# Effects of meaningfulness and organization on problem solving and computability judgments

RICHARD E. MAYER and JAMES G. GREENO
*The University of Michigan, Ann Arbor, Michigan 41804*

When subjects were required to calculate answers for computable problems and answer questions, an interaction was found corresponding to that obtained by Kieras and Greeno (1975) from judgments of computability. With nonsense formulas, much longer times were required to identify noncomputable problems than to compute answers, with a much smaller difference when formulas consisted of meaningful concepts. The better performance on noncomputable problems and questions with meaningful formulas corroborates an interpretation that those items test the connection of algorithms with general conceptual knowledge. Finally, it was found that for relatively complex problems, solution times and time to judge computability were longer if nonsense formulas were learned in separate sets than if they were learned in a single set; however, no such effect was found with meaningful formulas. It was concluded that learning conditions influenced the integration of cognitive structure in the case of nonsense formulas, while subjects were able to adjust organization of the meaningful formulas.

One purpose of the present experiments was to examine effects of meaningfulness of formulas studied by Kieras and Greeno (1975) in a task requiring actual problem solving rather than just a judgment of computability. A second purpose, involved in Experiments III and IV, was to study the effect of overall organization of a set of formulas, manipulating organization in a different way than by the supplementary training used in Kieras and Greeno's Experiment II.

The varying meaningfulness of formulas also is related to a theoretical distinction used by Mayer (1973) and Mayer and Greeno (1972) in interpreting results of experiments where subjects learned mathematical concepts. We concluded that subjects whose instruction emphasized general concepts acquired new cognitive structures with strong external connectedness, meaning that the new material was strongly embedded in the subjects' general knowledge structures. On the other hand, subjects whose instruction emphasized algorithmic calculation were said to acquire new structures with relatively weaker external connectedness, but strong internal connectedness, meaning that the elements of the new concepts were strongly connected to each other in the operations involved in computation according to the learned formula. This conclusion was based on performance on different kinds of test problems in which subjects with emphasis on algorithmic calculation excelled on test problems that involved straightforward computation, while subjects with emphasis on meanings of concepts excelled on test problems that involved interpretation. A similar variety of problems was used in

the present experiments, permitting comparison with results obtained in the earlier instructional experiments.

## EXPERIMENT I

This study compared use of meaningful and nonsense formulas in solution of a variety of problems, including computational problems, questions in which subject had to retrieve information about the formulas, and problems that cannot be solved using the formulas where the subject had to recognize the impossibility of solution.

### Method

**Subjects, Design, and Materials.** Twenty female University of Michigan undergraduates served as subjects and were paid for their participation. The design was a 2 by 5 factorial with repeated measures, where the between-groups factors were the meaningfulness of formulas used and the order in which test items were presented. There were five different types of test problems given, and the same problems were given to all subjects.

The formulas were the same as those used by Kieras and Greeno (1975) in their Experiment I (see their Table 1).

Test problems were of five types. Type F (familiar) problems used a memorized formula directly. For example, "Distance = 90 miles, Gas used = 5 gal, Find Gas mileage." Type T1 (single transformation) used a single formula, but the unknown was not the left-side member of the memorized equation. For example, "Preparation time = 1/2 h, Total time = 2 h, Find driving time." Type T2 (two transformations) required use of two formulas, solving for an intermediate result and then using that to find the unknown quantity. For example, "Arrival time = 6:00, Leaving time = 4:00, Average speed = 25 mph, Find distance." Type Q (question) asked for information about the formulas. For example, "Given arrival time and leaving time, what else is needed to find total time?" And Type U (unanswerable) gave inconsistent or incomplete information so the answer could not be computed. For example, "Average speed = 50 mph, Driving time = 1 h, Find gas mileage." Test items based on the nonsense formulas were constructed by substituting the corresponding letter variables for the variable names in the meaningful items.

## Table 1
## Mean Times* (Seconds) to Solve Problems

| | Problem Type | | | | |
|---|---|---|---|---|---|
| | F | T1 | T2 | Q | U |
| Meaningful | 16.7 | 21.5 | 37.2 | 24.0 | 27.5 |
| Nonsense | 20.5 | 22.6 | 39.3 | 36.5 | 63.9 |

*Approximate standard error for cell means is 5.5 sec.

The numerical values were the same, and the names of units (e.g., "hours") were omitted.

**Procedure.** Subjects who passed the pretest were assigned to treatment groups by a random, predetermined order and were tested individually in this experiment and in each of the experiments reported subsequently. First, the same pretest used by Kieras and Greeno in their Experiment I was given; four subjects made one or more errors and were replaced in the experiment. After the pretest, subject was given a sheet of paper with the four formulas she was to memorize. The meaningful formulas were accompanied by a paragraph describing a driving trip and mentioning the relations among the variables. The subject was instructed to memorize the formulas in any way she wished to, and to be prepared for a test involving them. After a 20-min study period, subject was asked to write out the equations from memory; all subjects succeeded.

Instructions for the test were given, including instruction to write "no answer" if subject decided a problem was impossible or there was not enough information. Subjects were permitted to write out the formulas and do figuring on the answer sheet provided. The test items were given to subjects on index cards, with one item per card. Five orders of the items were used. In each order, the four items of each type were presented successively. The order in which types were presented was determined by a 5 by 5 Latin square, and the orderings of the four problems within each type was determined by a series of 4 by 4 Latin squares. Subjects worked the problems at their own rates of speed, turning up a new card each time they wished to being a new problem. The experimenter timed subject on individual problems by noting the times between turning of cards.

## Results

Nearly all the problems were solved correctly; error rates were .04 with meaningful formulas and .08 with nonsense formulas.

The main data were mean solution times for problems solved correctly. Each subject's score on this measure was her mean solution time on those problems of a given type that she solved correctly. Table 1 shows the mean solution time for problems of the five types with the two kinds of formulas. The main effect due to meaningfulness was significant $F(1,10) = 30.6$, $p < .001$ as was the main effect due to type of problem $F(4,40) = 33.4$, $p < .001$. The most interesting aspect of the result was the strong interaction between meaningfulness and type of problem $F(4,40) = 14.8$, $p < .001$, where the difference between nonsense and meaningful formulas was much greater on questions and unanswerable items than on the more straightforward familiar and transforming problems. The average times found for unanswerable problems were nearly identical to those observed for the first unanswerable problem encountered which were 29.4 and 63.0 sec, respectively, for meaningful and nonsense formulas.

## Discussion

The results were consistent both with Kieras and Greeno's findings regarding judgments of computability, and with Mayer and Greeno's earlier work with instructional methods. The unanswerable items are analogous to negative items in the computability judgment task, and the long time taken by subjects to decide that a nonsense problem was unanswerable is analogous to the long times taken for negative computability judgments with nonsense formulas.

One feature of the problem-solving process here involves the temporal relationship between finding a method of solution and carrying out calculations. It is clear from the long times taken on unanswerable items that subjects proceeded to carry out calculations at the beginning of problem solving, rather than waiting to see how a complete solution could be obtained. Had subjects first worked out a solution method, they would carry out no calculations at all on the unanswerable problems, and times for those problems would have been quite short. The exceptionally long times for the unanswerable problems with nonsense formulas can be understood if we hypothesize that subjects in that condition carried out a relatively exhaustive set of calculations, using all the information they were given in the problems, only infrequently realizing that the information was insufficient until all possibilities had been tried. The much shorter times for unanswerable problems with meaningful formulas can be explained if we hypothesize a modest amount of look-ahead during problem solving, which would provide earlier recognition of the impossibility of solution.

The results provide a useful extension and confirmation of Mayer and Greeno's findings involving problem solving after different methods of instruction. The longer times needed to answer questions and discover impossibility of solution in the nonsense condition are analogous to the poorer performance on questions and unanswerable problems by subjects taught the binomial formula by a method emphasizing algorithmic calculation rather than meanings of concepts. Our earlier interpretation was that better performance on questions and unanswerable items resulted from having the material more strongly embedded in the subjects' semantic memory when instruction had emphasized the meanings of concepts. That interpretation is strengthened by the present result, where the same kind of interaction was found comparing meaningful formulas with formulas composed of uninterpreted letters. The results are both consistent with the idea that on items where subjects answer interpretive questions and on problems that are not answerable, performance depends strongly the presence of usable relations between the concepts involved in the items and the general structures of the subjects' semantic memories.

## EXPERIMENT II

This experiment used an important procedural control that was lacking in Experiment I. In the first experiment, some subjects wrote the formulas on the answer sheets provided by the experimenter, and other subjects did not use this form of external memory. An examination of the performance of the subjects who did and did not write out the formulas revealed no importance differences, but we decided to carry out a controlled comparison of performance under the two conditions.

## Method

**Subjects, Design, and Materials.** Subjects were 48 paid female University of Michigan students, who gave correct answers on all six items of an algebra pretest. Seven potential subjects were replaced because of pretest errors.

The design was a 2 by 2 by 6 factorial, with meaningfulness as the first factor, availability of the formulas during problem solving as the second factor, and order of test items as the third factor.

Each subject learned the four formulas using one of the two study sheets used in Experiment I. Six types of test items were constructed with items of Types F, T1, T2 and Q exactly the same as those used in Experiment I, and two types of unanswerable items. Items of Type U1 had two variables designated as givens (e.g., A = 50, V = 1, find M) and items of Type U2 had more than two givens (e.g., D = 100, A = 50, M = 25, find P). As in Experiment I, unanswerable problems either had incomplete information or inconsistent combinations of information. Four problems of each kind were used, making a test of 24 items in all.

**Procedure.** Procedures were the same as in Experiment I except for the use of answer sheets that varied between conditions. One group of subjects with each kind of formula had answer sheets with numbered blanks on the right side of the sheet, the rest of the page being available for calculations. Subjects in these groups were instructed not to write the formulas on the answer sheets. Subjects in the other groups had answer sheets with the formulas they had memorized reproduced in the upper left corner of the page.

## Results

Error rates were very low, ranging from .015 to .039 in the four conditions.

Mean solution times are given in Figure 1. Main effects of condition F(3,24) = 3.3, p < .05 and type of problem F(5,120) = 41.1, p < .001 were significant. Also significant was the interaction between condition and problem type F(15,120) = 4.3, p < .001. Again, this interaction constitutes the main result of the experiment, with problems of Types F, T1, and T2 requiring about the same amount of time with meaningful and nonsense formulas, and with much greater times needed for questions and unanswerable problems with nonsense formulas than with meaningful formulas. There was no substantial difference in this pattern of results between the conditions having visual access to the formulas and the conditions in which the formulas had to be remembered.
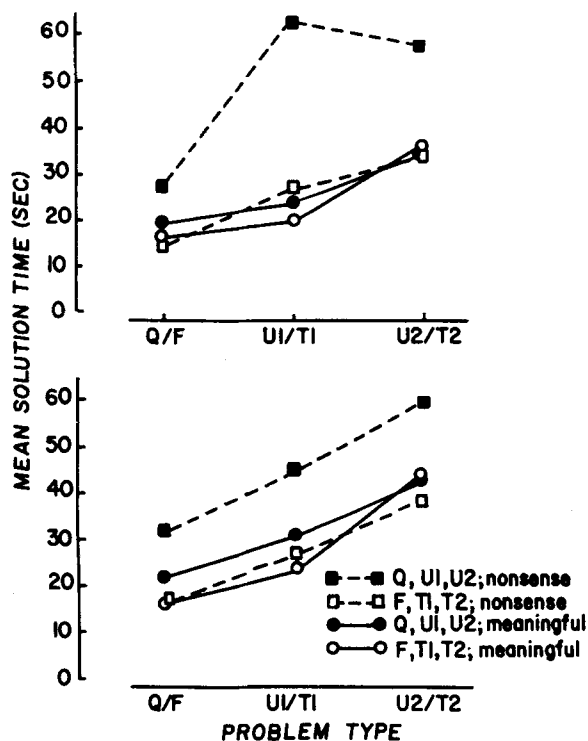


**Figure 1.** Mean times to solve problems, answer questions, and identify unanswerable problems in Experiment II. Upper panel shows times when formulas were visually present, lower panel shows times when formulas had to be recalled from memory. $SE(\bar{X})$ = 8.9 sec.

## Discussion

Since the interaction between problem type and condition occurred with about equal strength regardless of whether subject could see the formulas or had to remember them, we can conclude that the effect is not due to a process that involves retrieval from long-term or intermediate-term working memory. It seems more likely that the difference between meaningful and nonsense formulas that causes the interaction involves the ease with which they can be manipulated in subject's span of apprehension—that is, the ease with which subject can compare formulas, or, while considering one formula, also consider whether there is another formula with a certain property. In other words, it seems likely that the functional characteristics involved in the effects we obtained have to do with a system in which information is manipulated and processed, rather than by the function of holding or maintaining a representation of information by rehearsal.

## EXPERIMENT III

The purpose of this experiment was to investigate the effect of manipulating the strength of relationships between different formulas in the set used for solving problems. An attempt to strengthen relationships between formulas in cognitive structure was carried out by Kieras and Greeno (1975, Experiment II), without

**Table 2**
**Formulas Used in Experiments III and IV**

| | Meaningful | Nonsense |
|---|---|---|
| A1 | Volume = area × height | V = A × H |
| A2 | Height = stopping point − starting point | H = E − B |
| A3 | Area = length × width | A = L × S |
| B1 | Word = weight × distance | W = O × D |
| B2 | Potential energy = weight × height | M = O × H |
| B3 | Power = work/time | R = W/T |
| C1 | Density = weight/volume | N = O/V |
| C2 | Weight = mass × acceleration | O = Y × C |
| C3 | Pressure = weight/area | F = O/A |

significant success. In the present experiment, we taught a set of nine equations, in three sets of three equations each. Then problems were given, and performance was compared on problems solved with formulas taken from a single set, and problems requiring use of formulas from two or three different sets.

## Method

**Subjects, Design, and Materials.** Subjects were 48 female University of Michigan undergraduates, paid for their participation. Twelve subjects made errors in the algebra pretest and were replaced.

The design was a 2 by 2 factorial, involving meaningfulness of formulas as one of the factors. The other factor varied whether a subject's test problems came from one of the sets of formulas learned or from different sets presented during learning. There were also 12 different conditions involving which of two organizations of equations was used, which of two sets of test problems was given, sequence of presenting equations during training, and sequence of test items. For statistical purposes, there was one subject in each cell of a 2 by 2 by 12 factorial design.

The formulas used are shown in Table 2, in both their meaningful and nonsense forms. The formulas were constructed so they could be grouped into sets of three in two different ways. Note that formulas denoted A1, A2, and A3 form a group in which each formula shares a variable with at least one other formula in the group. A1 and A2 both have height and A2 and A3 both have area. Groups of the same kind are formed by Formulas B1, B2, and B3 and by Formulas C1, C2, C3. The other grouping arrangement uses A1, B1, and C1 as a group. A1 and C1 both include volume, and B1 and C1 both include weight. In this arrangement, groups also are formed from A2, B2, C2 and from A3, B3, C3.

Each subject learned the formulas in three separate teaching booklets. For one-half of the subjects, one booklet contained Formulas A1, A2, A3; another booklet contained Formulas B1, B2, B3; and the other booklet had Formulas C1, C2, C3. For the other subjects, the booklets had A1, B1, C1; and A2, B2, C2; and A3, B3, C3. Each booklet had seven pages. (1) The three formulas were presented. For the meaningful formulas, there was also a brief paragraph involving a physical situation such as "things involved in dropping an object from a certain height," and explaining how the variables were related. (2) A blank page was included to be used by subject for writing out the formulas she was memorizing. (3) One or two questions asking which variable(s) were involved in more than one formula. (4) Answers to questions on Page 3, with instructions to study the formulas again if answers were incorrect. (5) Questions asking subject to write out the formulas containing the variable or variables occurring in more than one formula. (6) Answers to the questions on Page 5, with instructions to study the formulas again if answers were incorrect. (7) Instruction to write out all three of the formulas.

Two sets of test problems were constructed using the meaningful formulas, with a set of problems using nonsense formulas matched with each set, giving four sets in all. One of the two sets using each set of formulas used Equations B1, B2, and B3. The other problem set used Formulas A2, B2, and C2. The problems were selected for individual subjects so that one-half of the subjects receiving each set of problems had received the formulas she needed as a single set during training, and the other subjects had received the three formulas needed for problem solving as members of separate training sets.

Each of the problem sets included 18 problems, with three of each type of problem used in Experiment II.

**Procedure.** Subjects were tested one, two, or three at a time. The subjects were given the pretest and a brief explanation of the experiment. Then the first of the three teaching booklets was given. Subjects were told they could spend 4 min memorizing the three formulas shown on the first page, using the second page as scratch paper. After 4 min elapsed, the experimenter instructed subjects to continue through the teaching booklets, answering questions, checking answers, and finally writing out the three formulas from memory. This was repeated for the other two teaching booklets. After the third booklet was completed, subjects were asked to write out all nine formulas on blank sheets of paper. If a subject failed to do this within 2 min, the booklets containing missed formulas were returned to the subject for further study.

After subjects could write all nine of the formulas, instructions were given for the problem-solving test. Subjects were given answer sheets with instructions that they could use the space on the sheet for figuring and that they could write out formulas needed in specific problems but they could not write out the nine formulas at the beginning. Subjects were also instructed to write "no answer" if they felt a problem was impossible or there was not enough information.

Test problems were typed on index cards in sequences determined by Latin squares, and were worked by subjects at their own pacing, as in the earlier experiments.

## Results

Error rates were somewhat higher than in earlier experiments, ranging from .10 to .14 in the various conditions. There were no substantial differences among the conditions in the number or kinds of problems missed.

Mean solution times for correct answers were calculated in the same way as in Experiments I and II, obtaining each subject's mean time for each kind of problem, and then averaging these mean times. The results are in Figure 2. Main effects due to meaningfulness $F(1,11) = 5.5$, $p < .05$ and type of problem $F(5,55) = 45.8$, $p < .001$ were significant. A significant interaction between type of problem and meaningfulness was obtained $F(5,55) = 5.5$, $p < .005$; note that as in Experiments I and II, the time taken on questions and unanswerable problems was disproportionately large with nonsense formulas. We replicate the finding that with nonsense formulas a greater difference between unanswerable and question items on the one hand and straightforward calculation problems on the other.

The new factor in the present experiment was the difference in organization of the formulas involved in the problems to be solved. This factor of grouping did not have a significant main effect, but the more

interesting question is whether grouping had selective effects on some kinds of problems. The statistical analysis indicated that such a selective effect occurred. The three-way interaction involving meaningfulness, type of problem, and grouping was significant $F(5,55) = 7.3$, $p < .001$.

A reasonable expectation is that the effect of grouping should be concentrated on items of Type T2 and U2, and to some extent on Type Q. The reason is that with Types F, T1, and U1, only one formula needs to be retrieved from memory to determine the answer. However, for Types T2 and U2, and for about one-half the items of Type Q, two or more formulas need to be remembered and related to each other for the answer to be obtained. The difference in organization gave a difference in the strength of organization or connection between the formulas involved in problems; thus, it would be expected that organization would particularly affect difficulty on the items where more than one formula had to be used. This would produce a two-way interaction between type of problem and grouping, with a greater difference between complex problems (T2, U2, and Q) and simpler problems (F, T1, and U1) for the condition in which formulas came from three sets and a smaller difference between complex and simple items when the formulas came from a single set. This particular partitioning of the test items is supported by the fact that the response time of [(nonsense separate sets—nonsense one set) — (meaningful separate sets—meaningful one set)] is highest for T2, U2, and Q items (8, 13 and 19 sec, respectively) and lowest for F, T1, and U1 items (-12, 0, and 2 sec, respectively). An analysis of variance based on the difference in response times on T2 + U2 + Q - F - T1 - U1 reveals a reliable difference among the four treatment groups $F(3,44) = 4.03$, $p < .025$, with a two-way interaction of 166 sec and a SE (contrast) of 83 sec.

It is of considerable interest that the two-way interaction spelled out above was virtually nonexistent for the meaningful formulas. The mean times for complex and simple problems from a single set were 53.0 and 25.2 sec, while from three different sets the times were 51.9 and 21.3 sec. Thus, there was a 2.8-sec interaction. However, for the nonsense formulas there was a substantial interaction. For the formulas from a single set, times for complex and simple problems were 60.9 sec and 36.4 sec; from three different sets the times were 73.1 sec and 27.8 sec; the interaction was 20.8 sec. The obtained three-way interaction involves more complex comparisons, of course, but this difference between interactions surely made a contribution to the significant three-way interaction and may have been the major factor.

### Discussion

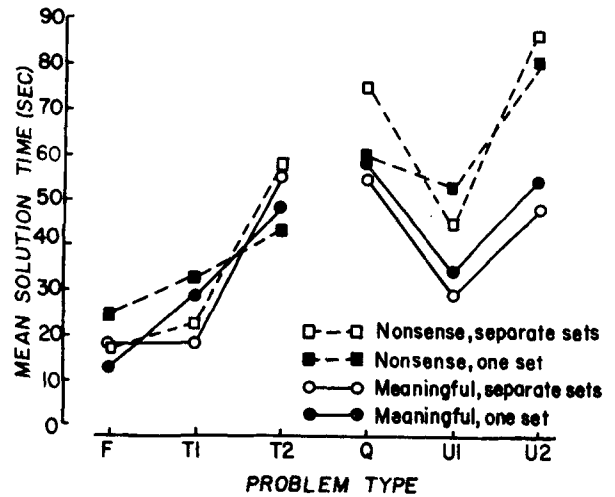An interpretation of the three-way interaction is that with meaningful formulas, either the nine equations



Figure 2. Mean times to solve problems, answer questions, and identify unanswerable problems in Experiment III. SE($\overline{X}$) = 10.1 sec.

formed an integrated structure at the time the formulas were learned, or when subject encountered the first few problems involving the set of formulas used she easily formed a structure integrating the three formulas. However, it seems that, with the nonsense formulas, the integration of the formulas was not accomplished, at least to the same degree, when the formulas had been learned as parts of separate sets. Thus, the results provide evidence that meaningfulness can supply a basis for achieving a cognitive structure with stronger integration among the various components.

### EXPERIMENT IV

This study was carried out to determine whether the effect of organization obtained in Experiment III would also occur when subject's task was to judge computability.

### Method

**Subjects, Design, and Materials.** Twenty-four female and 24 male University of Michigan students participated for pay. Each subject served in one cell of a 2 by 2 by 2 factorial design. The factors were meaningfulness of the formulas and grouping of the formulas used for test items, as in Experiment II, and the gender of the subjects. In addition, the specific combination of formulas learned and test items used was balanced as in Experiment II.

The formulas, instructional booklets, and pretest were the same as those used in Experiment II. In addition, 40 test items were constructed, each consisting of a list of variables with one designated as the unknown and the other two, three, or four as given variables. These test items were given in four subblocks of 10 items each, with each block having 2 computable and 2 noncomputable items with two givens, 2 computable and 2 noncomputable items with three givens, and 1 computable and 1 noncomputable item with four givens.

**Procedure.** Subjects were tested individually. Procedure for study from the booklets was the same as in Experiment III. For the test items, subject was seated in a desk about 1 m from a screen and was given a control box with two buttons marked Yes and No. The subject was told that two, three, or four variables
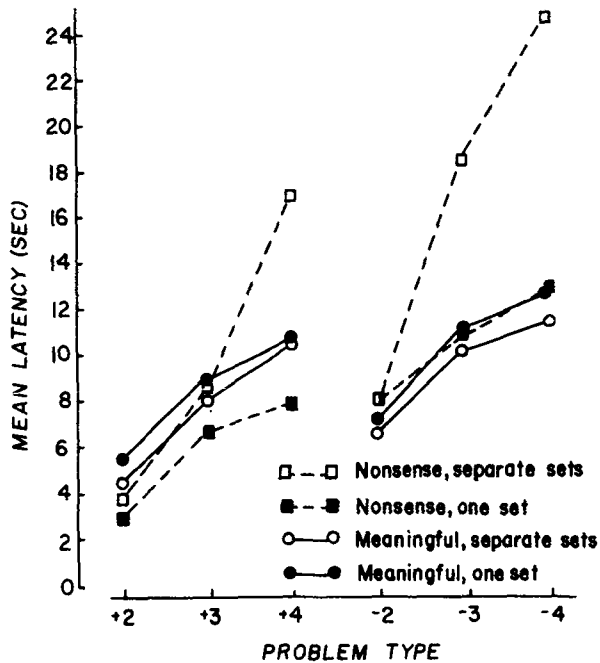
Figure 3. Mean latencies of judgments of computability in Experiment IV. Problem types labeled + were computable, - were not computable. Numbers 2, 3, and 4 denote numbers of given variables. $SE(\bar{X}) = 4.1$ sec.

would be shown above a line, and one variable would be below the line on each trial. The subject was told to consider the variable below the line as the unknown and press the Yes button if it would be possible to calculate the value of the unknown if values were given for the other variables shown, otherwise to press the No button. The subject was asked to respond "as soon as you are fairly sure you are correct."

Two practice trials were given, with feedback, to ensure that subject understood the task, and then the 40 test items were presented using a Carousel slide projector. On each trial, the experimenter said "Ready," and pressed the advance button which displayed the next slide and started a timer. When subject pressed a response button, the timer was stopped and the projector advanced to a blank slide. No feedback was given. After recording the response and the time and resetting the timer, the experimenter said "Ready" again and presented the next trial.

The four 10-item subblocks were given in an order determined by a 4 by 4 Latin square, and the ordering of items within subblocks was randomized. If subject gave an incorrect response to an item, the item was repeated at the end of the 40-item block in order to obtain latency measures for correct responses for all items. After all 40 items were given and answered correctly, subject was given a short rest while the experimenter rearranged the slides for the second block of trials and the 40 items were presented a second time.

## Results

Eight subjects failed to solve at least four of the six pretest problems, and three subjects failed to reproduce the to-be-memorized equations after repeated study. Data for these subjects were eliminated, and new subjects were tested in their places. In the test, error rates were very low, with an average of .05 of the items needing to be repeated, and with most of these occurring in the first block.

A large practice effect was obtained during the first block, with mean latency of 21.5 sec during the first 10-trial subblock and 12.2 sec during the fourth subblock. On the other hand, latency was reasonably stationary during the second block, averaging 9.5 sec in the fifth subblock and 8.5 sec in the final subblock. Differences in latency between male and female subjects, and interactions involving gender of subjects, were small and unreliable.

The main results of the experiment are in Figure 3, which shows mean latencies obtained during the final 40-trial block of trials. Analysis was also carried out for the total set of 80 responses, with no important differences in the effects of treatment variables.

The main results of Kieras and Greeno's experiments was an interaction between meaningfulness of formulas and correct answer of items, with negative items taking much longer than positive items with nonsense formulas and a smaller difference with meaningful formulas. That interaction appears in the data shown in Figure 3, and it was highly reliable $F(1,40) = 9.69$, $p < .005$. The two-way interaction between meaningfulness and number of givens was significant $F(2,80) = 3.93$, $p < .025$ as was the interaction between grouping and number of givens $F(2,80) = 5.66$, $p < .01$. However, the major finding involved the three-way interaction of meaningfulness, grouping, and number of givens $F(2,80) = 5.85$, $p < .005$. There was virtually no interaction between grouping and number of givens for meaningful items, but there was a strong interaction between these variables for the nonsense formulas.

## Discussion

In making judgments of computability, as in computing solutions of problems, grouping of formulas during training had an effect when formulas were not meaningful and did not affect performance when formulas were meaningful. The nature of the effect was that problems requiring use of more than one formula took much longer if formulas had been learned in different sets than if formulas had been learned together.

## GENERAL DISCUSSION

These results provide some information on the effect of meaningfulness of learned information in problem solving. A fairly straightforward idea—one which defines "meaning" in purely quantitative terms—is that nonsense information may take longer to learn or require longer retrieval time overall, but there is no qualitative difference between nonsense and meaningful information in terms of its storage or processing. A second position is that information is stored and processed in qualitatively different ways depending on its meaning to a learner. According to this view, meaningful statement of the formulas encouraged the development of what we have called "external

connections" (Mayer & Greeno, 1972)—links between the information presented and general ideas or contexts already in subject's semantic memory. For example, the "potential energy = weight x height" formula may be externally related to subject's experience with a heavy object being held by a crane (i.e., subject's experience provides a context for assimilating the new information). This context would not be provided by formulas such as "carrots = shoes/distance" or "M = O x H" because subject's experience here is purely mnemonic. Thus nonsense formulas encouraged the development of "internal connections"—strong rigid links between the variables in the formula as presented, and between triplets of presented formulas.

While the first model predicts no difference in the effect of presentation organization for the two groups, the second idea—that qualitatively different problem solving structures are built—is consistent with the main results, namely a three-way interaction among meaningfulness, problem type and organization (Experiments III and IV) and a two-way interaction between meaningfulness and problem type (Experiments I and II). The fact that presentation organization seems to be retained more for nonsense information than for meaningful suggests that allowing subject to relate problem solving information to his general cognitive structure produces a restructuring and integration of the new information.

The results of Experiments III and IV show that problem-solving performance can be influenced by the degree of integration among components of the cognitive structure serving as the knowledge base for solution of the problem. The evidence for this overall integrative function was not of a surprising kind; it involved specific facilitative effects on problems requiring subjects to use two or more formulas in solving the problem. A straightforward interpretation is that the additional time for the longer problems involving nonsense formulas indicates greater cognitive distance or less cognitive integration among the nonsense formulas that we learned in different sets. There was apparently little or no difference in the time needed to retrieve a single formula and use it to solve a problem or judge whether a problem was or was not computable. However, when the calculation or judgment required use of more than one formula, extra time was needed to retrieve or process the relations between formulas, and this additional time was greatest for nonsense formulas that were learned in different sets. Such an effect of added time would be expected if the time to retrieve and process a set of formulas depended on the degree of integration of those formulas in memory, and if the nonsense formulas were less strongly integrated when they were not learned together as a group.

It is of considerable interest that a facilitative effect of having the needed formulas learned as a set occurred only with the nonsense formulas. Apparently when the formulas were meaningful, subjects' familiarity with the concepts to which they referred was sufficient to provide the basis for spontaneous integration into the needed subset for efficient performance either in problem solving or in judgments of computability.

It would be pleasant if the results regarding structural integration could shed light on the interaction between meaningfulness and computability found consistently by Kieras and Greeno (1975) and corroborated in the findings reported here. Unfortunately, the connection between the two results is not conclusive. The data of this study lack strong indications of interaction between the two effects. The interaction between meaningfulness and computability appears to have occurred both for the formulas learned as single sets and for formulas learned in separate sets. Although the amount of the interaction was somewhat larger for the separate sets than for the single sets, the three-way interaction was not significant, and since the larger difference occurred in the condition with longer times, it would be difficult to justify strong theoretical conclusions even if a reliable interaction had occurred.

## REFERENCES

Kieras, D. E., & Greeno, J. G. Effect of meaningfulness on judgments of computability. Memory & Cognition, 1975, 3, 349-355.

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

## NOTE

1. Mayer, R. E. Acquisition and resilience under test stress of structurally different problem solving procedures. University of Michigan: Human Performance Center Technical Report No. 42, 1973.