DOCUMENT RESUME

ED 205 401 SE 035 465

AUTHOR	Mayer, Richard E.
TITLE	Different Solution Procedures for Algebra Word and Equation Problems. Technical Report Series in Learning and Cognition, Report No. 80-2.
INSTITUTION	California Univ., Santa Barbara. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	80
GRANT	NIE-G-78-0162
NOTE	66p.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Algebra: *Cognitive Processes: *Cognitive Style: College Mathematics: *Educational Research: Higher Education: *Mathematics Education: *Problem Solving: Secondary Education: Secondary School Mathematics
IDENTIFIERS	Equations (Mathematics): *Heuristic Models: Heuristics: *Mathematics Education Research: Word Problems (Mathematics)

ABSTRACT

The development of a theory of algebraic problem solving is discussed. The domain of the investigation is restricted to single sentence word problems. There were three experiments, the first involving 42 college students from the University of California, Santa Barbara. These subjects were divided into two groups, half in an equation group and half in a word group. The results of experiment one suggested that the solution process used by subjects in the equation group involves planning, while the translation and solution process of the word group does not show the same sort of planning. A second experiment was conducted using another group of 42 students, with half writing a numerical answer for each word problem, and the other writing an equation for each problem. Both groups displayed a pattern of performance similar to that displayed by the word group in in experiment one. A supplemental study was conducted to provide additional data on the problem-solving process of subjects, through interviews of eight college students. Evidence was provided that equation subjects engage in goal stacking. Results conflicted with a two-stage model of problem solving consisting of separate translation and solution processes. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED205401

TECHNICAL REPORT
SERIES IN LEARNING AND COGNITION

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
RICHARD E. MAYER

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

Department of Psychology
University of California
Santa Barbara, California 93106

035 465

ERIC

SERIES IN LEARNING AND COGNITION

Different Solution Procedures for Algebra Word and Equation Problems

Richard E. Mayer

Report No. 80-2

This research was supported by Grant NIE-G-78-0162 from the National Institute of Education.

ERIC

Abstract

In Experiment 1, 42 subjects solved algebra word problems in either work or equation format. In Experiment 2, 42 subjects solved or simply translated word problems into equations. The pattern of response latencies, by problem length, was different for the treatment groups. Results conflicted with a two-stage model of problem solving consisting of separate translation and solution processes.

ERIC

This paper is concerned with the development of a theory of algebraic problem solving. The rationale for studying algebraic problem solving includes the fact that work in this area would contribute to existing work on prose comprehension (Hinsley, Hayes & Simon, 1977), that a theory of algebraic problem solving would become a major component of a larger theory of deductive reasoning (Mayer & Revlin, 1978), and that the resulting theory could be tested in a wide variety of real world situations involving mathematics instruction (Kilpatrick, 1970).

Traditionally, studies of algebraic problem solving have dealt with one of three tasks:

(1) Equation Problems -- consisting of single or multiple equations and a goal of solving for a certain value. For example, 3X = 9.

(2) Word Problems -- consisting of a single or multiple sentences that can be translated into equations and solved without any additional semantic knowledge. For example, find a number such that 3 times the number is the same as 9.

(3) Story Problems -- consisting of single or multiple sentences that require additional semantic knowledge to be translated into equations and solved. For example, John drove his jet 9 miles on 3 gallons of gas. What was his mileage? (You must know that mileage = distance ÷ gas used.)

The domain of the present study will be restricted to single sentence word problems, although these tasks are closely related to the others.

Translation and Solution

An extensive literature review is beyond the scope of this paper. However, much prior work falls into two categories: computer simulations of algebraic problem solving processes (e.g., Bobrow, 1968; Hayes & Simon, 1974;

Bundy, Note 1; Bregar & Farley, Note 2; Novak, Note 3); and instructional studies aimed at improving mathematics education, such as published in Journal for Research in Mathematics Education or NCTM meeting proceedings (Barnett, Vos & Sowdder, Note 4; Clement, Lochhead & Soloman, Note 5). Unfortunately, there has not been much basic experimental research to fill the gap between high level theory (i.e. computer models) and educational application.

Perhaps the strongest single theme of the existing experimental research on algebraic problem solving is that the process of translating the problem into a representation is a crucial one. For example, Paige & Simon (1966) found that subjects who drew integrated pictures to represent word problems were more likely to perform correctly on the problems than subjects who drew fragmented pictures. Hayes and his collegues (Hinsley, Hayes & Simon, 1977 Robinson & Hayes, 1978) found that students have "schemas" for various types of story problems and that they use these schemas in selection of information they will pay attention to. Work on arithmetic word problems by Greeno and his associates (Riley & Greeno, Note 6; Heller & Greeno, Note 7) demonstrates that children can interpret an add/subtract operation in a word problem in different ways such as "cause/change" or "combine" or "compare," and that the difficulty of the problem depends on which interpretation is involved. In similar studies, Loftus & Suppes (1972) and Rosenthal & Resnick (1974) found that difficulty of word problems was affected by the linguistic structure of the problem. Schwartz (1971) found that subjects who used a matrix representation for information in a deduction problem were more successful. Also, Mayer & Greeno (1975) and Mayer (1978a; 1978b) reported that the same algebraic information was stored in memory in qualitatively different ways by different subjects. In review of representative studies in human deductive reasoning, Mayer & Revlin (1978,

p. 21) concluded:

> The most striking commonality among the studies is that emphasis on how reasoners encode the presented information... each researcher seems to focus on a slightly different aspect of the encoding process.

This conclusion seems particulary relevant to work on algebraic problem solving.

A second strong theme of experimental work involves mainly the process of solution (rather than translation). A detailed analysis of the solution protocols of individual problem solvers involved in solving physics problems suggests that students acquire a set of heuristic strategies that can be represented formally (Simon & Simon, 1978; Larkin, Note 8). Similarly, Greeno (1976, 1978) has provided a detailed analysis of the problem solving processes used by high school students solving geometry problems. These studies may be summarized by stating that detailed models of the problem solving process for story problems can be fit to the data of real students solving real science problems; furthermore, a major feature of the solution process for skilled performers is a reliance on various forms of heuristic planning.

The research on encoding and solution processes, as well as one's own introspections, might lead one to assume that solving word problems consists of two successive stages: translation from words to equation, and solution of the equation using the rules of algebra and the strategic planning of a good problem solver. Indeed, many algebra textbooks explicitly teach this procedure of translation followed by solution. Further, computer models of algebraic problem solving generally make a distinction between a program to translate a problem from English into an internal representation, and a program to solve the problem by applying operations to the internal representation. For example, Bobrow's (1968) STUDENT and Hayes & Simon's (1977) UNDERSTAND program make a clear distinction between translation and solution and tend to emphasize the first process; in addition, Bundy (Note 1) and Berger & Farley (Note 2) rely on the translation/solution distinction but tend to emphasize the solution process in their programs.

For the purposes of the present paper, the term "two-stage model" will refer to the idea that solving word problems involves:

(1) Translation Stage -- which takes words as its input and gives a formal representation (such as an equation) as its output, using encoding rules.

(2) Solution Stage -- which takes the output of the translation stage as its input and gives a numerical value as its output, using algebraic rules and strategic rules.

Although the two-stage model seems to be a well-accepted fact of life -- forming the basis for computer models as well as textbook lessons -- there has not been sufficient research evidence to establish its validity.

Problem Representation

This paper provides an experimental investigation of the two-stage model in the context of one type of word problem. In particular, this study investigates whether subjects use different solution procedures when an algebra problem is presented in equation format versus when it is presented as a word problem. The effects of problem representation on algebraic problem solving have been investigated in several previous studies.

In a previous set of experiments on algebraic substitution, subjects learned a set of four interlocking equations (Mayer, & Greeno, 1975, Experiment 1). For some subjects the information was presented as a story such as, Mileage = Distance/Gas Used, Total Time = Preparation Time + Driving Time, Driving Time = Arrival Time - Leaving Time, Speed = Distance/Driving Time.

For other subjects, the information was presented as an isomorphic set of nonsense equations, such as M = D/G, T = P + V, V = A - L, S = D/V. The results indicated that there were no major differences between groups in terms of response latencies or reported solution strategy for simple problems such as, "Preparation Time = 1/2 hour, Total Time = 2 hours, Find Driving Time". Thus if we had stopped at that point we would have concluded that the same problem solving model could account for the performance of both groups. However, two additional types of problems were also used: "unanswerable problems" such as "Speed = 50 mph, Driving Time = 1 hour, Find Mileage", and "questions" such as "Given Arrival Time and Leaving Time what else is needed to find Total Time?" For these types of problems, which require checking many relations, the story subjects were much faster (e.g., twice as fast for the unanswerable problems) as compared to the letter group. Thus, this study provided our first hint that the story format provides for a more integrated memory representation of the four equations and/or faster search speed when many relations must be tested.

If the story subjects tend to integrate the information with their existing knowledge, as is suggested by the previous result, one prediction is that they should be less influenced by the presentation order of the equations. In order to test this idea, another study (Mayer & Greeno, 1975, Experiments 3 and 4) was conducted in which subjects learned 9 interlocking equations presented either as meaningful stories or as nonsense equations. Subjects learned 3 equations to criterion, then learned a second set of three, and then learned a final set of three. On a subsequent test subjects were asked about only three of the equations such as, W = O * D, M = O * H, R = W/T. Problems required either no substitutions, such as "Given M and H could you find O?", one substitution such as, "Given M, H and D could you find W?", or two substitutions such as, "Given M, H, D and R could you find T?". For some subjects the 3 equations had all been learned in the same set (same set) and for other subjects the 3 equations came from 3 different sets during learning (separate sets). For the story subjects, performance was the same on problems requiring substitutions regardless of whether one or separate sets were involved; however, for the letter group response time on substitution problems was much longer (e.g. twice as long on two-substitution problems) for the separate set presentation as compared to the one set presentation. Thus the story subjects appeared not to be influenced by the organization of presentation while the letter subjects were. This result provides replicatory support for the idea that the story format allows for a more integrated memory representation, and thus faster performance on problems requiring search across several equation.

Since the above results encouraged the idea that presentation format influenced problem solving performance, a subsequent series of experiments was conducted which investigated the same issue using a simplier and better analyzed task -- linear reasoning. For example, in one study (Mayer, 1978, Experiment 2) subjects learned the remote pairs of linear ordering such as: F>H>R>B. For some subjects, premises were presented as a story such as: "The frog gets 10 times as many votes as the rabbit. The hawk gets 20 times as many votes as the bear. The frog gets 40 times as many votes as the bear." Other subject received isomorphic equations in nonsense form, such as, "F = 10 * R, H = 20 * B, F = 40 * B". Subjects were asked to answer questions involving no substitutions such as "Is F>R?", involving one substitutions such as "Is F>H?", and involving two substitutions such as "Is H>R?". If subjects remember the three premises and then use them to make inferences, the response latencies or error rates should increase with the number of required substitutions. This is the pattern obtained and best fit by the story subjects, but it was not obtained for the latter subjects. The latter group's data was best fit by a model in which each term was tagged, and question answering involved tag matching rather than making inferences. Thus, in this case problem format influenced both the mode of representation and the problem solving procedure.

The previous two sets of studies highlighted the negative aspects of equations -- i.e., requiring more time in complex situations, and encouraging a superficial solution strategy. However, the present experiment uses a task in which equations may hold an advantage; first, a single long equation is used rather than a list of substitutable equations, and second, the necessity of real-world knowledge is minimized. In this case, equations may hold an advantage over word format because they require less memory load (Hayes, 1978).

Problem Space

For example, consider the operators that a person uses to solve a word problem such as the following:

> Find a number such that if 8 more than 3 times the number is divided by 2, the result is the same as 11 less than 3 times the number.

For purposes of this paper, we will call this an algebra word problem, and subjects who are asked to solve these types of problems will be called the word group.

An isomorphic way of stating the problem in equation form is:

$$(8 + 3X)/2 = 3X - 11$$

We will call this format an algebra equation problem, and subjects who are asked to solve these types of problems will be members of the equation group.

There are two basic classes of actions that can be used to generate problem states in this problem:

MOVEs -- a variable or number is moved from one side of the equation to the other by addition, substraction, multiplication, or division of both sides by the same value, and

COMPUTEs -- combining two consecutive numbers or two instances of the same variable on one side of the equation by carrying out a computation such as addition, subtraction, multiplication or division.

For example, given the problem state,

$8 = 6X - 22 - 3X$

a MOVE operator would be to add 22 to both sides, yielding the new problem state,

$8 + 22 = 6X - 3X$

An alternative operator that could be applied to the problem state,

$8 = 6X - 22 - 3X$

is the COMPUTE operator; since there are two variables on one side of the equation they can be combined, yielding the new problem state,

$8 = 3X - 22$

Each problem state may be characterized by indicating the minimum number of MOVEs and of COMPUTEs that would have to be applied in order to move from the given state to the end state (i.e., X = a number). For example, the initial state,

$(8 + 3X)/2 = 3X - 11$

requires 5 computations and 4 moves to generate a value for X. Thus it is labeled as problem state 54; the first number indicates the number of required computations and the second number gives the number of required moves. If one MOVE is made, such as to multiply both sides by 2, the resulting state,

$$8 + 3X = 2(3X - 11)$$

requires 5 computations and only 3 moves; thus, it is labeled as problem state 53. Further, if the COMPUTE operator is applied twice on the right side of the equation, the resulting state,

$$8 + 3X = 6X - 22$$

requires only 3 computations and 3 moves; thus it is labeled as problem state 33. The remaining 3 moves are to get all the X's on one side, get all the numbers of the other side, and then to divide both sides by the number of Xs; the remaining 3 computations are to add 8 and 22, to subtract 3X from 6X, and to carry out the final division 30/3 to get a value for X.

Table 1 gives a list of some of the possible problem states, and labels each according to the minimum number of MOVEs and COMPUTEs that are required.[1] Labels with prime marks (') indicate that they have the same number of MOVEs and COMPUTEs as some other non-identical state. For example,

$$8 + 22 = 3X$$

requires a COMPUTE (add 8 and 22), a MOVE (divide both sides by 3) and a COMPUTE (divide 30 by 3). Similarly,

$$30 = 6X - 3X$$

requires a COMPUTE (subtract 3X from 6X), a MOVE (divide both sides by 3) and a COMPUTE (divide 30 by 3). Thus both problem states require 2 COMPUTEs and 1 MOVE, and are labeled as problem 21 and 21' respectively.

Insert Table 1 About Here

ERIC

There is more than one single path to the solution state. For example, one could move from state 53 to state 52 by applying a MOVE operator, or one could move from state 53 to state 33 by applying COMPUTE operators. A partial problem space is given in Table 2. The given state (54) is on the left and the goal state (00) is on the right; thus, any change in state to the right represents progress towards the goal. The problem space does not include any state with a fraction and does not include backwards moves from state 54. Table 2 consists of number pairs--representing problem states--connected by labeled lines--representing MOVEs or COMPUTEs. The first number indicates the number of required COMPUTEs and the second number indicates the number of required MOVEs, as is shown in Table 1. The M above a line refers to the application of a MOVE operator and the C refers to the application of a COMPUTE operator. For example, one sequence of problem states is 54,53,52,51,31,21,11,10,00; another possible sequence is 54,53,33,32,22,21,11,10,00. As can be seen, not all of the 14 states shown in Table 2 are needed for any given solution path.

Insert Table 2 about here

In the present experiments, however, subjects were asked to solve problems which began at each of the 14 states listed in Tables 1 or 2. Some subjects were given problems in word form (word group) and others were given problems in equation form (equation group). Thus the resulting data is a 2 x 14 array showing the response latencies for each group by problem state.

<u>Contrasting Versions of a Two-State Model</u>

<u>Simple Two-State Model</u>. The simple two-stage model states that solving an algebra word problem involves two states--translation plus solution--while

solving an isomorphic algebra equation involves only one state--solution. In addition, it is assumed that the solution processes are identical for corresponding word and equation problems; the only difference between subjects solving word and equation problems is that word problems require a translation and equation problems do not.

Finally, it is assumed that translation requires a constant amount of time regardless of problem state.[2]

According to this version of the two-stage model, the response time to answer a word problem is represented as,

$$RT_w(p) = RT_t + RT_s(p)$$

while the response time to answer a corresponding equation problem is,

$$RT_e(p) = RT_s(p)$$

where RT_t is the time to go from words to an equation for all problems, $RT_s(p)$ is the time to go from an equation to responding with the correct answer for problem p, RT_w is the total time to solve a word problem p and RT_e is the total time to solve a corresponding equation problem.

This model allows for a straightforward prediction concerning differences in performance of the two treatment groups on each of the 14 problem states: There should be an overall main effect in which the equation group is faster than the word group, and since this difference is constant over all problem types there should be no interaction between treatment and problem state.

Modified Two-Stage Model. The simple two-stage model makes an assumption that translation time is a constant regardless of problem state. In order to eliminate this dubious assumption, a modified two-stage model is introduced to this section. The modified two-state model is identical to the earlier model except that translation time varies as a function of problem state. Like the simple two-state model, it is still assumed that the solution process is the same for both types of problems.

The solution time to solve word problem p may be represented as follows,

$RT_w(p) = RT_t(p) + RT_s(p)$

and the time to solve a corresponding equation problem is,

$RT_e(p) = RT_s(p)$

where $RT_e(p)$ is the time to translate a given problem into an equation, and $RT_s(p)$ is the time to produce a correct response for an isomorphic equation.

This model also predicts an overall main effect in which the equation group performs faster than the word group. In addition, this model also predicts an interaction between treatment and problem state. The form of the interaction, however, is that the difference in latencies should increase as a function of problem length (i.e., as the complexity of the translation process). Thus, word problems requiring much translation should generate higher differences from corresponding equation problems than word problems that are short. Since there are several ways of defining problem length, these will be discussed in the results section.[3]

Different Solution Procedures. The previous two models have been based on the idea that responding to algebra word problems involves translation plus solution, and that the solution process for word problems is the same as the solution process for equation problems. An alternative presented in this section is that the solution process for a word problem is not the same for its corresponding equation problem.

According to this model the latency for responding to a word problem may be represented as follows.

$RT_w(p) = RT_t(p) + RT_{sw}(p)$

and the time to solve a corresponding equation problem is,

$RT_e(p) = RT_{se}(p)$

where $RT_t(p)$ is the time to translate a problem p, $RT_{sw}(p)$ is the time to solve the problem using the word-solution procedure, and $RT_{se}(p)$ is the time to solve the problem using the equation-solution procedure.

Since this model assumes qualitatively different solution procedure for the two treatment groups, it is important to specify the nature of the differences. The nature of the differences can be conceived in terms of differences in the conditions that are attached to the possible actions (i.e., MOVE and COMPUTE). In general, problem solvers in this task report two general types of goals: (1) Reduce expression--trying to carry out any indicated operations or clearing any parantheses as soon as possible, and (2) Isolate variables--trying to move all the Xs on one side and move all the numbers onto the other side. It may be noted that the conditions for reduce expression involve looking at only one side of the equality such as noting whether there is a parenthesis or whether there are two numbers on one side (i.e., 8 + 22 is on one side in problem state 21); in contrast, the conditions for applying operators with respect to the isolate goal always involve looking at both sides of the equation such as noting whether there is an X on both sides.

For purposes of this paper, it is reasonable to suppose that equation format requires less memory load than word presentation format. Since word format is cumbersome, it is more likely that the controlling strategy of subjects in this group is to reduce the expression. Since the equation allows for a unified visual representation of the entire problem, it is more likely that the controlling strategy of subjects in this group is to isolate the variable. Specific production system models which are based on these difference in the conditions will be discussed in the next section of this paper, and are presented in detail in a companion paper (Larkin & Mayer, 1980).

Based on these differences it is possible to predict differences in the pattern of solution performance that must result in a treatment x problem state interaction. In particular, the equation group should show evidence of planning and of setting "isolate variables" as its top goal while the word group should show a pattern of performance that does not involve planning but rather sets as its goal to "reduce expression".

Reduce Versus Isolate Strategies

In this section we present a more formal model which outlines the differences in performance between the equation subjects--who presumably use an isolate strategy--and the word subjects--who presumably use a reduce strategy. As a first step, Table 3 presents a general condition-action list (or means-ends table). The left side of each statement gives a general description of a situation (or difficulty) that may be encountered in solving the algebra equation, and the right side gives a general description of the action to be taken. Thus, the condition-action pairs listed in Table 3 are much more general than those needed for a running program but they are consistent with Newell & Simon (1972) system and will allow us to derive some predictions in this experiment.

We assume that a subject represents the information in a form that will allow for testing of the six conditions listed in Table 3. When a subject is given a problem in some state p, the subject searches for the conditions that are met. If more than one condition is met, the equation subjects will choose one to work based on the priority ordering, I-1, I-2, R-1, R-3, R-4; similarly, the priority ordering for word subjects is R-1, R-2, R-3, R-4, I-1, I-2.

Another way to state this difference is to say that equation subjects look first for isolate conditions and word subjects look first for reduce conditions. In the present experiment, there are several cases in which actions associated with isolate conditions (I-1 and I-2) cannot be carried out without further setting of subgoals. However, if subjects use the reduce strategy it is never necessary to stack subgoals.

An example of these differences is given in Tables 4 and 5. For the problem in state 54, three conditions are met: there are Xs on both sides (I-1), there are numbers on both sides (I-2), and there is a parenthesis (R-3). Equation subjects select I-1 as the condition-pair to be executed, but when trying to apply the action operators find that there is a constraint--i.e., there is a parenthesis. In order to remove this constraint, a new subgoal must be stacked on top of the I-1 subgoal, and so on. As can be seen, problem 54 involves 2 failures to carry out a subgoal (i.e., 2 instances of subgoal stacking), problem 53 requires 1 instance of subgoal stacking, and 33 requires none. For word subjects, there is never a failure to carry out a subgoal and thus no need for subgoal stacking. The results of applying the reduce and isolate strategies to all 14 problem states is summarized in Table 6. In order to predict response latencies it is necessary to predice that each move requires some constant amount of time, each computation requires some constant amount of time, and each instance of goal stacking (planning) requires an additional amount of time.[4]

Like the modified two-stage model, this formulation predicts an interaction between treatment and problem length. However, unlike the previous model, the pattern of interaction should reflect different solution strategies. In particular, the interaction should reflect the fact that there is goal stacking required for problem states 54 (2 time), 53 (1 time), and 51 (1 time) for the equation group but not the word group.

Tables 3, 4, 5, 6 about here

Experiment 1

Subjects solved problems in each of the 14 problem states with some subjects receiving all problems in word format and some subjects receiving them in equation format. Experiment 1 was conducted in order to test the predictions of each of the models discussed above. The simple two-stage model predicts an overall main effect but no interaction. The modified two-stage model predicts an overall main effect as well as a monotonically increasing difference as a function of problem length. Finally, if the subjects in the groups use qualitatively different solution strategies such as isolate versus reduce, more specific differences in the goal stacking procedures can be described.

Method

Subjects and design. The subjects were 42 college students recruited from the Psychology Subject Pool at the University of California, Santa Barbara. Twenty-one subjects served in the equation group and 21 subjects served in the word group. All subjects solved the same 14 states of problems so state of problem is a within-subject factor.

Materials and apparatus. Two sets of 98 problems were constructed. The equation set presented the problems in equation form such as:

$(8 + 3X)/2 = 3X - 11$

The word set presented problems identical to those in the equation set but used words such as:

Find a number that if 8 more than 3 times the number is divided by 2 the results is the same as 11 less than 3 times the number.

The 98 problems were generated using a 7 by 14 design. The first factor indicates that seven different problems were used such as the one indicated above. All problems were of the same form but specific values were different. The second factor indicates that each problem was broken down into 14 different problem states by manipulating how many computations and moves had to be made. Table 1 shows the 14 states of each problem, and Table 2 shows a partial problem space.

The apparatus consisted of an ADM-3 terminal screen and keyboard for presentation of the stimuli and entering of responses; an ADM-3 terminal that was used as an experimenter monitor; a Micropolis dual floppy disk drive used for storing the stimuli, the experiment program, and the data; a Teletype printer used for printing out the data; and an Altair 8080 computer used to control these devices.

Procedure. Each subject was run individually in a session that lasted approximately 30 to 60 minutes. First, subjects were given verbal instructions to make sure that they could solve algebraic equations similar to those in the experiment. Three subjects were unable or unwilling to solve the trial equation, so new subjects were run in their places. Instructions stated that an equation (or a word problem) would be presented on the screen. The subject was to logically solve the problem for the unknown, and then press the corresponding button on the keyboard. The buttons were labeled from 1 to 30 by integers and included one button labeled "none." Subjects were explicitly told not to guess or to use any shortcut method; rather they were told to carefully work out the solution step by step and to press the answer button only when they were sure of their answer.[5] Subjects were allowed 60 seconds to answer; as soon as the subject answered, or as soon as 60 seconds passed, the question was removed from the screen and a new question appeared in 2 seconds. Each new question was preceeded by a beep.

The questions were presented in random order except for the constraint that each set of 14 items consisted of all 14 problem states and that each set of 14 items contained 2 different instances of each problem. After each set of 14 problems, subjects were given a brief rest period.

First a set of 14 practice items were given, one at a time. Then seven sets of 14 experimental items were given. At the end of the experimental session subjects were asked about their solution strategies to make sure that subjects had followed directions.

Results and Discussion

The average response time for each of the 14 problem states was computed for each subject in each group.[6] Table 7 summarizes the mean response time for each of the two groups by problem state. An analysis of variance was performed on the data with treatment as a between-subject factor and problem state as a with-subject factor. As expected, the ANOVA revealed that the equation group was significantly faster than the word group in overall performance, $F(1,40) = 71,244$, $p < .001$, and there were overall differences in response time for different problem states, $F(13,520) = 197.11$, $p < .001$.

The main focus of this experiment, however, is on the question of whether subjects in the two treatment groups followed different solution procedures. The "translation plus solution" theory states that subjects in the word group simply translate the problem into an equation and then solve the equation in the same was as the equation group. A simple version of this theory assumes that translation requires a constant amount of time for each problem, and thus predicts that the word group should take longer than the equation group overall but there should be no interaction with problem state. The ANOVA revealed a significant pattern of interaction between treatment and problem state, $F(13, 520) = 17.78$, $p < .001$, and thus provides evidence against a simple version of the "translation plus solution theory."

Insert Table 7 about here

These results allow one to reject a simple version of the "translation plus solution" theory. However, a modified version of the theory predicts an interaction if its assumed that the translation stage takes longer for longer problems. This modified version of the two-stage theory predicts an interaction between treatment and problem state in which there is a monotonically increasing difference as problem length (indicated by number of computations and moves) increases. The difference in response time (in seconds) between the two treatment groups for each of the 14 problem states are listed in the last row of Table 7. A linear regression[3] using number of steps (i.e., computations or moves) as the independent variable and the difference in RT as the dependent variable produced an R^2 of only .65, and a multiple regression using number of computations and number of moves as the independent variables and the difference in RT as the dependent variable produced an R^2 of only .70. These results indicate that the treatment by problem state interaction cannot be neatly described as consisting of monotonically increasing differences, if problem difficulty is defined as a function of the number of steps.

An alternative theory that predicts an interaction between treatment and problem state is that the two groups engage in qualitatively different solution procedures. For example, the data presented in Table 7 suggests that there are three general stages for the equation group indicated by a sharp drop from state 54 to 53, and from each of the 5-computation states (53, 52 or 51) to its respective 3-computation state (33, 32 or 31). One hypothesis developed earlier is that the equation subjects may plan ahead and stack goals; for example, they can stop at state 54 and see several moves ahead, at the other 5-computation states they can see several computations ahead, and at the 3-computation states they can see to the end of the problem. The three major jumps in response time suggest that subjects are able to form subgoals (Thomas, 1974). However, the word subjects do not show the same subgoal pattern; for them each additional computation or move tends to add a constant amount to solution time.

In order to provide more information on this observation, several multiple regressions were fit to the means for the 14 problem states for each group. These are summarized in Table 8. First, a simple linear regression was used with the independent variable being number of steps. (A step was defined as either a move or a computation such that 54 state required 9 moves, and state 10 required 1). As shown in Table 4, the model fit the word group reasonably well ($R^2 = .95$) but did not fit the equation group well ($R^2 = .83$). Second, a multiple regression was used with the independent variables being number of computations and number of moves. As shown in Table 8, this model fit the word group slightly better than the one-variable model, but did not fit the equation group well ($R^2 = .84$). Finally, a multiple regression was used that included three independent variables: number of computations, number of moves, and stage level. Problem state 54 was defined as stage level 2, problem states 53, 52 and 51 were designated as stage level 1, and all the lower states were designated as stage level 0. This model resulted in no improvement of fit for the word group, but did result in an enormous improvement of fit for the equation group (R^2 = .99). Thus, the equation group was best fit by the three variable (or stage model) while the word group was best fit by a simple step model or two variable step model. These results suggest that the pattern of interaction between the two groups cannot be adequately described by a simple or modified "translation plus solution" theory. In contrast, there is evidence of different processing strategies as indicated by the fact that the group response patterns are best fit by different models -- a planning or stage model for the equation group and a step model for the word group.

Insert Table 8 about here

The previous analyses explored trends in the group means. An additional analysis was carried out to determine whether the individual data encouraged the same conclusion as the group data. Each of the three regression models--one variable, two-variable and three-variable -- was fit to the pattern of mean response latencies for 14 problem state for each individual subject. Table 9 shows the value of R^2, indicating the goodness of fit, for each model applied to each of the 42 subjects. Table 10 shows the number of subjects in each group (n = 21) who were best fit by each regression model; the top portion (strict criterion) of Table 10 defines "best fit" as yielding a R^2 value that is more than .01 higher than the next simpliest model, and the bottom portion (linient criterion) defines "best fit" as yielding an R^2 value that is more than .03 higher than the next simpliest model. As can be seen, 20 of the 21 subjects in the equation group are best fit by the three-variable (or planning) model, while most the the word group subjects are best fit by the one or two-variable (or step) models. Chi-square tests were conducted for the data in the top and bottom portions of Table 10; there was a significantly different classification pattern for the two treatment groups using the strict criterion, $\chi^2 = 20.88$, df = 2, $p < .001$, and using the linient criterion, $\chi^2 = 25.04$, df = 2, $p < .001$. Thus these results are consistent with the conclusions presented on the basis of group data.

Insert Tables 9 and 10 about here

A final analysis was conducted in order to determine whether the differences between the treatment groups was an artifact of the many repeated trials in the experiment. For example, since each of seven problems was repeated in 14 different problem states, it is possible that subjects by the end of the session learned to respond on the basis of distinctive cues in the problem rather than actually computing an answer. It should be pointed out that the instructions clearly stated that the subject should compute the answer rather than try to find ways of guessing, and further, that subjects indicated that they had followed directions when questioned after the experiment. However, in order to provide more data on this question, an analysis was conducted using data from the first experimental trial (there was a practice trial before this trial) and the last experimental trial; each of these consisted of one of each of the 14 problem states with all problems counterbalanced. The mean response latencies for each group on each of the problem states occurring in the first and last experimental trials is shown in Table 11.

Insert Table 11 about here

As can be seen in comparing the pattern of the equation and word groups, the same general pattern is obtained for both the first and last trial. In both cases, the word group shows a pattern of consistently more time required for each new step while the equation group shows a jump from S4 to S3 and from the 5-computation state to its corresponding 3-computation state. These patterns are similar to those presented and discussed in conjunction with Table 3. An analysis of variance was conducted using treatment as a between subjects factor and trial and problem state as within subject factors. As expected, the equation group was faster overall, $F(1,13) = 75.23$, $p < .001$., and performance on the last trial was faster than on the first, $F(1,40) = 104.95$, $p < .001$. In addition there was a pattern of interaction between treatment and problem state that was similar to that shown in Table 7, $F(13,520) = 9.42$, $p < .001$. There was, however, no evidence that the treatment x problem state interaction was different for the first versus the last trial; this observation is consistent with the failure to obtain a three-way interaction involving treatment, problem state and trial, $F(13,520) = 1.05$, n.s. In addition, a separate ANOVA was performed on the data for the first trial and a separate ANOVA was performed on the data for the last trial with treatment and problem state as the factors. In both cases there was the same significant treatment x problem state interaction; for the first trial data, $F(13,520) = 7.03$, $p < .001$; for the last trial data, $F(13, 520) = 4.81$, $p < .001$. These results provide evidence that the differences in performance between the groups was not based on an artifact of the design that allowed for learning of a short cut as the session progressed.

Experiment 2

The results of Experiment 1 suggest that the solution process used by subjects in the equation group involves planning. However, the translation and solution process of the word group does not show the same sort of planning. One reasonable conclusion is that the two groups use qualitatively different solution processes, as represented by the isolate and reduce strategies

In order to provide another test the implications of the models discussed above, a second experiment was conducted in which one group wrote a numerical answer for each word problem (word-to-solution group), and another group wrote an equation for each word problem (word-to-equation group). According to the two-stage model performance of the word-to-solution task involves the following:

$$RT_{w\text{-}to\text{-}s} = RT_t(p) + RT_s(p) + RT_w(p)$$

and the performance of the word-to-equation group involves:

$$RT_{w\text{-}to\text{-}e} = RT_t(p) + RT_w(p)$$

where $RT_t(p)$ involves the time to go from the word problem to a statement of the underlying equation, RT_s involves moving from the equation to a numerical answer (a process that requires planning when performed separately) and RT_w involves the time to write an equation or a number on a sheet of paper. The previous experiment suggests that the RT_s component involves planning.

The previous experiment demonstrates that when the solution stage is performed alone (i.e., for equation problems) there is evidence of planning. If there are two independent stages in solving algebra word problems, and if the solution stage requires planning, then the performance of the two groups in Experiment 2 should differ. The two-stage model predicts that the word-to-solution group should show a pattern involving planning--since a solution process is required--while the word-to-equation group should show a pattern in which performance depends only on the length of the problem.

As an alternative it may not be possible to cleanly separate the translation and solution processes for word problems. In this case, both groups will show the same general pattern of performance since both groups engage in the same step-by-step translation process; however, since each step requires more computation for the word-to-solution group then each step should take more time.

Method

Subjects and design. The subjects were 42 college students from the University of California, Santa Barbara. Twenty-one subjects served in the word-to-solution group and 21 served in the word-to-equation group. All subjects solved the same 14 states of problems so state of problem is a within subject factor.

Materials and apparatus. The same 98 word problems and the same apparatus were used in Experiment 1.

Procedure. The procedure was similar to that used in Experiment 1, except that subjects were asked to write their answers on a sheet of paper and then press any button on the terminal. For the word-to-equation subjects, the task was to write down an equation--using X as the unknown--to represent the word problem. For the word-to-solution subjects, the task was to write down a number for the answer. All subjects were allowed to circle their final answer. In addition, subjects were told to press any button on the keyboard immediately after circling the answer they had written. Each subject solved 14 practice and 42 target problems; these were presented in sets of 14 each and were counterbalanced as in Experiment 1.

Results and Discussion

The average response time[7] for each of the 14 problem states was computed for each subject in each group, as in Experiment 1. Table 12 summarizes the mean response time for each of the two treatment groups by problem state, and is comparable to Table 7 for Experiment 1. As in Experiment 1, an analysis of variance was performed on the data with treatment as a between-subjects factor and problem state as a within-subject factor. The overall mean response time for the two treatment groups was identical (23.3 sec) so the overall main effect for treatment produced an F of zero. As expected there was an overall main effect due to problem state, reflecting the fact that longer problems required more time, $F(13,520) = 226.1$, $p < .001$.

The main focus of this experiment was on the pattern of performance of the two groups by problem state. Both groups appear to show a pattern that is similar to that displayed by the word group in Experiment 1 -- monotonically increasing response time as a function of problem length. However, the treatment by problem state interaction is statistically significant, $F(13,520) = 14.18$, $p < .001$. The interaction can be described by saying that the word-to-solution group is faster than the word-to-equation group on simple short problems but the word-to-equation group is faster than the word-to-solution group on long problems. This interaction is consistent with the idea that translation to equations is not needed and may interfer with problem solving, especially for problems that do not require much planning.

Insert Table 12 about here

In order to better understand the nature of the treatment by problem state interaction, several muliple regressions were performed. The goal of these analyses was to determine which models best fit the performances of the two groups. In Experiment 1, the performance of the word group was best described by a simple one-variable or two-variable step model while the equation group was best fit by a three-variable planning model. The group means by problem state were fit to each of these models, as in Experiment 1; in addition, another one-variable model was tested in which the number of calculations served as the independent variable.

The two-stage model (or translation plus solution theory) predicts that the word-to-solution group will demonstrate a solution process that uses planning heuristics -- as indicated by a fit to the three-variable planning regression model -- while the word-to-equation group will demonstrate a linear step-by-step process -- indicated by a fit to a one or two-variable step model. However, the alternative model predicts that the two treatment groups will show similar patterns of performace -- both being fit by a simple step model.

The results of these analyses are summarized in Table 13. As can be seen, both groups were best fit by simple one-variable or two-variable step models, and neither group required the planning model that characterized the equation group in Experiment 1. Apparently, both groups performed more like the word group in Experiment 1 than like the equation group. For example, a comparison of Tables 8 and 13 shows that the word-to-solution group replicates the general form of performance of the word group in Experiment 1. The word-to-equation group shows a similar trend except that each additional step requires less time than for the word-to-solution group. One interpretation that is consistent with these findings is that both groups engage in segment-by-segment translation process but the word-to-solution group must also perform some operations on each newly translated segment.

Insert Table 13 about here

A more detailed inspection of the results shown in Table 13 shows the following: (1) The intercept for the word-to-solution group is higher than for the word-to-translation groups; this is consistent with the idea that the RT_{write} process takes longer for equations which consist of many symbols than for writing a single number. (2) The time required for additional calculation steps is greater for the word-to-solution group; this is consistent with the idea that this group must actually make a computation as part of the ongoing translation/compacting process while the other group does not. Further, the number of moves required increases the solution time for the word-to-solution group, presumably because each move must actually be executed as part of the translation/compacting process. However, the number of moves does not increase the required time for the word-to-equation group and in fact serves to decrease it slightly; this is consistent with the idea that the number of moves does not increase the number of segments that must be translated. For example, problem states 54, 53, 52 and 51 do not differ in terms of the number of variables and relations that must be translated.

Supplemental Study

The previous studies suggest that equation format allows goal stacking for long problems. However, this conclusion is based only on the pattern of response latencies. In order to provide additional data on the problem-solving process of subjects, an interview study was conducted.

The subjects were eight college students from the Introductory Psychology Subject Pool at the University of Pittsburgh, with four subjects in the equation group and four subjects in the word group.

Subjects were randomly assigned and run individually. Each subject was asked to solve two of the state 54 equations used in previous experiments. For the equation treatment, an equation was printed on a chalkboard by the experimenter. The subject was asked to "think aloud" and tell the experimenter what to do with the equation. Each action that the subject called for was carried out by the experimenter writing a new equation on the board under the previous one. When a subject finished a problem, he/she was asked to go back to the first step of the problem; in particular, the subject was asked whether he/she first thought about "getting rid of the parentheses" or about "getting the Xs on one side". A similar procedure was used for the word treatment, except that the experimenter wrote the problem on the board in word form. The experimenter transcribed all of the problem states generated on the chalkboard and all of the subjects' comments concerning search for conditions; the entire session was also tape recorded.

The main interest in this study is to determine whether there is any evidence that equation subjects engage in goal stacking. Two of the four subjects in the equation group gave clear evidence of goal stacking but none of the word subjects did. For example, the solution process and comments of one of the equation subjects is given in Table 14. As can be seen, this subject sets "isolate X" as a major goal but this strategy leads to several failures. Of the two subjects in the equation group who did not show signs of goal stacking, one used fractions and one gave a fast, textbook proper description. However, even the subject who used fractions gave a hint of the "isolate strategy"; "I wanted to get X on one side, but to do that I had to do something with that fraction. Separate 8 from 3X and remembering the basic rules of algebra you can't subtract them out when they are over a fraction. You have to separate them." Thus, this subject also expresses a goal stacking approach in which moving X cannot be accomplished until it is separated from the fraction. None of the word subjects gave any evidence of setting "isolate X" as a goal, or of goal stacking.

In addition, several equation subjects expressed difficulty in using the slash (/) as a division symbol. Hayes (1973) has found similar evidence for the role of spatial factors in solving equations. Many subjects expressed difficulty in verbalizing their thought process, and there is reason to believe that this task is not condictive to protocol analysis. However, the results provide support for the reality of goal stacking in equation subjects.

Insert Table 14 about here

General Conclusion

This study provides new information concerning the problem solving process for algebra word problems. First, Experiment 1 compared solving word problems (which presumably require separate translation and solution phases) with solving corresponding equation problems (which presumably require only the solution phase). Results indicated that is was not possible to characterize the behavior of the word problem solvers as consisting of a translation phase followed by a solution phase like that of the equation problem solvers. Rather there was evidence, consistent with earlier results by Mayer & Greeno, (1975) that the solution process was qualitatively different for the two treatment groups. The word problem solvers displayed a pattern of monotomically increasing response increasing response times as a function of steps, as previously noted by Loftus

& Suppes (1974); the equation problem solvers displayed heuristic planning similar in some respects to those suggested by Bundy (Note 1).

Second, Experiment 2 compared the processes involved in translation from words to equations and the processes involved in translation from word to answer. Solving word problems produced a pattern of behavior that indicated a modified version of tranlation alone. There was, again, no evidence that the solution phase involves the planning procedures as produced by the equation group in Experiment 1. Thus these two studies cast serious doubts on the applicability of the two-stage model of algebraic problem solving to the current task.

Finally, these results suggest directions for future research on how humans solve algebra word problems. The results of these studies, as well as other studies comparing word versus equation problem solving (Mayer & Greeno, 1975; Mayer, 1978a, 1978b), indicate that theories of problem solving based on solution of equations may be different than theories based on solution of corresponding word problems. Hueristic solution models must take the problem representation into account. The present results encourage the idea that work on comprehension of prose might be relevant to work on the solution of word problems. This is so because the present results show that the comprehension process and the solution process are far more intertwined than was previously assumed. It seems likely that when subjects are given complex word problems that overload their working memories, they do not rely on the straightforward translation-plus-solution strategy. Rather, they appear to rely on a successive chunking procedure in which segments are translated and operated upon in a piece-by-piece way. Several researchers have pointed to the important role of working memory in prose comprehension (Kintsch & van Dijk, 1978; Britton, Holdredge, Curry & Westbrook. 1979). Furthermore, individual differences in working memory (Hunt, Lunneborg & Lewis, 1975) might be important in encouraging the use of different problem solving strategies.

ERIC

Reference Notes

1. Bundy, A. Analysing mathematical proofs, or reading between the lines. (Research Report No. 2). Edinburgh: University of Edinburgh, Department of Artificial Intelligence, May, 1975.

2. Bergar, W. S. Interactive problem solving in elementary albegra. (Report 78-3-1). Corvallis, Oregon: Oregon State University, Computer Science Department, March, 1978.

3. Novak, G. S., Jr. Computer understanding of physics problems stated in natural language. (Technical Report NL-20). Austin: University of Texas, Department of Computer Science, 1976.

4. Barnett, J., Vos, K. and Sowder, L. A review of selected literature in applied problem solving research. Unpublished manuscript). Nothern Illinois University, Department of Matehmatics, May, 1978.

5. Clement, J., Lochhead, J. and Soloway, E. Translating between symbol systems: Isolating a common difficulty in solving algebra word problems. (Unpublished manuscript). Amherst, Massachusetts: University of Massachusetts, Department of Physics and Astronomy, March, 1979.

6. Riley, M. S., & Greeno, J. G. Importance of semantic structure in the difficulty of arithmetic word problems. Paper presented at Midwestern Psychological Association, Chicago, May, 1978.

7. Heller, J. I. & Greeno, J. G. Semantic processing in arithmetic word problem solving. Paper presented in Midwestern Psychological Association, Chicago, May, 1978.

References

Bobrow, D. G. Natural language input for a computer problem-solving system. In M. Minsky (Ed.), Semantic information processing. Cambridge, Mass.: MIT press, 1968.

Britton, B. K., Holdredge, T. S., Curry, C. & Westbrook, R. D. Use of cognitive capacity in reading identical texts with different amounts of discourse level meaning. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 262-270.

Greeno, J. G. Cognitive objectives of instruction: Theory of knowledge for solving problems and answering questions. In D. Klahr (Ed.), Cognition and instruction. Hillsdale, N. J.: Erlbaum Associates, 1976.

Greeno, J. G. A study of problem solving. In R. Glaser (Ed.), Advances in instructional psychology, Vol. 1. Hillsdale, N. J.: Erlbaum, 1978.

Hayes, J. R. On the function of visual imagery in elementary mathematics. In W. G. Chase (Ed.) Visual information processing. New York: Academic Press, 1973.

Hayes, J. R. & Simon, H. A. Understanding written problem instruction. In L. Gregg (Ed.), Knowledge and cognition. Hillsdale, N. J. Erlbaum Associates, 1974.

Hinsley, D., Hayes, J. R. & Simon, H. A. From words to equations. In P. Carpenter & M. Just (Eds.), Cognitive processes in comprehension. Hillsdale, N. J.: Erlbaum Associates, 1977.

Hunt, E., Lunneborg, C. & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Kintsch, W. & Van Dijk, T. A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-94.

Larkin, J. and Mayer, R. E. Models of performance on algebra word and equation problems with alternative solution strategies. In preparation.

Loftus, E. F. & Suppes, P. Structural variables that determine problem-solving difficulty in computer-assisted instruction. Journal of Educational Psychology, 1972, 63, 531-542.

Mayer, R. E. Effects of meaningfulness on the representation of knowledge and the process of inference for mathematical problem solving. In R. Revlin & R. E. Mayer (Eds.), Human reasoning. New York: Wiley/Winston, 1978.

Mayer, R. E. Qualitatively different storage and processing strategies used for linear reasoning tasks due to meaningfulness of premises. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 5-18.

Mayer, R. E. & Greeno, J. G. Effects of meaningfulness and organization on problem solving and computability judgments. Memory & Cognition, 1975, 3 356-362.

Mayer, R. E. & Revlin, R. An information processing framwork for research on human reasoning. In R. Revlin, & R. E. Mayer (Eds.), Human reasoning, New York: Wiley/Winston, 1978.

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

Paige, J. M. & Simon, H. A. Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Eds.), Problem Solving. New York: Wiley, 1966.

Robinson, C. S. & Hayes, J. R. Making inferences about relevance in understanding problems. In R. Revlin & R. E. Mayer (Eds.), Human reasoning. New York: Wiley/Winston, 1978.

Rosenthal, D. J. A. & Resnick, L. Children's solution processes in arithmetic word problems. Journal of Educational Psychology, 1974, 66, 817-825.

Schwartz, S. H. Modes of representation and problem solving: Well evolved is half solved. Journal of Experimental Psychology, 1971, 91, 347-350.

Simon, D. P. & Simon, H. A. Individual differences in solving physics problems. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, J. J.: Erlbaum Associates, 1978.

Thomas, J. C., Jr. An analysis of behavior in the hottits-orcs problem. Cognitive Psychology, 1974, 6, 257-69.

Footnotes

This research was supported by Grant NIE-G-78-0162 from the National Institute of Education. The author appreciates the helpful comments of James Greeno and Jill Larkin. This paper was written while the author was on sabbatical leave at the Learning Research and Development Center, University of Pittsburgh. Requests for reprints should be sent to: Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106.

1. Problem states involving fractions or involving backwards moves from the given state are not included.
2. The strategy used in this paper is to begin with as simple a model as possible, and to add more parameters only as needed. Thus, we begin with a simplistic assumption that translation time is a constant, but later modify that assumption below.
3. Based on the results of Experiment 2, the best indicators of translation difficulty are number of moves and computations required.
4. The isolate strategy and the reduce strategy are not the only possible models but they both are the simpliest and correspond to the reports of problem solvers. In addition, the assumption that all computations take an equal amount of time, all computes take an equal mount of time, and each instance of goal stacking takes an equal amount of time, are made in the interests of simplicity. If there are gross differences between individual MOVEs or COMPUTEs or STAGEs then the fit of our models should suffer and we would be encouraged to add even more parameters. Fortunately, this is not necessary in the present experiments.

5. One strategy that we did not want our subjects to engage is that of estimating values to plug into X. The instructions emphasized the fact that subjects should use logical deduction, and a post-experimental questionnaire confirmed that subjects followed instructions.

6. Error rates were low, and the distribution of errors was similar for the two groups. The equation group averaged 4% errors and the word group averaged 7% errors with errors defined as not giving a correct response within 60 seconds.

7. Error rates were low, and the distribution of errors was similar for the two groups. The word-to-solution group averaged 5% errors and the word-to-equation group averaged 1% errors.

Table 1

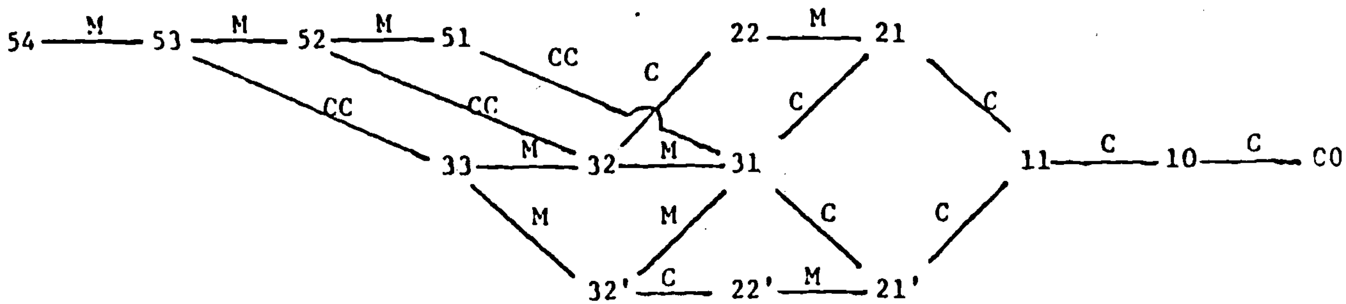
Fourteen Problem States for (8 + 3X)/2 = 3X - 11

Computations and Moves Required	Problem State
54	(8 + 3X)/2 = 3X - 11
53	8 + 3X = (3X - 11)·2
52	8 = 2 (3X - 11) -3X
51	8 + 2(11) = 2(3X) -3X
33	8 + 3X = 6X - 22
32	8 = 6X - 22 - 3X
32'	8 + 22 + 3X = 6X
31	8 + 22 = 6X - 3X
22	8 = 3X - 22
22'	30 + 3X = 6X
21	8 + 22 = 3X
21'	30 = 6X - 3X
11	30 = 3X
10	30/3 = X

ERIC

Table 2

Problem Space for (8 + 3X)/2 = 3X - 11



Note. - Problem states are represented by two digits in circles; the first digit indicates the number of required calcuations, the second digit indicates the number of required moves. Operations are represented by letters next to arrows; M indicates a move, C indicates a calculation, CC indicates two calcuations.

Table 3

Some Condition-Action Pairs for Solving Problem 54

Isolate Variable

(I-1) Xs are both sides of the equation -->

Move X to left side and combine with other X

(I-2) Ns are both sides of the equation -->

Move N to right side and combine with other N

Reduce Expression

(R-1) 2 Xs on one side of the equation -->

Combine them

(R-2) 2 Ns on one side of the equation -->

Combine them

(R-3) Parenthesis on one side of the equation attached to division -->

Move divided term to other side of equation

(R-4) Parenthesis on one side of equation attached to multiplication -->

Carry out the multiplications

Table 4

Solution of Problem 54 Using Isolate Strategy

Problem State	Events	Calculation	Move	State
(54)	3X - 11 (8 + 3X)/2			
	Conditions: I-1, I-2, R-3			
	Goal: I-1			
	Fail due to PARENS (R-3)			1
	Goal: R-3			
	Succeed		1	
(53)	2(3X - 11) = (8 + 3X)			
	Conditions: I-1, I-2, R-4			
	Goal: I-1			
	Fail due to PARENS (R-4)			1
	Goal: R-4			
	Succeed	2		
(33)	6X - 22 = 8 + 3X			
	Conditions: I-1, I-2			
	Goal: I-1			
	Succeed	1	1	
(22)	3X - 22 = 8			
	Conditions: I-2			
	Goal: I-2			
	Succeed	1	1	
(11)	3X = 30			
	Conditions: I-2			
	Goal: I-2			
	Succeed	1		
(00)	X = 10			
	From Problem State 54	5	4	2
	From Problem State 53	5	3	1
	From Problem State 33	3	3	0
	From Problem State 22	2	2	0
	From Problem State 10	1	1	0

Table 5

Solution of Problem 54 Using Reduce Strategy

Problem State	Event	Calculation	Move	Stage
(54)	3X - 11 = (8 + 3X)/2			
	Conditions: I-1, I-2, R-3			
	Goal: R-3			
	Succeed		1	
(53)	2(3X - 11) = 8 + 3X			
	Conditions: I-1, I-2, R-4			
	Goal: R-4			
	Succeed	2		
(33)	6X - 22 = 8 + 3X			
	Conditions: I-1, I-2			
	Goal: I-1			
	Succeed	1	1	
(22)	3X - 22 = 8			
	Conditions: I-2			
	Goal: I-2			
	Succeed	1	1	
(11)	3X = 30			
	Conditions: I-2			
	Goal: I-2			
	Succeed	1	1	
(00)	X = 10	1	1	
	From Problem State 54	5	4	0
	From Problem State 53	5	3	0
	From Problem State 33	3	3	0
	From Problem State 22	2	2	0
	From Problem State 11	1	1	0

Table 6

Number of Moves, Computations, and Subgoal Stackings for 14 Problem States by Two Solution Strategies

Probl me State	Isolate Strategy			Reduce Strategy		
	COMPUTEs	MOVEs	STATES	COMPUTEs	MOVEs	STATES
54	5	4	2	5	4	0
53	5	3	1	5	3	0
52	5	2	1	5	2	0
51	5	1	1	5	1	0
33	3	3	0	3	3	0
32	3	2	0	3	2	0
32'	3	2	0	3	2	0
31	3	1	0	3	1	0
22	2	2	0	2	2	0
22'	2	2	0	2	2	0
21	2	1	0	2	1	0
21'	2	1	0	2	1	0
11	1	1	0	1	1	0
10	1	0	0	1	0	0

Table 7

Mean Response Time by Problem State for Two Treatment Groups --Experiment 1

Treatment	Problem State													
	10	11	21	21'	22	22'	31	32	32'	33	51	52	53	54
Equation Group	2.7	2.1	3.1	3.7	4.6	5.0	5.0	5.8	6.1	7.9	12.6	14.7	15.2	25.0
Word Group	3.4	4.0	6.2	10.7	11.0	13.8	14.6	18.1	18.4	20.7	27.4	28.8	31.3	34.3
Difference	.7	1.9	3.1	7.0	6.4	8.8	9.6	12.3	12.3	12.8	14.8	14.1	16.1	9.3

Note. - Main effect of treatment, $p < .001$. Main effect of problem state, $p < .001$. Introduction of treatment and problem state, $p < .001$.

ERIC

Table 8

Values of R^2 and Variable Weighings for Three Multiple Regressions Fit to Two Treatment Groups--

Experiment 1

TREATMENT	ONE VARIABLE	TWO VARIABLES	THREE VARIABLES
EQUATION GROUP	R^2 = .83 STEP = 2.63 sec Intercept = - 4.40 sec	R^2 = .84 COMPUTATION = 3.02 sec MOVE = 2.04 sec Intercept = -4.59 sec	R^2 = .99 COMPUTATION = .69 sec MOVE = 1.42 sec STAGE = 7.25 sec Intercept = .90 sec
WORD GROUP	R^2 = .95 STEP = 4.37 sec Intercept = 3.58 sec	R^2 .98 COMPUTATION = 5.63 sec MOVE = 2.47 sec Intercept = 3.94 sec	R^2 = .98 COMPUTATION = 5.42 sec MOVE = 2.43 sec STAGE = .61 sec Intercept = -3.39 sec

ERIC

Table 9

Values of R^2 for 42 Subjects on Three Models -- Experiment 1

Individuals Ss	Word Groups			Equation Group		
	Regression 1 (One Variable)	Regression 2 (Two Variable)	Regression 3 (Three Variable)	Regression 1 (One Variable)	Regression 2 (Two Variable)	Regression 3 (Three Variable)
1	.881	.930*	.934	.843	.870	.913*
2	.926	.969*	.970	.635	.685	.825*
3	.882	.968*	.972	.707	.785	.869*
4	.934	.965*	.967	.753	.758	.888*
5	.902	.902	.982	.735	.747	.848*
6	.699	.905*	.911	.705	.927	.975*
7	.901	.939*	.978*	.763	.765	.878*
8	.933	.962*	.964	.798	.798	.884*
9	.648	.694	.790*	.863	.910	.951*
10	.935	.946*	.947	.710	.849	.908*
11	.895	.934*	.937	.715	.715	.850*
12	.922	.954*	.954	.797	.825	.917*
13	.890*	.892	.893	.779	.734	.869*
14	.852	.854	.963*	.839	.850	.910*
15	.862	.892	.921*	.818*	.821	.828*
16	.933	.945*	.955	.728	.736	.861*
17	.870	.952	.971*	.772	.798	.911*
18	.911	.968*	.966	.803	.832	.929*
19	.881	.966*	.957	.831	.870	.910*
20	.945*	.948	.949	.693	.698	.794*
21	.935	.948*	.945	.861	.863	.945*

Note. - Asterisk (*) indicates best fit using strict criterion.

ERIC

Table 10

Number of Subjects Who Were Best Fit By Each of Three Regressions

In Two Treatment Groups -- Experiment 1

Treatment	Regression 1	Regression 2	Regression 3
Strict Criteria			
Equation Group	1	0	20
Word Group	2	13	6
Linient Criteria			
Equation Group	1	0	20
Word Group	8	9	4

Note. - For Strict criteria, $\chi^2 = 20.88$, df = 2, $p < .001$.

For linient criteria, $\chi^2 = 25.04$, df = 2, $p < .001$.

Table 11

Mean Response Time by Problem State for Two Treatment Groups on First and Last Trials -- Experiment 1

Treatment	Problem State													
	10	11	21	21[1]	22	22[1]	31	32	32[1]	33	51	52	53	54
First Trial														
Equation Group	3.2	2.6	3.6	4.8	5.6	5.5	7.9	7.2	7.7	10.9	15.1	19.9	18.6	35.6
Word Group	5.0	5.4	9.6	13.7	12.5	19.2	20.9	22.8	24.0	28.1	35.3	39.6	36.1	40.9
Last Trial														
Equation Group	2.0	1.9	2.5	2.8	3.5	4.7	3.7	4.4	4.6	5.5	11.6	13.2	12.8	21.2
Story Group	2.3	2.9	4.7	8.9	10.0	9.1	11.8	14.2	15.7	17.6	22.8	25.3	26.2	25.9

Note. - Main effect for treatment, $p < .001$; main effect for trial, $p < .001$; main effect for problem state, $p < .001$; interaction between treatment and problem state, $p < .001$; interaction between treatment, problem state and trial, n.s.

Table 12

Mean Response Time By Problem State for Two Treatment Groups - Experiment 2

Treatment	Problem State													
	10	11	21	21[1]	22	22[1]	31	32	32[1]	33	51	52	53	54
Word-to-Solution	6.4	9.1	12.8	15.3	14.3	17.1	23.0	21.4	24.1	24.4	35.9	41.1	38.6	42.8
Word-to-Equation	12.9	12.4	16.2	18.5	15.5	17.1	24.5	24.0	24.8	22.1	40.5	35.9	30.5	31.4

Note. - Main effect for treatment, p = n.s.; main effect for problem state, $p < .001$; interaction between treatment and problem state, $p < .001$.

Table 13

Values of R^2 and Variable Weighings for Four Multiple Regressions Fit to Two Treatment Groups

Experiment 2

Treatment	One Variable (Step)	One Variable (Calc)	Two Variables	Three Variables
Word-to-Solution	R^2 = .95 Step = 5.1 sec Intercept = -.9 sec	R^2 = .97 Computation = 8.1 sec Intercept = -.9 sec	R^2 = .98 Computation = 7.5 sec Move = 1.4 sec Intercept = -1.6 sec	R^2 = .98 Computation = 6.9 sec Move = 1.2 sec State = 1.8 sec Intercept = -.3 sec
Word-to-Equation	R^2 = .66 Step = 3.1 sec Intercept = 8.4 sec	R^2 = .92 Computation = 5.7 sec Intercept = 6.3 sec	R^2 = .97 Computation = 6.6 sec Move = -2.2 sec Intercept = 7.4 sec	R^2 = .97 Computation = 6.7 sec Move = -2.2 sec Stage = -.3 sec Intercept = 7.2 sec

Table 14

Transcription of Subject 3 for Problem 54

Equation	Subject's Comments
(8 + 3X)/2 = 3X - 11	The first thing I want to so is get all variables on one side. So, I was doing a lot of stuff. (Pause) First, I would add 11 to both sides.
(8 + 3X)/2 + 11 = 3X	Then I would, lets see. Oh, first I would have to divide, I mean multiply both sides by 2. (Points to first equation.) I usually write it with the fraction different. OK, so I mean here (points to first equation) multiply by 2.
8 + 3X = 2(3X - 11)	Oh, I would multiply the whole thing at one time.
8 + 3X = 6X - 22	In one step I would subtract 6X and add 8. Subtract 8.
-30 = -3X	Then X = 10.
X = 10	

ERIC

TECHNICAL REPORT SERIES IN LEARNING AND COGNITION

Report No.	Authors and Title
78-1	Peper, R.J. & Mayer, R.E. Note-taking as a Generative Activity. (Journal of Educational Psychology, 1978, 70, 514-522.)
78-2	Mayer, R.E. & Bromage, B. Different Recall Protocols for Technical Text due to Advance Organizers. (Journal of Educational Psychology, 1980, 72, 209-225.)
79-1	Mayer, R.E. Twenty Years of Research on Advance Organizers. (Instructional Science, 1979, 8, 133-167.)
79-2	Mayer, R.E. Analysis of a Simple Computer Programming Language: Transactions, Prestatements and Chunks. (Communications of the ACM, 1979, 22, 589-593.)
79-3	Mayer, R.E. Elaboration techniques for Technical Text: An Experimental Test of the Learning Strategy Hypothesis. (Journal of Educational Psychology, 1980, 72, in press.)
80-1	Mayer, R.E. Cognitive Psychology and Mathematical Problem Solving. (Proceedings of the 4th International Congress on Mathematical Education, 1980.)
80-2	Mayer, R.E. Different Solution Procedures for Algebra Word and Equation Problems.
80-3	Mayer, R.E. Schemas for Algebra Story Problems.
80-4	Mayer, R.E. & Bayman, P. Analysis of Users' Intuitions About the Operation of Electronic Calculators.
80-5	Mayer, R.E. Recall of Algebra Story Problems.
80-6	Mayer, R.E. Solving Nonsense and Meaningful Equations.