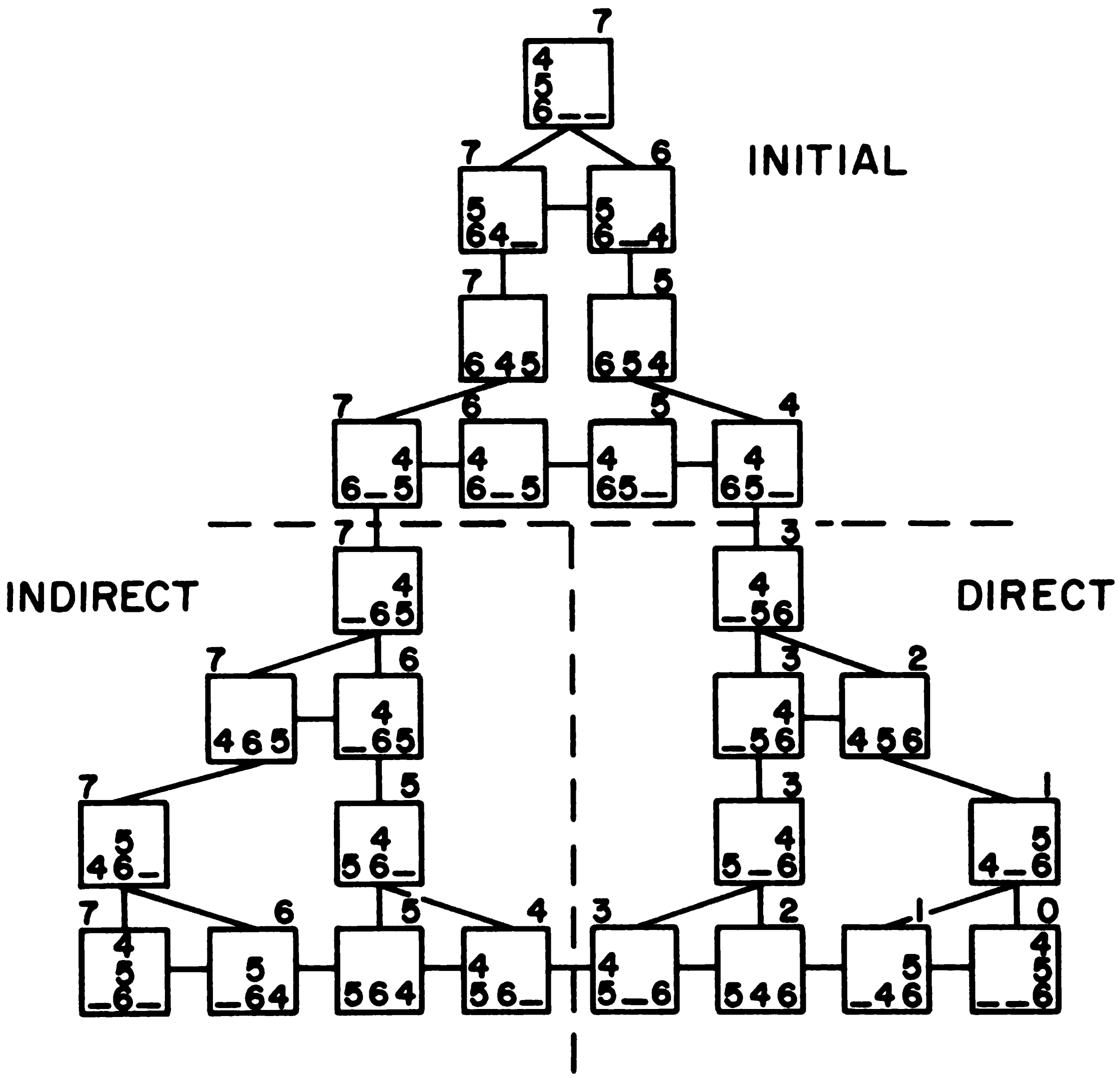

Fig. 3. Episodes for the six disk problem.

to another involves a minimum of eight moves, unless the subject merely moves directly back to the previous episode. Each square in the figure represents a possible configuration of the three largest disks on the three pegs, with the start peg on the left and the goal peg on the right. In the top square of the figure, the three largest disks, numbered 4, 5, and 6, are all on the start peg, with the middle and goal pegs empty of the major disks. The sequence of episodes has been divided into three parts, with the top part called the initial portion of the episode space. The direct portion of the figure, on the lower right, contains the episode necessary for solution, with all the major disks on the goal peg. The indirect part is the only portion that does not include episodes that are part of the ideal solution. All episodes on the left side of the triangular figure represent episodes that are as far away from the solution as the initial episode. The episodes on the right side of the triangle are episodes involved in the ideal solution.

In order to evaluate the difficulty of moving from episode to episode, the average number of moves per episode was obtained for each subject. Then an analysis of variance was calculated for a 3(verbalization conditions) X 2(pacing groups) design. For the six disk problem, there was no difference between the verbalization groups as far as the number of moves per episode. Each verbalization group experienced the same degree of difficulty as far as moving the larger disks. The only significant effect for the number of moves per episode involved the pacing variable. The self-paced group ($\overline{X}$ = 11.66, SD = 1.54) produced significantly more moves per episode, $F(1, 36) = 7.08$, $p < .01$, than the delayed group ($\overline{X}$ = 9.72, SD = 4.38).

The number of episodes involved in solving the six disk problem was highly correlated with the excess number of moves ($r = .86$). The group differences for the six disk problem were the same for both dependent measures. The analysis of variance for the number of episodes is presented in Appendix A-4.

The major interest in the episode analysis was to trace the general sequence of episodes involved in reaching a solution. The verbalization groups differed according to the portions of the episode space (see Figure 3) involved in their solutions. None of the solutions for subjects in the analytical group included episodes from the indirect portion, but 42 per cent of the silent group and 92 per cent of the thinking aloud group produced indirect solutions. The aloud group used the indirect solution more often than the silent group, $X^2(3) = 8.00$, $p < .05$. The difference between the silent group and the analytical group approached significance, $X^2(3) = 7.26$, $p < .10$.

An episode analysis for the five disk problem was less meaningful. The same definition of an episode was logical, but the episode space was much smaller. Many subjects solved the five disk problem in only a few episodes. The difference between the verbalization groups was best reflected by the number of episodes involved in solution, and that dependent variable was highly correlated with excess number of moves ($r = .78$). The difference in number of episodes again reflected that the analytical group correctly positioned the larger disks during solution while the silent and thinking aloud groups moved them about superfluously.

Ratio Score

A major problem of interpretation in comparing the five disk to the six disk problem concerns the possible cost of making an error. As an illustration, say the problem solver obtains the problem state where all the disks were moved from peg A to peg B rather than to peg C. To reach this intermediate state requires the minimum number of moves, so any further moves will be excess moves. The six disk problem will require 63 additional moves, whereas the five disk problem will only require 31 additional moves. Hence, the difference between difficulty of the six disk problem and the smaller five disk problem can be explained by the consequences of making a mistake. The extra moves are needed in the six disk problem to correct the mistake.

The cost of making a mistake was how Bourne, Ekstrand, and Dominowski (1971) accounted for the effects of problem size on the difficulty of the disk problem.

An alternative to dealing with excess number of moves to solution involves equating the different problem sizes by a ratio: total number of moves/minimum number of moves to solution. The proportion equates the problems as far as the moves necessary to correct for a mistake. Regardless of the state of the problem in terms of the search space, the minimum number of moves (or less) is all that is needed in order to reach the solution state. To return to the illustration, regardless of whether the problem solver was solving the five or six disk problem, the problem solver would obtain a ratio score of 2.0 if he first reached the problem state where all the disks were moved to peg B rather than peg C and then managed to move the disks to peg C in the minimum number of moves. Using the number of excess moves as the denominator establishes a "recovery unit" as far as mistakes and corrections are concerned.

The two different measures, number of excess moves and the ratio score, were highly correlated ($r = .94$). As is shown in Figure 4, the relationship between either the five or six disk problem and its corresponding ratio score and excess number of moves is simply a linear transformation.

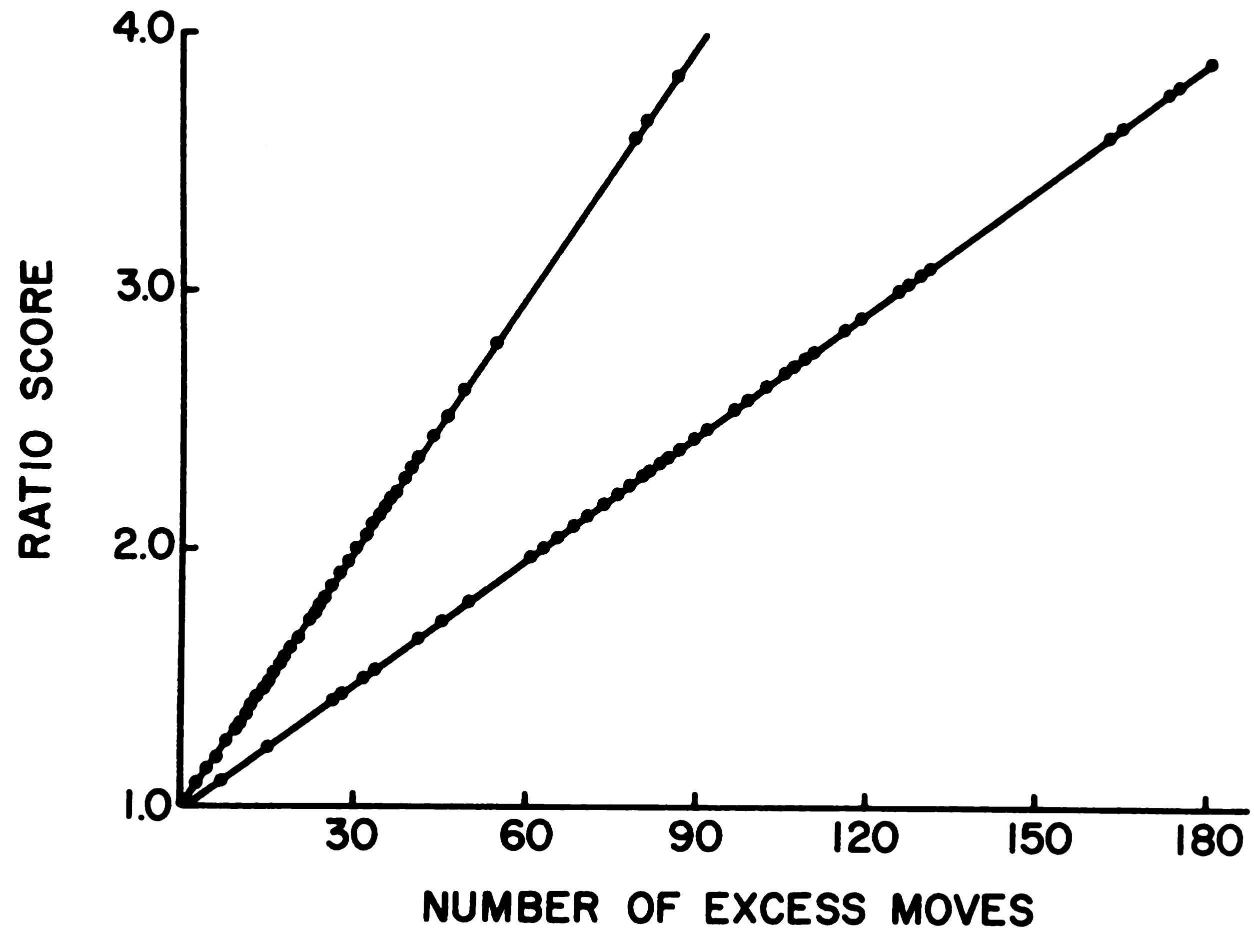

Fig. 4. Scatter plot for relationship between excess moves and ratio score.

Empirically, there is little difference between the ratio score and number of excess moves. Analysis of variance was calculated and results were identical to those presented for excess number of moves. The analysis is presented in Appendix A-5.

Total Time to Solution

The amount of time required to complete the problem was not of primary interest. The instructions stressed that time was not a major factor, and the subject should try to solve the problem in as few moves as possible. Total time also would not indicate whether thinking aloud would actually affect the efficiency of problem solving. Thinking aloud could slow down problem solving without actually resulting in any different problem solving processes. The solution time could be slower simply because verbalization slowed each move.

There was a highly significant correlation between total time and excess moves, $r = .77$. An analysis of variance also showed identical results (see Table A6). The means and standard deviations for each group are presented in Table A7.

DISCUSSION

The major purposes of the present study were (1) to study the effect of different verbalization instructions on problem solving efficiency and the processes involved in the solution and (2) to determine how the size of the problem and the time between moves modified the effects of verbalization. Briefly, the results indicated that (1) analytical instructions aided problem solving while thinking aloud impaired problem solving, (2) the effect of verbalization instructions increased as the size of the problem increased, and (3) the effect of verbalization instructions was independent of the amount of time between problem solving moves. Slowing down the subjects did decrease the number of excess moves, but only for the larger problem.

Analytical Verbalization

The effect of verbalization was dependent upon what the subject was asked to verbalize. The analytical instructions required that the subjects produce reasons for each move as they solved the disk transfer problem. The hypothesis that analytical verbalization would aid problem solving was supported. Previous research had indicated that analytical verbalization helped (Gagné and Smith, 1962; Davis, Carey, Foxman, and Tarr, 1968; Wilder and Harvey, 1971).

Correspondence to Previous Research

Earlier studies (Gagné and Smith, 1962; Wilder and Harvey, 1971) had indicated that the analytical group took more time to produce each move. The extra time alone could have accounted for the analytical group's superior performance when compared to a silent group. However, the present study showed that even when all the verbalization groups were delayed, and each group consumed the same amount of time per move, the analytical group was still superior to both the silent and the thinking aloud groups. The effect of the analytical instructions was not merely to emphasize a time-accuracy trade-off in problem solving.

The present experiment also demonstrated that the effect of the analytical instructions was not dependent upon a transfer paradigm. Previous researchers investigating the use of analytical instructions (Gagné and Smith, 1962; Davis, Carey, Foxman, and Tarr, 1968; Wilder and Harvey; 1971) had required their subjects to begin with a simple disk transfer problem, involving three disks, and increased the next problem by one disk, until the subjects had solved the six disk problem. Gagné and Smith's experiment has been cited to support advantages of discovery learning.

The present experiment indicated that an analytical approach to problem solving was in itself an advantageous strategy without any previous learning necessary.

The effect of the analytical instructions was more than just task facilitation. Performance was facilitated in terms of excess moves to solution, but a difference between the verbalization groups also was indicated in terms of stereotypic behavior and in the episode analysis.

Analytical Instructions Induce Active Processing

Requiring a person to state reasons for each move in solving a problem improves performance. In order to produce plausible reasons, the problem solver was forced to evaluate moves, consider different moves, and seek relationships between moves. The analytical group was able to correctly anticipate the most direct solution to the problem. As the episode analysis showed, analytical problem solvers moved from the initial portion of the search space to the direct search space. The reasons verbalized by the analytical group indicated that they were solving the problem by establishing subgoals. The episode analysis also indicated that the analytical group moved the largest disks to the best positions in terms of the ideal solution.

Several observations on good and poor problem solvers led to the expectation that active processing, involved in following analytical instructions, would improve problem solving. A study by Goor and Summerfield (1975) indicated that problem solvers who scored high on divergent thinking tasks were able to maintain productive work on convergent problem solving tasks. Subjects who scored low on the divergent thinking tasks showed a more rapid decrease of productive work and consequently performed worse on the problem solving tasks. Bloom and Broder (1950) found that low aptitude college students tended to engage in "one shot thinking" rather than attacking a problem with a thorough sequential analysis to get the solution. Both studies indicated that the good problem solver constantly attacked the problem in terms of producing ideas and hypotheses. Encouragingly, analytical verbalization has been used as a training procedure to improve problem solving. Whimbley (1976) has reported that requiring low aptitude students to solve problems using analytical verbalization has led to improved grades in school, on SAT and GRE scores. Using analytical verbalization as a training procedure to improve problem solving has been successful.

Thinking Aloud

Contrary to the claim of some researchers (Goor and Summerfeld, 1975; Benjafield, 1969) requiring a subject to talk while solving a difficult problem does not always improve performance. Newell and Simon (1972) claimed there was a basic difference between asking a subject to think aloud and asking a subject to verbalize analytically. The present study demonstrated a difference between the two types of verbalization. However, the question remains whether thinking aloud or analytical verbalization gives a better picture of problem solving processes. Both methods have been used for process tracing. Obviously, neither is ideal. When compared to a group which solved the problems silently, the thinking aloud group performed worse on the problem while the analytical group performed better. Contrary to the claims of Newell and Simon, though, thinking aloud did have an effect on the problem solving processes. When Newell and Simon (1972) tested the hypothesis that thinking aloud affected problem solving, they examined the number of steps to solution and the directions of search in terms of the logical expressions generated by the subjects. Their experiment has been criticized on methodological grounds because of the nonrandom selection of subjects for the groups, the small number of subjects, and the lack of statistical tests. The present experiment also examined the problem solving in terms of excess moves, stereotypic behavior, and an episode analysis (the direction of search for a solution). The results were consistent with the hypothesis that thinking aloud would interfere with the problem solving efforts of the subject.

Stereotypic Behavior

The performance of the thinking aloud group on the six disk problem was hindered by stereotypic behavior. When a subject is working with a complex problem, Kahneman (1974) has claimed that a secondary task (such as thinking aloud) would narrow the focus of attention. Stereotypic behavior was defined as repeating the same sequence of moves three or more times. Subjects in the delayed condition produced fewer stereotypic moves, so the time pressure definitely increased the tendency to use set responses.

Methodological Questions

The differences between the silent and thinking aloud groups in terms of problem solving efficiency and stereotypic behavior raises methodological questions. Requiring subjects to think aloud while solving a problem has been a traditional method used to trace the actual problem solving processes involved in problem solving. The method has become more popular with the ascendance of information processing models of problem solving. However, the results of the present study raise questions of internal and external validity of the theories that are formulated from verbal protocols.

Newell and Simon (1972) have been the pioneers as far as using verbal protocols, collected from subjects thinking aloud, to build models of problem solving. The models can be described as a hierarchy of production rules. A production rule consists of a condition which must be satisfied, and a consequent action. In problem solving, the condition consists of a certain state of knowledge about the problem, and the action consists of a change to be made on the present state of the problem. For a simplified example, if a certain configuration of pieces existed on a chess board, the condition of a production rule could be satisfied, and the resulting action would be a certain chess move. Once the move was completed, the state of the problem would have changed.

What effect would stereotypic behavior have on the models constructed from verbal protocols? The production rules are organized into a hierarchical structure called a production system. The organization of a system depends upon the sequence of the rules found in the verbal protocols. Newell and Simon develop a production system by formulating a matrix and tracing the sequential dependence among the production rules derived from verbal protocols. Therefore, any stereotyped behavior will have the possible effect of establishing an invalid hierarchy. The errors in stereotypic behavior are the repetition of certain sequences of production rules not characteristic of the silent subject. The basic data, used to build the production system, is distorted by the requirement to think aloud, and the final models reflect the basic distortions.

The episode analysis indicated that the thinking aloud group was arriving at a solution differently than the silent group. When the aloud group was compared to the silent group on the six disk problem, the episode analysis showed that the thinking aloud group was more likely to enter the indirect search space. If an information processing model was constructed using the verbal protocols from the thinking aloud group, the model would solve the problem by going through the indirect search space. The production system, constructed from protocols, would differ from a model of the silent group which would be more likely to avoid the indirect search space.

According to Newell and Simon (1972), a distinction can be made between the demands of the task and the psychology of the subject. To the degree that behavior is exactly required by the task, it gives the researcher information about the task environment. But to the extent the behavior departs from perfect rationality, the researcher gains information about the characteristics of the subject, about the nature of the psychological processes that limit his problem solving performance. Unfortunately, when a subject is required to think aloud while solving a difficult problem, departure from perfect rationality is caused by psychological processes involved in solving the problem and by processes involved in a communications task. In order to build a valid model from verbal protocols, it would be necessary to eliminate or modify the processes affected by verbalization.

The question remains whether the differences between the silent and thinking aloud groups would involve any major distinctions in terms of broad generalities used to characterize problem solving. Certainly Newell and Simon's concepts of the problem space, the use of different heuristics such as means-end analysis, the basic limitations of human information, etc., are descriptive of problem solving regardless of the quality of the problem solving. However, much of the power of information processing models involve the detailed description of problem solving processes. The present study showed that thinking aloud affected the finer level of analysis in terms of excess moves and stereotypic moves and also the more general analysis of problem solving in terms of episode analysis.

Problem Size

The size of the problem had the expected effect on the quality of the solution. The hypothesis was that as the size of the problem increased, the number of excess moves required for solution would also increase. The five disk

problem required fewer excess moves, less time per move, and less stereotypic behavior than the six disk problem. The increase in difficulty with the increase in problem size has been found in a number of studies with a variety of different problems (Davis, 1967; Neimark and Wagner, 1964; Solley and Snider, 1958) As the size of the problem increases, one would expect that a problem would require more time to solve it or more moves because the problem involves more problem elements in the solution or more minimum moves to solution. However, the difficulty caused by problem size goes beyond any discrepancy in materials or moves involved in solution. With the disk transfer problem, the dependent measure of excess moves equates the different sized problems as far as the difference in moves to solution is concerned. Even with this basic correction, a number of researchers (Ewert and Lambert, 1932; Cook, 1937; Gagné and Smith, 1962) found that as the number of disks increased, the difficulty of the problem increased exponentially.

Problem Size and the Cost of an Error

With the disk transfer problem, the exponential increase in problem difficulty as problem size increases cannot be explained as the cost of recovering from an error as suggested by some researchers (Bourne, Ekstrand, and Dominowski, 1971). In some problems, such as the spy problem (Hayes, 1965), the search space involves blind alleys, so that certain mistakes lead to retracing previous moves in order to recover. With some problems, as the problem size increases, the number and/or length of the blind alleys increases. This is not true of the disk transfer problem. The search space does not involve blind alleys, and no problem state in the search space is more than the minimum number of moves from solution. However, as the search space expands with the larger problem size, the number of problem states that are the minimum number of moves away from the goal also increases. When the problem solver reaches these problem states, the move has failed to make any progress toward the goal. Using the ratio score neutralizes the number of excess moves needed to recover from any single or series of mistakes, but even when the ratio score was used as the dependent variable, the difficulty of finding a solution increased with problem size. The size of the search space increases at an exponential rate as the size of the disk transfer problem increases. With larger sized problems, the need for planning moves becomes more necessary and difficult, and the possibility of making mistakes or unnecessary moves also increases.

Verbalization Effects Depend Upon Problem Size

The problem size interacted with the verbalization variable. The generality of the verbalization results depends upon the size of the problem the subject is asked to solve. Problem size is a major determinant of problem difficulty and is also an important characteristic for developing a problem typology. The analytical group was superior to both the silent and thinking aloud groups for both the five and six disk problems. However, the difference between the groups was more pronounced for the six disk problem. The analytical group used the advantage of an active problem solving approach. The results were consistent with the hypothesis that the analytical group's ability to reduce the effective search space by planning would become more critical as the problem size increased.

The difference between the thinking aloud and silent groups was significant only for the six disk problem. Also, the thinking aloud group displayed stereotypic behavior only on the six disk problem. The interaction with problem size is consistent with the hypothesis that a limited capacity for performing two mental tasks (problem solving and communicating) would be exceeded only when the problem was sufficiently complicated or mentally taxing (Kahneman, 1974). Most information processing theorists have worked with complicated problems while requiring their subjects to think aloud, such as chess, cryptarithmetic, or logic problems. Simple problems are often less useful for collecting good verbal protocols because important problem solving processes occur too quickly for verbal description.

Problem Solving Tempo

The general hypothesis on problem solving tempo was confirmed. Slowing down subjects between moves on the disk transfer problem improved the quality of performance. Requiring a subject to stop and think is by itself an effective strategy. Johnson (1972) has suggested that the immediate production (and acceptance) of solutions interferes with the chances of cognitive reorganization or even identifying alternative response patterns. The instructions each subject received did emphasize that the number of moves was important and not to worry about the amount of time the solution involved. Certainly the general atmosphere of a testing situation would impose a time pressure, but it is also possible that in general, subjects do not take enough time between steps involved in complex problems. Any training program in problem solving should emphasize a thorough appraisal at the beginning by the problem solver and possible check points afterwards.

Contrary to prediction, the different problem solving tempos did not interact with the verbalization variable. Regardless of the verbalization instruction, slowing down the problem solver did not effect the basic differences between the verbalization groups.

There was a significant interaction between the tempo and the problem size. Slowing down the problem solver improved problem solving only on the six disk problem. The six disk problem involves more alternatives in its larger search space, and correspondingly, there is a greater probability of making mistakes. If the subjects in the delayed condition used the extra time to consider alternatives and to plan sequences of moves, then the six disk problem also presented more possibilities to prevent errors and mistakes by careful analysis.

LIST OF REFERENCES

Benjafield, J. Evidence that "thinking aloud" constitutes an externalization of inner speech. Psychonomic Science, 1969, 15, 83-84.

Bloom, B. S. and Broder, L. Problem solving processes of college students. University of Chicago, 1950.

Bourne, L., Ekstrand, B. and Dowminowski, R. L. The psychol-of thinking. Englewood Cliffs: Prentice Hall, 1971.

Byers, J. L. and Davidson, R. E. The role of hypothesizing in concept attainment. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 595-600.

Carpenter, J. A., Moore, O. K., Snyder, C. R. and Linsanky, E. S. Alcohol and Higher-order problem Solving. Quarterly Journal of Studies on Alcohol, 1961, 22 183-222.

Claparède, Edourand. La psychologie de l'intelligence. Scientia, 1917, 22, 353-368.

Cook, T. W. Amount of material and difficulty of problem solving. II. The disc transfer problem. Journal of Experimental Psychology, 1937, 20, 288-296.

Dansereau, D. An information processing model of mental multiplication. Unpublished doctoral dissertation, Carnegie-Mellon University, 1969.

Dansereau, D. and Gregg, L. W. An information processing analysis of mental multiplication. Psychonomic Science, 1966, 6, 71-72.

Davis, G. A. Detrimental effects of distraction, additional response alternatives, and longer response chains in solving switch-light problems. Journal of Experimental Psychology, 1967, 73, 45-55.

Davis, J. G., Carey, M. H., Foxman, P. N., and Tarr, D. B. Verbalization, experimenter presence, and problem solving. Journal of Personality and Social Psychology, 1968, 8, 299-302.

De Groot, Adriaan D. Thought and choice of chess. The Hague: Mouton, 1965.

Dominowski, R. L. Required hypothesizing and the identification of unidimensional, conjunctive, and disjunctive concepts. Journal of Experimental Psychology, 1973, 100, 387-394.

Duncan, C. P. Effect of instructions and information on problem solving. Journal of Experimental Psychology, 1973, 100, 387-394.

Duncker, K. On problem solving. Psychological Monographs, 1945, 58, No. 270.

Egan, D. and Greeno, J. G. Theory of rule induction: knowledge acquired in concept learning, serial pattern learning, and problem solving. In Gregg, L. W. (ed.) Knowledge and cognition. Potomac, Maryland: Lawrence Erlbaum, 1974.

Ewert, P. H. and Lamber, J. F. Part II: The effect of verbal instructions upon the formation of a concept. Journal of General Psychology, 1932, 6, 400-413.

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

Gardner, Martin. The Scientific American book of mathematical puzzles and diversion. New York: Simon & Schuster, 1959.

Goor, A., and Sommerfeld, R. E. A comparison of problem solving processes of creative students and noncreative students. Journal of Educational Psychology, 1975, 67, 495-505.

Hayes, J. R. Problem typology and the solution process. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 371-379.

Johnson, D. M. A systematic introduction to the psychology of thinking. New York: Harper & Row, 1972.

Kahneman, D. Attention and Effort. Englewood Cliffs: Prentice-Hall, 1973.

Kendler, H. H. and Kendler, T. S. Vertical and horizontal processes in problem solving. Psychological Review, 1962, 69, 1-16.

Luchins, A. S. Mechanization in problem solving: The effect of Einstellung. Psychological Monographs, 1942, 54, No. 248.

Maier, N. R. F. An aspect of human reasoning. British Journal of Psychology, 1933, 24, 144-155.

Mayzner, M. S., Tresselt, M. E. and G. Helbock. An exploratory study of mediational responses in anagram problem solving. Journal of Psychology, 1964, 57, 262-274.

Nance, R. D. and Sinnot, W. Level of difficulty and consistency of performance in anagram solving. Psychological Reports, 1964, 14, 895-898.

Neimark, E. D. and Wagner, H. Information gathering in diagnostic problem solving as a function on the number of alternative solutions. Psychonomic Science, 1964, 1, 329-330.

Neisser, U. The multiplicity of thought. British Journal of Psychology, 54, 1963, 1-14.

Newell, A. and Simon, H. A. Human problem solving. Englewood Cliffs: Prentice Hall, 1972.

Nilsson, N. F. Problem solving methods in artificial intelligence. New York: McGraw-Hill, 1971.

Peterson, L. R. Concurrent verbal activity. Psychological Review, 1969, 76, 376-386.

Ray, W. S. Mild stress and problem solving. American Journal of Psychology, 1965, 78, 227-234.

Reed, S. K., Ernst, G. W. and Banerje, R. The role of analogy in transfer between similar problem states. Cognitive Psychology, 1974, 6, 436-450.

Rokeach, M. The effect of perception time upon rigidity and concreteness of thinking. Journal of Experimental Psychology, 1950, 40, 206-216.

Scheerer, M. Problem-solving. Scientific American, 1963, 208, 118-128.

Solley, C. M. and Snyder, F. W. Information processing and problem solving. Journal of Experimental Psychology, 1958, 55, 384-387.

Whimbley, A. You can learn to raise your I. Q. score. Psychology Today, 1976, 9, 27-29, 84-85.

Wickelgren, W. A. How to solve problems. New York: Freeman, 1974.

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

APPENDICES

APPENDIX A

TABLES

Table A1. Excess Moves: Analysis of Variance

Source	SS	df	MS	F
Verbalization (V)	37556	2	18778	39.14**
Pronlem Size (S)	106003	1	106003	220.93**
Pacing (P)	12580	1	12580	26.22**
V x S	15663	2	7831	16.32**
V x P	1207	2	603	1.26
S x P	6103	1	6103	12.72**
V x S x P	117	2	58	.12
Within Cell	34542	72	479	

** $p < .01$

Table A2. Stereotypic Moves (Six Disk Problem) Analysis of Variance

Source	SS	df	MS	F
Verbalization	275.54	2	137.77	10.88**
Pacing	87.38	1	87.38	6.90*
Verb. x Pacing	16.27	2	8.14	.64
Within Cell	455.71	36	12.66	

** $P < .01$ * $p < .05$

Table A3. Time Per Move: Analysis of Variance

Source	SS	df	MS	F
Verbalization (V)	105.99	2	53.00	5.08**
Pacing (P)	2000.72	1	2000.72	191.82**
Problem Size (S)	58.70	1	58.70	5.63@
V x P	102.98	2	51.49	4.94*
V x S	4.12	2	2.06	.20
P x S	4.80	1	4.80	.46
V x P x S	.65	2	.33	.03
Within Cell	750.97	72	10.43	

** $p < .01$ * $p < .02$ @ $p < .05$

Table A4. Number of Episodes for Six Disk Problem: Analysis of Variance

Source	SS	df	MS	F
Pacing	48.21	1	48.21	6.68*
Verbalization	358.34	2	179.17	24.82**
Pacing X Verb.	9.57	2	4.80	0.66
Within Cell	260.00	36	7.22	

** $p < .01$ * $p < .05$

Table A5. Ratio Score: Analysis of Variance

Source	SS	df	MS	F
Verbalization (V)	12.09	2	6.05	24.47**
Problem Size (S)	11.88	1	11.88	48.10**
Pacing (P)	3.29	1	3.29	13.32**
V x S	3.15	2	1.58	6.38**
V x P	1.51	2	0.76	3.08
S x P	1.12	1	1.12	4.53*
V x S x P	0.40	2	0.20	0.81
Within Cell	17.78	72	0.25	

** $p < .01$ * $p < .05$

Table A6. Total Time: Analysis of Variance

Source	SS	df	MS	F
Verbalization (V)	1226.3	2	613.2	16.13**
Pacing (P)	582.5	1	582.5	15.32**
Problem Size (S)	5116.9	1	5116.9	134.65**
V x P	437.3	2	218.7	5.75**
V x S	684.9	2	342.5	9.01**
P x S	98.5	1	98.5	2.59
V x P x S	166.5	2	83.2	2.19
Within Cell	2736.6	72	38.0	

** $p < .01$

Table A7. Group Means and Standard Deviations for Total Time (in Minutes)

	Problem Solving Tempo			
	Self-paced		Delayed	
Group	5 disk	6 disk	5 disk	6 disk
Analytical	$\bar{X}$ = 2.0	$\bar{X}$ = 14.2	$\bar{X}$ = 4.3	$\bar{X}$ = 9.7
	SD = 1.4	SD = 7.2	SD = 4.1	SD = 5.9
Silent	$\bar{X}$ = 3.3	$\bar{X}$ = 13.2	$\bar{X}$ = 7.4	$\bar{X}$ =26.9
	SD = 2.6	SD = 3.9	SD = 4.6	SD = 9.5
Aloud	$\bar{X}$ = 4.7	$\bar{X}$ = 21.8	$\bar{X}$ = 6.4	$\bar{X}$ =34.9
	SD = 3.9	SD = 8.0	SD = 3.7	SD = 8.1

APPENDIX B

GENERAL INSTRUCTIONS

APPENDIX B

GENERAL INSTRUCTIONS:

You see the problem on the table consists of three pegs, lettered A, B, and C, and a pile of disks on peg A. The disks are numbered from 1 to 5 (6) with the largest on the bottom, the next largest next to the bottom and so on, with the smallest on top. Your task is to move all the disks from peg A to peg C in the least number of moves possible so that they will be in the same order. You may move only one disk at a time. You may move only disks that are on top of a pile or that are alone. You must not place a larger disk on top a smaller. A disk moved from one of the pegs must be placed on another disk. There is no time limit, so do not hurry your moves. Remember, the object is to solve the problem in as few moves as possible.

<u>In addition for the analytical group</u>:

State out loud a full reason for each move as completely as you can. If you are not certain of a reason for a specific move do not hesitate to say so. However, try to produce a reason for each move.

For the thinking aloud group:

As you work on the problem, think aloud as you work, saying all thoughts that come to mind. Try to produce a steady stream of talking. As long as you are thinking, you should try to verbalize those thoughts.

All subjects:

When you have completed a move, I will press this button to turn on this light. Wait until the light goes out before you complete your next move.