DOCUMENT RESUME

ED 027 582 24 CG 003 650

By- Johnson, Donald M.

Improvement of Problem Solving Processes. Final Report.

Michigan State Univ., East Lansing. Dept. of Psychology.
Spons Agency- Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No- BR-5-0705
Pub Date Dec 68
Contract- OEC-6-10-013
Note- 84p.
EDRS Price MF-$0.50 HC-$4.30
Descriptors- College Stude .ts, Decision Making Skills, Divergent Thinking, *Learning Processes, *Problem Sets, *Problem Solving, Productive Thinking, *Psychological Studies, *Transfer of Training

Different groups of college students wrote one solution or many solutions to problems, like the plot-title problem, of verbal, numerical, and pictorial material. Instructions to write many solutions yielded solutions of lower mean quality, but more superior solutions. Information about criteria for good solutions raised quality. Large quantity was associated with low quality, both for variations in conditions and for individual differences within conditions. Differences between problems in the quality-quantity relation were dependent on the number of superior solutions. These three types of judgment training: (1) individual, (2) dyadic, and (3) tutorial, interpolated between production of solutions and selection of best solution, were generally successful and, under certain favorable conditions, improved over-all performance. The response hierarchy model can be modified to apply to the present data, and transfer of training can be demonstrated between tasks. (BP)

BR 5-0705
PA-24
OE-BR

FINAL REPORT
Project No. 5-0705
Contract No. 6-10-013

IMPROVEMENT OF PROBLEM SOLVING PROCESSES

Donald M. Johnson
Department of Psychology
Michigan State University
East Lansing, Michigan 48823

December 1968

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

Project No. 5-0705
Contract No. 6-10-013

IMPROVEMENT OF PROBLEM SOLVING PROCESSES

Donald M. Johnson

Department of Psychology

Michigan State University

East Lansing, Michigan

December 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

PREFACE

This is the final report on a project, Improvement in Problem Solving Processes, conducted at Michigan State University from October, 1965 to December, 1968. The project was supported by the University and by the United States Office of Education, Basic Studies Branch, Project No. 5-0705. The principal investigator was Donald M. Johnson, but in the final year George L. Parrott and R. Paul Stratton carried most of the responsibility for design and execution. Bruce Leibrecht joined the project for the summer of 1967, Stuart Agres for the summer of 1968.

The studies reported in this volume continue earlier research, which is reported under the title, "Production and Judgment of Solutions To Five Problems", in a Supplementary Monograph, Journal of Educational Psychology, December, 1968. Hence, for the convenience of the reader, an abstract of the monograph is reprinted as part of the Summary.

SUMMARY

Some of the research conducted under the present contract has been published as a monograph: Production and Judgment of Solutions to Five Problems. The abstract of the monograph follows.

> Different groups of college Ss wrote one solution or many solutions to problems, like the plot-title problem, of verbal, numerical, and pictorial material. Instructions to write many solutions yielded solutions of lower mean quality but more superior solutions. Information about criteria for good solutions raised quality. Large quantity was associated with low quality, both for variations in conditions and for individual differences within conditions. Differences between problems in the quality-quantity relation were dependent on the number of superior solutions to the problem. Three types of judgment training, individual, dyadic, and tutorial, interpolated between production of solutions and selection of the best solution, were generally successful and, under certain favorable conditions, improved overall performance.

Previous research has demonstrated that students instructed to write many solutions to a problem write solutions of lower mean quality than students instructed to write one solution. To determine the relation between quality and quantity more precisely different groups of students were instructed to write 1, 2, 4, or 6 solutions. The results were consistent with previous research and could be reasonably well described by a decreasing function of the form: $Y = K - C \log X$. The decline of mean quality with the increase in quantity requested was greater for the Conclusions problem than for the Sentences problem, presumably because the former has a small number of good solutions. The quality-quantity relation can be viewed as a quantification of the difference between convergent and divergent thinking.

An attempt to examine the processes underlying productive thinking manipulated the time allowed and the instructions. Previous research has shown that the quality of solutions to certain problems increases with continued production when standard instructions are given but not when quality is emphasized. To clarify these effects students were given .5, 1, 2, or 4 minutes to write one solution, and some were given standard instructions, some were given the criteria of good solutions, and some were given the criteria and a rating guide for judging their solutions. Others were instructed to write many solutions in 4 minutes, with criteria cues. There were four problems: Sentences, Conclusions, Cartoon Captions, and Plot Titles.

In general the results support the hypotheses. Information about the criteria of good solutions was helpful for students writing one solution as it has been, in previous experiments, for students writing many solutions. Those who write one solution with such information write much better solutions, on the average, than those who write many solutions with the same information. The effects of time restrictions were apparent for the Sentences and Conclusions problems but not for the other two.

The response hierarchy model that has been applied to simpler tasks can be modified to apply to the present problems and variations. When subjects are asked to write many solutions to a problem, they write the more conventional solutions first, either because these solutions have more habit strength from past experience or because they are more readily constructed from obvious relations between familiar components. As the conventional solutions are exhausted, fewer but less common ones are produced. Thus solutions produced in the last half of a production sequence are often of higher quality than those produced in the first half. This production order effect does not occur when information about good solutions is given because such information restricts production of the low-quality solutions. When subjects are asked to write one solution, they may think of more than one but write only the best. In the case of Sentences and Conclusions, some time is required for the writing itself, hence only those subjects with two or four minutes working time proceed to the better solutions. In the case of Plot Titles and Cartoon Captions, which can be written quickly, time limitations are not restrictive and subjects in all conditions were able to proceed to their better solutions.

The final study, on facilitation of production and judgment, raised the critical question of transfer. Will the beneficial effects of judgment training transfer to other problems, and if it does, just what is it that transfers? The experiments used two parallel sentence problems which required the subjects to make sentences out of four given words and gave all groups the criteria for good solutions. In the multiple-solution experiment there were three groups of 40 subjects each, one of which produced solutions to Sentence 1, then had judgment training in pairs, followed by Sentence 2; another group had practice in production and judgment of solutions to Sentence 1, followed by Sentence 2, and a control group had only Sentence 2. The judgment training following Sentence 1 decreased the number and increased the quality of solution produced to Sentence 2. The effects of the judgment training showed up on Sentence 1 in that the solutions the subjects selected as their best were better than their nonpreferred solutions, but this effect was not so noticeable on Sentence 2 because, as a result of the training, the subjects wrote fewer low-quality solutions to Sentence 2. Overall, the solutions which the trained group produced and selected as their best,on Sentence 2 were distinctly superior to those of the control group.

A companion experiment with the same problems and the same procedures and three comparable groups of 40 subjects required each subject to write only one solution, and this experiment also demonstrated significant transfer from the first problem to the second in the quality of solution produced and in scores on a standard judgment test. In fact those who wrote only one solution reached a distinctly higher level of performance on Sentence 2 than those who wrote many and selected their best.

In general, these experiments with brief but challenging problems have shown that (1) college students who write many solutions to problems write more superior solutions (and many more inferior solutions) than those who write only one; (2) instructions to write many solutions reduce mean quality of the solutions; (3) information about the criteria of good solutions improves the quality of the solutions produced; (4) without criteria information quality increases as production continues; (5) students who write many solutions can select the best of these with fair accuracy; (6) the accuracy of this selection can be improved by carefully programmed training in judgment; (7) some of the improvement transfers to a second problem; (8) the highest performance is attained by those who write only one solution with the benefit of the criteria of good solutions and of training in judgment of solutions.

INTRODUCTION

Problem solving, like learning, is not a simple homogeneous activity. This truism is important for the analysis of problem solving because a statement that holds for one component process may not hold for another. It is important also for attempts to improve problem solving because a procedure that facilitates one process may not facilitate another. Therefore the research to be described separates problem solving into three different but functionally interdependent processes: preparation, production, judgment.

Intellectual tasks begin with some kind of preparation, most often the acquisition and organization of information, as by listening to instructions or by reading a printed paragraph. After preparation, some tasks are primarily productive, as in writing many uses of a brick, and some are mostly matters of judgment, as in selecting the best answer on a multiple-choice test. The present research is concerned with that large class of problems, not often investigated, for which both production and judgment are required. The subject (often abbreviated S) produces several possible solutions, then examines these and picks one as his best effort. In such cases production is different from preparation but depends on it. Judgment is different from production, but it is the solutions produced that are judged. The different processes are interdependent, but they can be studied separately, and conditions that influence each can be experimentally manipulated.

The present research is focused on problems with many solutions that cannot be dichotomized as right or wrong but can be graded in respect to such qualities as usefulness, appropriateness, cleverness, and originality. This might be called productive thinking as well as problem solving. In Guilford's terminology it is divergent thinking as opposed to convergent thinking. Another consideration was that the problems should be substantial problems that would offer some challenge to college students. Tasks as simple as writing uses for a brick or giving uncommon associations to words have been criticized as trivial. Any punster knows that the set for simple verbal productions is highly vulnerable to variations in instructions and social atmosphere.

It was necessary to choose problems yielding solutions that vary considerably in quality and that can be reliably rated as to quality. Finally, within the restrictions of the design, the problems should vary in content. The literature on problem solving contains many findings that apply, as far as is known, to only one problem. The use of several problems permits general principles to emerge as well as differences between problems.

The problems selected, after considerable preliminary research, are briefly described below, together with instructions for multiple solutions.

Plot Titles. A paragraph gave the plot of a story or movie, with the following instructions. "Your task is to think of titles for the story. Read the plot then write as many titles for it as you can."

Table Titles. A table of agricultural data, showing four columns of statistics for seven time periods, was printed, with the following instructions. "Table X below has reference to United States statistics, and was taken from a past volume of the World Almanac. You are to examine the table and then write as many titles for it as you can."

Conclusions. A chart was printed, with column diagrams representing social welfare expenditures under five public programs for six time periods, and the following instructions. "Fig. XVI is taken from The Statistical Abstracts of the United States, 1964. What can you conclude from this table? Write short sentences, as many as you can, each of which summarizes a generalization from this table."

Sentences. Wallach and Kogan (1965) asked children to write short stories using four words. For present purposes with adult Ss the integration of four words in a single sentence seemed preferable. "Write many sentences, each of which contains these four words."

happy expensive horse lake

Cartoon. A cartoon was presented in four squares, with the printing removed from the last square. The instructions were to "write as many different quotes for the last square as you can."

Although the solutions were all written in words, the materials presented included verbal, numerical, and pictorial materials in order to achieve some variety of content. Samples of materials are included in the appendix.

THE QUALITY-QUANTITY RELATION IN PRODUCTIVE THINKING

Previous research demonstrated that instructions to write many solutions lower the average quality of the solutions produced, relative to instructions to write one solution. The present experiment altered the instructions so as to explore the quantity-quality relation more thoroughly.

Method

Two problems which had yielded somewhat different results in previous research were selected in the interests of generality. One, called Sentences, required the subjects to construct sentences from four words: happy, expensive, horse, lake. The other, called Conclusions, required the subjects to examine a chart displaying expenditures for several social welfare programs over a period of several years and write conclusions or inferences from the data. Information about the criteria of a good solution was given for each problem: "A good sentence reads smoothly; the four words fit unobtrusively into the structure of the sentence." "A good conclusion would be a valid generalization which integrates the table as a whole."

In previous experiments the subjects were instructed to write many solutions; in the present experiment the subjects were instructed specifically to write one, two, four, or six solutions in seven minutes. The experiment was conducted in small groups of psychology students recruited and assigned to conditions randomly.

As in previous research on this project, the solutions were typed on cards, coded, and rated by two experienced raters on a scale from 1 to 7. The sums of these ratings, which range from 2 to 14, constitute the data for analysis.

Results*

Tables 1 and 2 show that, as in previous experiments, those who wrote more solutions wrote solutions of lower mean quality. The data for a few subjects who wrote only five though asked to write six are also included in the tables. The standard deviations were computed. Each subject was given a score representing the mean of the quality ratings of his solutions, and the means and standard deviations shown in the tables are based on these scores.

* Stuart Agres had major responsibility for collection and analysis of the data.

TABLE 1

QUALITY OF SENTENCES PRODUCED UNDER DIFFERENT QUANTITY INSTRUCTIONS

Quantity requested	Number of subjects	Mean quality	Standard deviation	Theoretical value
1	27	8.41	2.02	8.29
2	28	7.54	1.95	7.81
4	30	7.57	1.23	7.33
(5)	9	7.60	.91	7.14
6	31	6.97	1.03	7.05

TABLE 2

QUALITY OF CONCLUSIONS PRODUCED UNDER DIFFERENT QUANTITY INSTRUCTIONS

Quantity requested	Number of subjects	Mean quality	Standard deviation	Theoretical value
1	27	9.25	5.04	9.50
2	25	8.28	3.41	7.97
4	25	6.39	2.38	6.53
(5)	12	6.15	2.27	6.07
6	25	5.68	1.46	5.69

The assumption of a linear drop in quality as a function of quantity would not be inconsistent with the data at hand but would lead to unbelievable values at more extreme points. A logarithmic decline would be more probable, hence the general form of the equation could be $Y = K - C \log X$. The equation of this form that best fits the data for Sentences is $Y = 8.3 - 1.60 \log X$, and for Conclusions $Y = 9.4 - 4.78 \log X$. The theoretical values are printed in Tables 1 and 2, and the theoretical curves are drawn in Fig. 1.

The linear correlation between number of solutions written and mean quality scores was -.271 for Sentences and -.384 for Conclusions, similar to the correlations obtained earlier with slightly different instructions. The corresponding correlations between log number and quality were about the same: -.282 for Sentences and -.389 for Conclusions.

Discussion

The generalization that mean quality declines with an increase in quantity has now been demonstrated by three variations in method. Previous research has shown that when subjects with and without information about the criteria of good solutions are instructed to write many solutions, they write solutions of lower mean quality than those instructed to write one solution, and within the groups instructed to write many there is a negative correlation between quantity and quality. In the present experiment with criteria cues, the same decline is demonstrated when definite numbers of solutions are requested.

Several interpretations of the decline in quality are plausible. It is reasonable to assume that the subjects take instructions to write more than one solution as permission to lower standards, but instructions to write four are probably not twice as permissive as instructions to write two. The standard-lowering effect of the number requested would probably be curvilinear, as in the equation above.

Another interpretation, not incompatible with the above, takes account of the quantity and quality of the available solutions. The number of good sentences that can be constructed from the four given words is almost unlimited, hence the decline with quantity would be small. The number of good conclusions to the chart is distinctly limited, however, hence the curve falls sharply. The high mean quality of the single conclusions suggests that the one good solution is easily constructed and differentiated from others of lower quality. The shape of the frequency distribution of solutions as a function of quality could influence the quantity-quality relation.

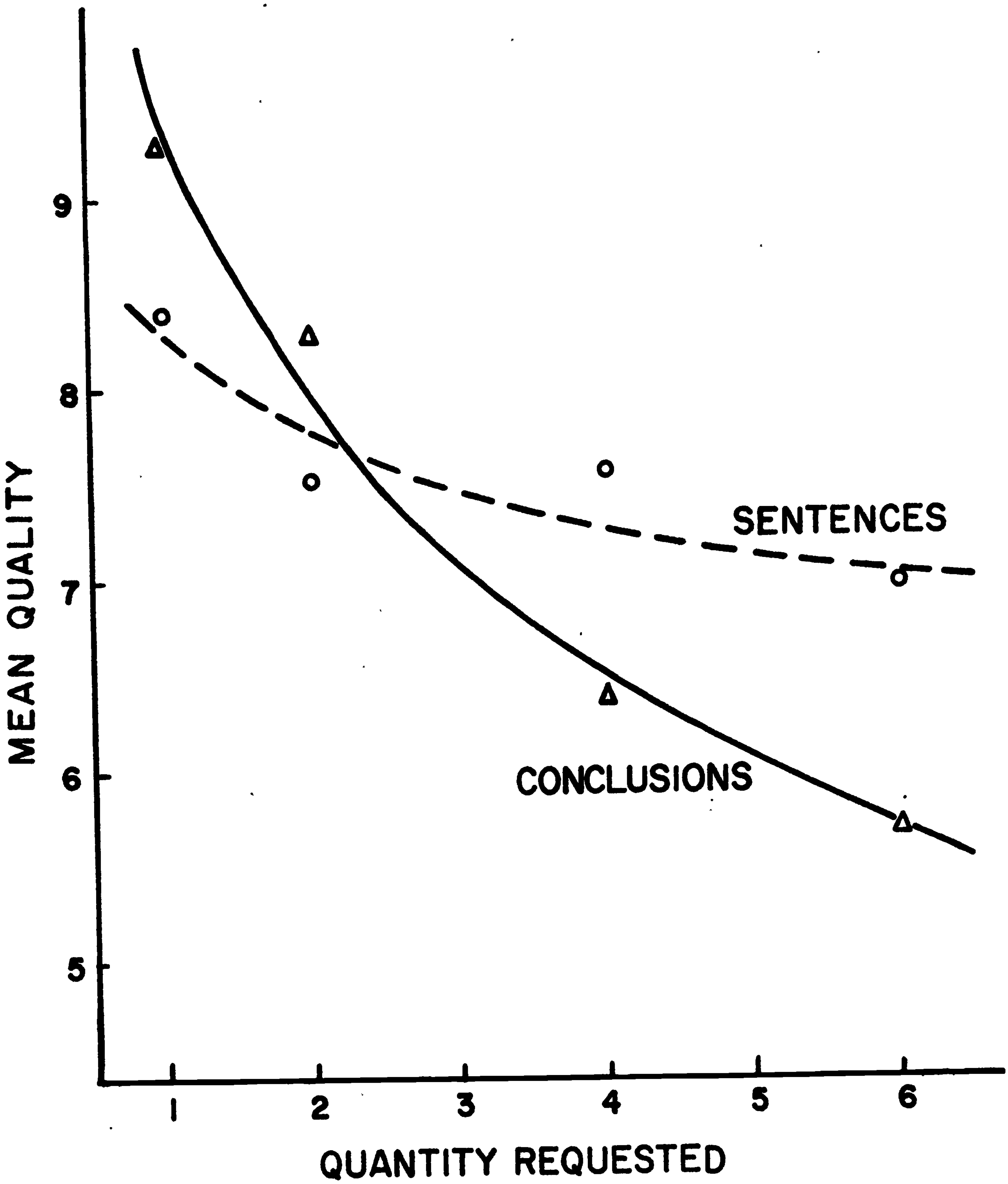


Fig. 1. The relation between quality and quantity for two problems.

EFFECTS OF RESTRICTED TIME AND INSTRUCTIONS ON PRODUCTION OF SINGLE SOLUTIONS TO FOUR PROBLEMS*

Much research effort has been directed toward the elucidation of individual differences in creativity (see MacKinnon, 1962). Progress in this direction, however, cannot advance very far without the elucidation of the processes underlining problem solving in general. Toward this end several studies (e.g., Johnson and Jennings, 1963; Johnson, Parrott, and Stratton, 1968) have analyzed the problem-solving effort into the three separable, but functionally interdependent, component processes of preparation, production and judgment.

Duncan (1967) proposes a more detailed model drawn from learning theory. In problems where there are several solutions possible, there is a habit strength connected to each solution. Duncan and others have demonstrated an active verbal response hierarchy based on associative habit strengths in anagram and verbal fluency problems. Gerlach, Schultz, Baker and Mazer (1964) used the same reasoning to explain the results found in brainstorming studies. In the unusual uses problem used by Gerlach, et. al., the subjects are required to produce many possible uses of a common object, such as a coat hangar. They hypothesized that a response hierarchy would exist with these problems such that the subjects would give the most dominant response first and subsequently work down the hierarchy. With the unusual uses problem, uncommon uses would have been reinforced less often previous to the experiment and would be lower on the response hierarchy; hence they would be given only after the more common responses. Their results demonstrated that uncommonness of responses was, in fact, linearly related to ordinal position in the production sequence. But, if instructions contained criterial information, uncommon responses occurred equally often in all ordinal positions. Parnes (1961) and Menske & Davis (1968) have also found the production-order effect with unusual uses problems and no criteria-cued instructions.

Tasks like writing uncommon associations or writing unusual uses are subject to the criticism that they require only a superficial verbal fluency. Problems may be constructed, however, which require the subject to absorb information and integrate it into an appropriate and clever solution. Tasks like the familiar plot-titles problem represent legitimate academic exercises which are complex as well as stimulating. If the extension of learning theory to problem solving is to be valid, it must be tested with problems which do relate strongly with intellectual activity.

*R. Paul Stratton had the major responsibility for this study.

Christensen, Guilford, and Wilson (1957) used the plot-title problem among others in their investigation of problem-solving processes. Whereas they found that the number of plot-title responses increased linearly with time, the quality of these solutions did not increase with time as it did for the verbal association and unusual uses problems. These results were explained as exemplifying the difference between "creative" and "recall" problems which can be accounted for by the directness of learned connections. Creative problems call upon connections which operate by way of a transfer of training, being more indirect than recall connections. Thus, the production-order effect would be expected when uncommonness and remoteness are the criteria of originality, but not when cleverness is the criterion since clever responses are just as closely related to a stimulus as nonclever responses.

The data from Gerlach, et. al., and Johnson, Parrott, and Stratton suggest another interpretation of these results. Gerlach, et. al., interpreted their results by asserting that criteria-cued instructions could disrupt the response hierarchy to allow uncommon responses to occur in any order. Johnson, Parrott, and Stratton used five problems similar to, and including, the plot-title problem. For each problem more superior solutions (above 90th percentile in quality) occurred in the second half of the production sequence, and over all problems almost twice as many superior solutions occurred in the second half. Importantly this was not the case in the three groups where there were instructions relating criterial or judgmental information.

The position of responses on the response hierarchy is theoretically determined by the amount of reinforcement received by the S-R contingency prior to the experiment. According to Gerlach, et. al., these criteria-cued instructions "bring to bear a habit-family hierarchy in line with the modified reinforcement contingencies." In word association problems (Duncan, 1966; Bremer, 1968) the interaction of instructions and the response hierarchy is described in terms of a search model. If instructions can guide the subject in his search for a solution through the response hierarchy, the order of production may remain intact, but the subject would be able to implicitly proceed through responses and only record better solutions which normally would appear later in the production sequence. Thus, the criteria-cued instructions would not disrupt the response hierarchy but would alter the recorded responses by eliminating those of lowest quality. This would also account for the reduction in quantity usually accompanying increases in quality with criteria-cued instructions.

The confusion that is apparent in the published reports on production order may be clarified when these factors are taken into consideration. Christensen, et. al., instructed subjects on the plot-title problem either to write "appropriate" solutions or to write "clever and appropriate" solutions. By the present analysis both instructions would guide the subject in his search and would allow him to proceed further down the response hierarchy reporting only the better solutions. The usual order of production would not occur in the reported solutions, even though it could occur implicitly. Thus, the response hierarchy notion is still a viable interpretation of the results obtained from complex and simple problems.

The concept of a response hierarchy may also be used to interpret another finding in the Johnson, Parrott, and Stratton study. Of five groups one was instructed to write only one solution and one was instructed to write as many solutions as possible. Otherwise their instructions were equivalent and did not contain any reference to quality or evaluation. The multiple-solution group, it was found, produced more superior solutions than the single-solution group but of lower average quality. The present interpretation of the problem-solving process may be applied to this data. Since both groups received equivalent instructions (with the exception of quantity instructions) and equivalent time to produce, the most parsimonious interpretation would indicate that both groups went through exactly the same processes. The subjects in the multiple-solution group recorded almost every solution they thought of while working through a response hierarchy and producing better and better solutions. The single-solution subjects, on the other hand, did not have to spend most of their time recording their solutions, but did spend the time covertly going over possible solutions or types of solutions. Their solutions would also get better through the production sequence, because they would progress through the response hierarchy as would the multiple-solution group. If they only advanced to the second half of the response hierarchy (i.e., covertly produced 3-4 solutions) and recorded the last solution produced, the single-solution group would produce better solutions on the average.

This interpretation would also predict what Johnson, Parrott, and Stratton called a "quantity effect." If there is no difference in the processes by which a single solution or many solutions are produced, they would produce essentially the same solutions, one overtly and one covertly. Thus, multiple-solution subjects produce more solutions at all quality levels, but they produce more inferior than superior solutions. Thus, multiple-solution subjects produce more superior solutions, but of lower average quality. Secondly, the distribution of solution quality by subjects indicates that the variability between solutions in one multiple-solution subject is the same as the variability between single-solution subjects. Thus, there is no difference in the quality of solutions or solution processes between one individual repeatedly writing solutions drawn from his own resources and several individuals writing one solution each from different resources.

The main purpose of the present experiment was to test the response hierarchy interpretation of the processes used in solving complex problems. If the subjects were instructed to record only one solution, the quality of that solution should vary with the number of solutions covertly produced. Thus, production time was restricted to .5, 1, 2, and 4 minutes, and it was hypothesized that quality would increase with increased time.

Other findings of the Johnson, Parrott and Stratton study are relevant here. When a multiple-solution group was given instructions containing criterial information, their solutions were better on the average than multiple-solution groups not so instructed but not significantly better than the solutions of the single-solution group. Secondly, when single-solution subjects were given the same criteria-cued instructions, there was an increment in solution quality, but it was not statistically significant. The later result would suggest that the solution quality reaches a peak under certain conditions. The later study, however, included only two problems and only the normal criteria-cued instructions. Single-solution quality could be so high that to go beyond would take more or better criteria-cued instructions.

The present experiment tested this possibility by varying the instructions of the single-solution group. It was hypothesized that criteria-cued instructions and additional specific information about criteria would increase the quality of single solutions. Secondly, single solution quality should be better than multiple-solution quality only when instructions are equivalent to or better than the multiple-solution group and only when the time given for a single-solution is equal to or greater than that given to write each solution in the multiple-solution group.

Including criteria-cued instructions for single solutions presents a second test of the response-hierarchy concept. Single solution quality should increase with time only when there are no instructions referring to quality or evaluation.

Method

Materials. Four problems were selected from those used in previous studies (see Appendix A). Sentences instructed the subjects

(Ss) to write a sentence containing the four words happy, expensive, horse, and lake. Conclusions instructed Ss to read a graph and write a conclusion drawn from that information. Cartoon Captions instructed Ss to write a caption for the last square of a four-square cartoon. Plot Titles instructed Ss to write a title for a story presented in a paragraph.

These problems were assembled in booklets in a rotating order such that each problem occurred in each position equally often. Each problem was assembled to control exposure times; the problem material (graph, four words, etc.), then the instructions.

Instructions. There were four instructional conditions. Group 1 Ss were instructed to write one solution for each problem in the allotted time. Group 2 Ss were instructed to write one good solution for each problem, a good solution being defined by criteria-cued instructions (See Appendix B). Group 3 Ss were given Group 2 instructions. In addition they were given the Rating Guide by which their solutions would be evaluated, and that relationship was explained. (See Appendix C). Group 4 Ss were instructed to write as many good solutions as possible in the allotted time and were given the criteria-cued instructions of Group 2. After each problem these Ss were told to reread their solutions and select the three best solutions, then the one best of the three.

There were four time conditions: .5, 1, 2, and 4 minutes. Each S worked all problems under the same instructional and time conditions, and was informed of the time limit. Passage of half the allowed time was announced. For the .5 minute time limit Ss were told to "write as fast as you can and write the first solution that comes to mind." For the 2 and 4 minute time limits, Ss were told to "take your time" and "do not write the first solution that comes to mind." If S was not done writing at the end of the production time, he was allowed to finish that solution. On the first problem of the sequence for .5 and 1 minute time conditions some Ss went over by 15 seconds but did do better on subsequent problems.

Procedure. Ss were gathered in small groups and assigned randomly to experimental conditions. They were then introduced to the task and were informed of the appropriate time limit. Problems were completed individually with the experimenter controlling the timing and sequencing of the exercise. The problem solving sequence was as follows: 60 seconds preparation time to read the problem material were allowed for all groups; instructions on what to do with this material were given verbally for Group 1 and repeated if necessary; Groups 2 and 4 were given 30 seconds to read and reread the instructions; Group 3 was given 90 seconds to do the same. The subjects then wrote their solution (s) in the time allotted. During the reading of the instructions the experimenter reminded Ss of what they should be doing if they appeared not to be reading.

Subjects. *S*s were 260 males and females who volunteered for the experiment as part of the introductory psychology course requirements. Males and females were roughly equated between groups, and there were 20 *S*s in each condition.

Results

Each solution was typed, coded, and judged according to the Rating Guide by two judges on a scale from 1 (low) to 7 (high). The criteria, in general, were cleverness, appropriateness, and accuracy. Interjudge agreement was above .80 for all problems.

The data for this experiment were the summed ratings (2- 14) of the single solutions produced by each subject in Groups 1 - 3 and the mean summed ratings of the multiple solutions produced by each subject in Group 4.

The means and standard deviations are given for each separate problem in Tables 1 - 4. Table 5 presents the same data based on the mean quality of all four problems per S. Note that Group 4 data are the mean summed ratings per S for all solutions.

To analyze the significance of the overall effects of instruction and time a 3 x 4 x 4 factorial design (Winer, 1962, p. 248 was used.) In the conclusions problem (Table 6) instructions showed a significant influence on solution quality ($p < .01$). Newman-Keuls analysis of individual means showed Group 2 to have better solutions than Group 3 both of which were better than Group 1 ($p < .01$). Also there is a significant three-way interaction with problem position which defies explanation.

Table 7 and Figure 2 present the analysis of the sentences problem. Both instructions and time show significant effects ($p < .01$). Newman-Keuls individual comparisons on instructions indicate Group 3 solutions to be better than Group 2 which both are better than Group 1 ($p < .01$). Varying the time allotted to think of and record a solution significantly affects solution quality ($p < .01$) such that .5 and 1 minute do not differ, but both are inferior to 2 and 4 minutes which differ from one another ($p < .01$).

An additional analysis is necessary with this problem, however, because sentence length is correlated to quality for Groups 1 - 3 (r = .35 - .49). Thus, time limits may only serve to restrict writing time, not any sort of processing time which would influence quality. Figure 1 presents mean sentence length per time and group. Notice how closely it corresponds to Figure 1 of sentence quality. The analysis of variance on sentence length in Table 8 shows the effects of time and instruction to be highly significant ($p < .001$). Individual comparisons show the same relationships as the first analysis. In addition .5 minute is less than 1 minute ($p < .05$).

TABLE 1. MEANS AND STANDARD DEVIATIONS OF RATINGS OF SOLUTIONS TO CONCLUSION PROBLEM BY INSTRUCTIONS AND TIMES

Groups	Time Limits in Minutes .5	1	2	4	Total
1. Standard instructions	6.10 3.94	5.10 4.11	9.05 4.08	7.00 4.77	6.81 4.48
2. Criteria cues	9.70 4.26	7.80 4.64	9.0 3.95	9.80 4.07	9.08 4.32
3. Criteria cues with Rating Guide	8.55 3.51	9.00 4.87	8.0 4.39	9.10 4.10	8.66 4.27
Totals for single-solution groups	8.12 4.19	7.30 4.84	8.68 4.17	8.63 4.49	8.18 4.67
4. Multiple-solution group with Criteria Cues				7.23 2.77	

TABLE 2. MEANS AND STANDARD DEVIATION OF RATINGS OF SOLUTIONS TO SENTENCE PROBLEM BY INSTRUCTIONS AND TIMES

Groups	Time Limits in Minutes .5	1	2	4	Total
1. Standard instructions	6.05 2.40	6.75 1.42	7.85 1.96	8.40 2.35	7.26 2.19
2. Criteria cues	8.95 2.01	7.40 2.31	8.25 2.16	9.70 2.15	8.58 2.32
3. Criteria cues with Rating Guide	7.80 2.87	8.70 3.46	9.40 2.87	10.05 3.17	8.99 3.22
Totals for single-solution groups	7.60 2.73	7.62 2.66	8.50 2.45	9.38 2.69	8.28 2.74
4. Multiple-solution group with Criteria Cues				7.37 2.10	

TABLE 3. MEANS AND STANDARD DEVIATION OF RATINGS OF SOLUTIONS TO PLOT TITLES PROBLEM BY INSTRUCTIONS AND TIMES

Groups	.5	1	2	4	Total
	Time Limits in Minutes				
1. Standard instructions	7.05 2.60	7.80 2.56	7.15 2.50	7.70 2.33	7.42 2.52
2. Criteria cues	6.50 2.71	7.50 2.48	7.00 2.39	7.70 2.49	7.18 2.56
3. Criteria cues with Rating Guide	6.50 2.11	6.80 1.86	7.50 2.11	6.10 2.17	6.96 2.14
Totals for single-solution groups	7.00 2.52	7.37 2.36	7.22 2.35	7.17 2.45	7.19 2.42
4. Multiple-solution group with Criteria Cues				6.36 0.97	

TABLE 4. MEANS AND STANDARD DEVIATION OF RATINGS OF SOLUTIONS TO CARTOON PROBLEM BY INSTRUCTIONS AND TIMES

Groups	.5	1	2	4	Total
	Time Limits in Minutes				
1. Standard instructions	6.95 2.13	6.50 1.99	7.30 2.90	6.00 2.93	6.69 2.57
2. Criteria cues	7.85 2.78	7.95 1.91	7.20 2.50	6.90 3.05	7.48 2.63
3. Criteria cues with Rating Guides	7.40 3.10	6.65 2.78	7.35 2.71	7.10 3.33	7.12 3.01
Totals for single-solution groups	7.40 2.73	7.03 2.35	7.28 2.71	6.67 3.14	7.10 2.76
4. Multiple-solution group with Criteria Cues				6.22 1.68	

TABLE 5. MEANS AND STANDARD DEVIATIONS OF RATINGS OF SOLUTIONS TO ALL PROBLEMS BY INSTRUCTIONS AND TIMES

Groups	Time Limits in Minutes .5	1	2	4	Total
1. Standard instructions	6.54 1.35	6.36 1.59	7.84 1.26	7.28 1.46	7.00 1.54
2. Criteria cues	8.25 1.55	7.66 1.68	7.86 1.52	8.52 1.82	8.08 1.68
3. Criteria cues with Rating Guide	7.80 1.51	7.79 2.08	8.06 1.40	8.09 2.03	7.93 1.79
Totals for single-solution groups	7.55 1.64	7.27 1.90	7.92 1.40	7.96 1.86	7.67 1.74
4. Multiple-solution group with Criteria Cues				6.77 1.16	

TABLE 6. ANALYSIS OF VARIANCE FOR CONCLUSIONS COMPARING ALL SINGLE-SOLUTION GROUPS

Source	df	SS	MS	F
Time (T)	3	74.23	24.74	1.37
Instruction (I)	2	232.31	116.16	6.44 **
Position (P)	3	37.10	12.37	n.s.
T x P	9	237.80	26.42	1.46
T x I	6	161.49	26.92	1.49
P x I	6	26.52	4.42	n.s.
T x P x I	18	551.68	30.65	1.70 *
Error	192	3464.80	18.04	
Total	239	4785.93		

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

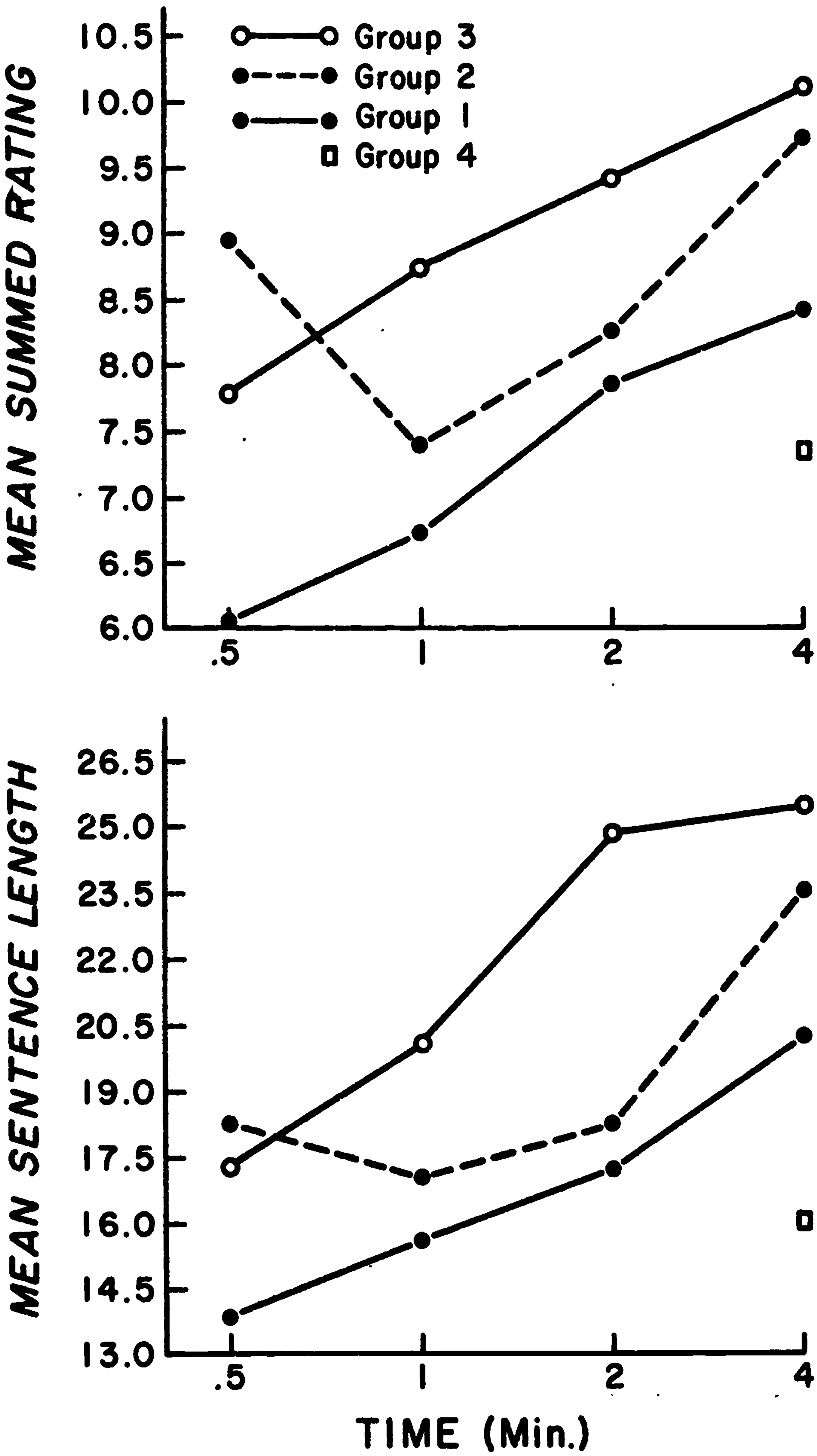


Fig. 1. Mean summed rating and sentence length as functions of time for each group. Sentence Problem.

TABLE 7. ANALYSIS OF VARIANCE FOR SENTENCES COMPARING ALL SINGLE-SOLUTION GROUPS

Source	df	SS	MS	F
Time (T)	3	130.08	43.36	5.33 **
Instruction (I)	2	129.83	64.92	7.98 **
Position (P)	3	22.55	7.52	n.s.
T x P	9	68.95	7.66	n.s.
T x I	6	51.04	8.51	n.s.
P x I	6	24.07	4.01	n.s.
T x P x I	18	120.13	6.67	n.s.
Error	192	1563.20	8.14	
Total	239	2109.85		

** $p < .01$

TABLE 8. ANALYSIS OF VARIANCE FOR SENTENCES COMPARING ALL SINGLE-SOLUTION GROUPS ON SENTENCE LENGTH

Source	df	SS	MS	F
Time (T)	3	1540.22	513.41	14.93 ***
Instruction (I)	2	1092.16	546.08	15.88 ***
T x I	6	322.47	53.74	1.56
Error	228	7841.15	34.39	
Total	239	10796.00		

*** $p < .001$

After the data were collected and analyzed, the solutions to the sentence problem were copied by hand by a naive volunteer who was instructed to write as fast as possible and to take frequent rest periods to prevent fatigue. The time to copy every solution was recorded. Under the .5 minute time condition an average of 29.45 seconds was required to simply write the solution, under the 4 minute condition 39.10 seconds. The average writing time per solution for Group 4 was 26.90 seconds. While writing time reflects the same conclusions as quality ratings, it further implies that Ss having only .5 minutes to think of and record a solution took most of that time to simply write the solution. Longer time conditions allowed Ss to do other things than write.

Table 9 presents the analysis for plot titles. For this problem the position in which the problem was encountered significantly influenced quality ($p < .05$) with the second ($\bar{X} = 7.63$) and fourth (7.63) positions being equal and more conducive to better solutions than the first (6.97: $p < .05$) or the third (6.52: $p < .01$).

Table 10 presents the analysis for cartoon captions. There are no statistically significant differences attributable to these main variables.

The analysis for all problems is presented in Table 11. Position effects could not be analyzed, because the mean for each S included each problem in each position. Hence, the effects of the position variable (if in fact there were any) were counterbalanced and averaged out. The effects due to time restrictions approach significance, but does not meet the .05 criterion ($F = 2.65$ necessary). A statistically significant influence was evidenced in the instructions variable ($p < .01$) over all problems. Instruction Groups 2 and 3 did not differ, but both differed from the standard Group 1 ($p < .01$). Apparently, the quality of a single solution can be increased substantially through instructions containing criterial information.

It was hypothesized that single solution quality should increase with increased time only when there are no instructions referring to quality or evaluation. Thus, the effects due to time were analyzed for each group separately within each problem and for all problems together. For conclusions (Table 12) only Group 1 gave a significant F value ($p < .05$), and 1 minute solutions were inferior to 2 minute solutions ($p < .05$) on a Newman-Keuls test.

For sentences (Table 13) time effectively increased quality in Group 1 ($p < .01$) and in Group 2 ($p < .05$). Within Group 1 individual comparisons showed 4 minute solutions to be superior to

TABLE 9. ANALYSIS OF VARIANCE FOR PLOT TITLES COMPARING ALL SINGLE-SOLUTION GROUPS

Source	df	SS	MS	F
Time (T)	3	4.11	1.37	n.s.
Instruction (I)	2	8.57	4.28	n.s.
Position (P)	3	53.78	17.93	3.29 *
T x P	9	37.87	4.21	n.s.
T x I	6	47.83	7.97	1.46
P x I	6	34.26	5.71	1.05
T x P x I	18	176.14	9.79	1.80
Error	192	1047.00	5.45	
Total	239	1409.56		

* $p < .05$

TABLE 10. ANALYSIS OF VARIANCE FOR CARTOON CAPTIONS COMPARING ALL SINGLE-SOLUTION GROUPS

Source	df	SS	MS	F
Time (T)	3	18.95	6.32	n.s.
Instruction (I)	2	24.91	12.46	1.63
Position (P)	3	18.58	6.19	n.s.
T x P	9	45.14	5.02	n.s.
T x I	6	22.59	3.76	n.s.
P x I	6	27.96	4.66	n.s.
T x P x I	18	204.27	11.35	1.48
Error	192	1468.40	7.65	
Total	239	1830.80		

TABLE 11. ANALYSIS OF VARIANCE FOR ALL PROBLEMS COMPARING ALL SINGLE-SOLUTION GROUPS

Source	df	SS	MS	F
Instructions	2	54.30	27.15	9.80 **
Time	3	19.66	6.55	2.37
Interaction	6	18.84	3.14	1.13
Error	228	631.57	2.77	
Total	239	724.37		

** $p < .01$

TABLE 12. ANALYSIS OF VARIANCE FOR CONCLUSIONS FOR GROUPS 1 - 3 COMPARING TIME CONDITIONS

Group	Source	df	SS	MS	F
1	Time	3	169.64	56.55	2.99*
	Error	76	1436.55	18.90	
	Total	79	1606.19		
2	Time	3	50.95	16.98	<1.0
	Error	76	1439.60	18.94	
	Total	79	1490.55		
3	Time	3	15.25	5.08	<1.0
	Error	76	1442.75	18.98	
	Total	79	1458.00		

* $p < .05$

TABLE 13. ANALYSIS OF VARIANCE FOR SENTENCES FOR GROUPS 1 - 3 COMPARING TIME CONDITIONS

Group	Source	df	SS	MS	F
1	Time	3	67.44	22.48	5.41 **
	Error	76	316.05	4.16	
	Total	79	383.49		
2	Time	3	57.85	19.28	3.92 *
	Error	76	373.70	4.92	
	Total	79	431.55		
3	Time	3	55.84	18.61	1.83
	Error	76	771.15	10.15	
	Total	79	826.99		

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

.5 minutes ($p<.01$), 1 minute ($p<.05$), and 2 minutes ($p<.05$). Within Group 2 only 1 minute was less than 4 minutes ($p<.01$).

For plot titles and cartoon captions there were no significant differences.

Over all problems (Table 14) time was an effective variable only in Group 1 ($p<.01$) where 2-minute solution quality was better than 1 minute ($p<.05$) and .5 minute ($p<.05$), and was equal to 4-minute quality. Apparently when criterial instructions are not present time is an effective determinant of single solution quality.

The effects of restricted time and criteria-cued instructions may also be evaluated by comparing single-solution groups under comparable conditions to multiple-solution Group 4 which had criteria-cued instructions and four minute production time. These effects may be evaluated separately.

Comparing single and multiple solution groups with four minutes the effect of quantity instructions and criteria-cued instructions may be seen. For the comparison the multiple solution condition (Group 4) served as a control against which each single solution condition with four minutes was compared.

A one-way analysis of variance across the four minute conditions of Groups 1 - 4 (Table 15) gave significant F values for the sentences ($p<.01$) and plot titles ($p<.05$) problems and for all problems together ($p<.01$). Individual comparisons with a two-tailed Dunnett's t-statistics and Newman-Keuls test showed Group 2 and 3 solutions to be significantly better than those of Group 4 at least at the .05 level of confidence on joint and individual comparisons for the sentence and plot title problems and for all problems together. Groups 1 and 4 did not differ.

Comparisons on hypothesized differences gave significant two-tailed t-values indicating the average solution quality of Group 2 with 4 minutes was superior to Group 4 for conclusions ($p<.05$), sentences ($p<.01$), plot titles ($p<.05$), and for all problems together ($p<.01$). Average solution quality of Group 3 with 4 minutes was superior to Group 4 for conclusions ($p<.05$), sentences ($p<.01$), and for all problems together ($p<.05$). Also for plot titles Group 1 with 4 minutes had better solutions than Group 4 ($p<.05$), but for all problems together the difference was not significant. Apparently under equivalent conditions writing a single solution is better than writing many solutions, and having more than the standard criteria-cued instructions does not increase this effect. For the most part multiple solutions with criteria-cued instructions are about as good as single solutions without them.

Comparing single and multiple solution groups with criteria-cued instructions the effect of quantity instructions and restricted

TABLE 14. ANALYSIS OF VARIANCE FOR ALL PROBLEMS FOR GROUPS 1 - 3 COMPARING TIME CONDITIONS

Group	Source	df	SS	MS	F
1	Time	3	27.95	9.32	4.39 **
	Error	76	161.24	2.12	
	Total	79	189.19		
2	Time	3	8.97	2.99	1.05
	Error	76	216.45	2.85	
	Total	79	225.42		
3	Time	3	1.59	0.53	<1.0
	Error	76	253.88	3.34	
	Total	79	255.47		

** $p < .01$

TABLE 15. ANALYSIS OF VARIANCE FOR THREE PROBLEMS COMPARING SINGLE-AND-MULTIPLE-SOLUTION GROUPS WITH FOUR MINUTES

		Conclusions		Sentences		Plot Titles		All Problems	
Source	df	MS	F	MS	F	MS	F	MS	F
Instruction	3	38.19	2.27	30.28	4.89**	14.63	3.22*	12.52	4.37**
Error	76	16.80		6.19		4.54		2.87	

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

time may be seen. For this comparison the multiple-solution condition served as a control against which single-solution conditions with criteria-cued instructions were compared.

The one-way analysis of variance--similar to that summarized in Table 15--gave significant values for the sentence problem ($F = 4.37$; $p < .01$) and for all problems together ($F = 3.55$; $p < .01$). Newman-Keuls individual comparisons for sentences showed the single solutions with four minutes produced solutions superior to the multiple-solution group ($p < .01$) and to the single-solution group with one minute ($p < .01$). For all problems together the multiple-solution group quality was inferior to the single-solution group with four minutes ($p < .01$) and with .5 minute ($p < .05$). Dunnett's t-statistic showed the same relationships except that in sentences 4 minutes was not better than 1 minute.

Comparisons with two-tailed t-tests were made on hypothesized differences. Single-solution quality should be better than multiple-solution quality when instructional conditions are the same and when single-solution time is equal to or longer than the time taken for the average single solution written by the multiple-solution group. In four minutes an average of 3.40 to 5.30 solutions were written by the multiple-solution group, giving 52 to 71 seconds per solution on the average. Thus, the 1 minute condition was equivalent to the multiple-solution condition for each solution. The hypothesis was confirmed for 4 minutes of Group 2 in conclusions ($p < .05$), sentences ($p < .01$), plot titles ($p < .05$), and for all problems ($p < .01$); for 2 minutes in conclusions ($p < .05$), and for all problems ($p < .01$); for 1 minute in plot titles ($p < .05$), cartoon captions ($p < .01$) and for all problems ($p < .05$); and for .5 minute in conclusions ($p < .05$), sentences ($p < .01$), cartoon captions ($p < .05$) and all problems ($p < .01$). Furthermore, the average quality of single solutions under all time conditions was better than multiple solutions in conclusions ($p < .05$), sentences ($p < .05$), cartoon captions ($p < .05$) and all problems ($p < .01$). Apparently time limitations even up to one-half that of the average multiple solution do not significantly reduce the quality of a single-solution quality when compared to multiple-solution quality.

Discussion

Using a variety of problems several researchers have shown that solution quality increases as S progresses through the production sequence. Response-hierarchy theory predicts such an improvement and stipulates the conditions under which such an improvement occurs. The order of production of several solutions, then, is determined by habit strength which has been found to vary with solution frequency, recency, and instructional set.

A response-hierarchy theory accounts for differences in solution quality between Ss producing one or many solutions by assuming that both groups go through the same processes--that they both produce several solutions in an order determined by habit strength. Thus, if single-solution Ss were to record the first solution they produced, quality would be lower than if they had produced several solutions and recorded the last one produced.

This interpretation was tested by varying the time allotted to solve each problem when only one solution was required. Presumably more time would allow S to produce several solutions even when required to record only one. Within Group 1 (no criteria-cued instructions) more time was found to be conducive to better solutions in the conclusions and sentence problems, and, since other problems showed no contradictions, the increase in quality was significant across all problems.

Simple writing time was recorded for the sentences problem after the experiment. In the 30-second time interval it appears that most of the time is spent writing. On the other hand, Ss with four minutes only used 40 seconds, on the average, to record their solutions. Writing time may approximate the minimum time necessary to formulate and record a solution. Sentences appropriate to this problem may be formulated and recorded in less than 30 seconds, the shortest time limit. Thus, although there must be a minimum time necessary to formulate a sentence, restricting time longer than that minimum must restrict the functioning of other processes. Response-hierarchy theory suggests that at least one of the other processes must be the evocation of alternative solutions to the problem.

For the plot title and cartoon captions more solutions were produced by the multiple solution Ss than for the sentences and conclusions problems. In addition, plot title and cartoon caption solutions are seldom longer than a few words. This would suggest that several alternative solutions could be considered in a short period of time, and, unlike the other problems, 30 seconds may still have been long enough for S to produce more than one solution. Additional time, then, would have had little beneficial effect.

Other problem differences point to the generality of these findings. Whereas the sentence problem may be considered to draw upon verbal skills, the conclusions problem primarily draws upon quantitative skills. Likewise, the sentence problem is more divergent, having an almost infinite number of good solutions. The conclusions problem has a limited number of good solutions and may be considered more convergent. The characteristic which differentiates these problems from the plot-title and cartoon-caption problems is that the solutions to the former are longer and require more time to write.

Criteria-cued instructions have been shown to greatly increase quality in several studies. An increase in quality due to quantity instructions was demonstrated by Johnson, Parrott, and Stratton. Producing a single solution on the average is better than producing several solutions when other conditions are equivalent. Providing criteria-cued instructions to multiple-solutions Ss increases their average solution quality, but it is still not better than single-solution quality.

By providing single-solutions Ss with criteria-cued instructions the present experiment demonstrated that single-solution quality could also be increased. Thus, although solutions produced singly are good, there is room for improvement, and the single-solution condition without criteria-cued instructions cannot be regarded as representing a ceiling of quality due to problem or judgmental idiosyncrasies. Whether quality may be increased further in single or multiple-solution conditions by a more rigorous training in judgment is subject for further research.

Single-solution quality may be directly compared to multiple-solution quality under several instructional and time conditions. Under the same time conditions writing one solution is better than writing many only when instructions are equivalent. When instructional conditions are equivalent, single-solutuon quality is best under all time restrictions. Thus, under equivalent instructions it is still better to produce one solution than to produce many.

Response-hierarchy theory also led to the prediction that time would be an effective determinant of quality only when there were no instructions referring to evaluation or quality. Results indicated that this was, in fact, the case. Obtaining similar results Gerlach, et al., interpreted the instructions to "bring to bear a habit-hierarchy in line with the modified reinforcement contingencies." Now a more detailed interpretation is possible. Such a disruption of the response hierarchy would occur if criterial information allowed S to search through a series of responses and select the one best meeting the criteria.

Thus, only the evaluation phase would be affected. Were this the case in the present experiment, a search would still be necessary and restricting time would restrict search time, thus reducing quality. If, on the other hand, criterial information guided S in the construction of each separate solution, the quality of the first solution produced would be as good as (or better than) the last solution encountered in an un uided search. This would predict that restricted time would not affect solution quality under criteria-cued instructions.

At this point the aforementioned difference in the nature of the problems becomes critical to the utility of the response-hierarchy concept as an interpretation of the present results. Response-hierarchy theory was formulated and tested on data from anagrams, unusual uses, association and other verbal problems. Solutions for simple verbal problems, like anagrams, have been previously encountered by S and stored in memory. Problem solution, then, principally involves memory scanning, and the order of solution recall--the order of scanning possible solutions--may be determined by such variables as word frequency (Duncan, 1967). On the other hand, Ss have never encountered solutions to the plot-title problem, even though they may have previously worked through similar situations. Thus, solution search does not depend on order of the recall of solutions, but upon the order of construction of solutions, and each solution must be constructed from scratch.

This does not, however, preclude the operation of a response hierarchy with these problems (Berlyne, 1965, p. 316), but it does necessitate some redefinition. A response hierarchy is usually derived from the products of thought and often based on the frequency of occurrence of solutions. In the unusual uses problem creative solutions are the ones which do not occur frequently in a large population of solutions. Although such a frequency may be determined for each solution type in more complex problems, frequency itself could not be expected to influence Ss' solution construction since Ss could not be aware of such a frequency--or "commonness" criterion--for solution evaluation. The inspection of solutions at various frequency levels could, however, give some insight into the characteristics of solutions which Ss change as they proceed through the production sequence.

Although a solution frequency derived from a large solution population could not be expected to influence Ss' order of production if he has not experienced the solutions previously, S must be considered to be capable of producing any solution in that population, and every solution must have some finite probability of occurrence. It would, then, be efficacious to regard every solution as being present within every S and having some probability of occurrence. If it is assumed that the probability associated with the solution population and the individual subject have some communality, the determinant of that probability in each would be the same and would be observable by looking at the solution population.

Other studies have shown that in terms of qualitative judgments "poorer" solutions usually occur first--if there are no criteria-cued instructions given. In unusual uses problems these are the more common uses in terms of normal frequency of usage (e.g., "bricks are to build with" vs "to use as boat anchors.") In the plot title problem the qualitative evaluation is more difficult. The present study used the criteria of "appropriateness and cleverness". The more a plot title describes the whole plot, the more appropriate it is. Cleverness, however, is more difficult to define. Mednick (1962) and others regard creativity as being a dimension based on the associative remoteness of the combined elements. Referring to the previous unusual uses problem, "boat anchor" is seen as a more remote associate to "brick use" than "to build with." Thus, a more creative solution is one which differs most from what you would expect in the problem situation. (Of course, the extremely unusual solution must still be appropriate to the problem). Creative responses, then, are ones which diverge from the original situation in some measurable way, such as associative strength.

Plot titles can also be measured on this divergence dimension. The plot is about a man who has lived a fast life and was domesticated by his wife after they moved to a new town. An old flame arrives, threatens exposure, and is poisoned by the competent wife. Thus, the title "The Gilson Story," while being very appropriate, is less clever by being more strongly related to the story than "Let No One Witness Yesterday."

A response hierarchy according to this analysis could be determined by those aspects of the problem situation which inhibit the tendency to produce divergent solutions; that is, a tendency to produce poor solutions first could be determined by the stimulus control inherent in the situation. Since S has never worked the problem before, reading through the problem could build an associative strength between the problem elements. The process of divergence, then, would be the breaking up of these associations and the forming of new ones. This suggests that the stronger the associations between problem elements, the more difficult it would be for S to produce solutions which are divergent from the problems. Thus, at least initially, solutions would be of lower quality and would be more closely related to the problem.

To account for the general increase in quality various mechanisms are possible: (1) a limited number of solutions which are very similar to the problem, (2) a dissatisfaction with these dull and uninteresting solutions, (3) a successive modification of existing solutions, and (4) a reinforcement of relevant criteria by recorded solutions.

Although response-hierarchy theory is not alone in attacking the processes of problem solving, it does fit the data from the present experiment. With the present extension of this theory the solution processes associated with a wide range of problems and instructions may be accounted for in terms of learning and the existing problem situation. Consequently the variables influencing the problem solving processes are accessible to empirical scrutiny.

FACILITATION OF PRODUCTION AND JUDGMENT OF SOLUTIONS TO A VERBAL PROBLEM BY JUDGMENT TRAINING *

The need for education to teach the young how to think, as opposed to teaching them what to think, has long been recognized. Problem solving represents a most important form of thinking, and an analysis of problem solving into discrete stages would show points of strength or weakness in particular efforts at problem solving.

Problem solving, in general, represents a complex interaction between the resources of an individual and the myriad aspects of a problem situation. The sequence of events from perception of the existence of the problem to final solution has been analyzed into the process stages of preparation, production and judgment (Johnson and Jennings, 1963). During preparation the individual confronts the problem situation which includes the relevant problem information and instructions. Most problems have many possible solutions, often varying in quality, and judgment follows the production of each solution, be it only partially complete, correct or incorrect. In problems with many possible solutions the interplay between these process stages may be investigated.

It seems reasonable that one way to improve the quality of solutions would be to inform the person how the solutions would be evaluated and to allow each person to construct solutions meeting these criteria. The proponents of brainstorming claim that this is not the case, and that significant gains may be made if production occurs with no regard to quality; by emphasizing quantity rather than quality more solutions are produced and more of them are of a superior quality (Meadow and Parnes, 1959). After a 30-week course in brainstorming, these researchers found an increase in the number of good ideas corresponding to an increase in the total number of ideas, and that this training showed some differential effect a year later (Parnes and Meadow, 1960).

Demonstrating that there are many ways to facilitate production, Gerlach, Schutz, Baker, and Mazer (1964) found instructions emphasizing quality and evaluation as effective as brainstorming instructions in producing more unusual uses. Similarly, other researchers have found that criteria-cued instructions facilitate production quality (e.g., Christensen, Guilford, and Wilson, 1957).

Whereas it is possible to increase the number of good ideas by increasing quantity without regard to quality, it is also possible to increase quality by emphasizing the evaluation of solutions. Johnson, Parrott and Stratton (1969) found that an emphasis on quality increased the average quality and decreased the number of solutions over instructions emphasizing quantity alone. Thus, an increase in quantity would produce more solutions at all quality levels, but a lower average quality. If overall quality is important, not just the number of superior ideas, the problem solver must at least be aware of the criteria of solution evaluation.

*R. Paul Stratton carried the major responsibility for this study. The assistanceof Bill Gould and John Jerome is greatly appreciated.

It would seem, then, that an increased awareness of the criteria of solution evaluation would increase average solution quality even more. In general, this has not been found to be the case. Johnson, et al. found that making Ss evaluate their own solutions after production to each of a series of problems produced no increase in the quality of subsequent solutions. Allowing Ss to practice evaluating 25 solutions, or to read these evaluations was found to be ineffective in changing solution quality or judgment accuracy on a transfer task (Johnson and Zerbolio, 1964).

A major limitation of studies attempting to demonstrate a production benefit from judgment practice has been the lack of adequate judgment practice. A primary criterion for an adequate judgment practice is the actual facilitation of judgment ability (which presumably would be transferred to production). By using an extended training procedure Johnson, et al. were able to increase the judgment accuracy of Ss on a series of problems. Ss produced several solutions to a problem, trained, then judged their own solutions. Each S worked on one problem. Over all types of problems judgment accuracy was found to increase with judgment training, and the greatest increment came when Ss trained in pairs.

The present study proposes to test the hypothesis that training in judgment will increase production quality, as well as judgment accuracy. By using the paired training program of Johnson, et al. this hypothesis will be tested with single and multiple production modes in hopes of finding the most productive combination.

Method

Subjects

During the Spring and Summer terms, 1968, 133 Ss were run in the multiple solution experiment. Incomplete test booklets were rejected and other Ss randomly eliminated to make three groups of 40 Ss each. Half of each group came from introductory psychology and education classes, and males and females were distributed without bias. During the Fall term 128 volunteer introductory psychology students were run in the single solution experiment. Three groups of 40 Ss were constructed as above, and males and females were equated across groups. All Ss were enrolled at Michigan State University.

Materials

Two sentence problems were used. Sentences 1 asked Ss to use the words happy, expensive, horse and lake in a sentence. Sentences 2 used the words alone, big, noise and money. (See Appendix A). Criteria-cued instructions which emphasized and defined "good" solutions were used for each. (See Appendix B).

The judgment training program (Appendix D) used in previous experiments applied to Sentences 1 and was used in both experiments. In general the program consisted of a detailed explanation of the evaluation procedure, examples of solutions at all levels of quality, and practice in judging solutions. The effectiveness of this training was evaluated in the multiple solution experiment by having Ss select one solution as their "best". For the single solution experiment a 12-item multiple-choice judgment test was constructed for each problem. (See Appendix E). The task here was to select the "best" solution of five for each item. Test items were of varying difficulty levels.

Procedure

Both experiments used a transfer-of-training paradigm controlling for practice in production and judgment to the first problem. Production for the multiple solution experiment (Exp. M) consisted of writing as many good solutions as possible in 7 minutes. Production for the single solution experiment (Exp. S) consisted of writing only one good solution in 7 minutes. Judgment in Exp. M allowed Ss to return to their own solutions and pick three, then one, "best" solutions. As much time as necessary was allowed for this. Judgment in Exp. S consisted of taking a standard multiple-choice judgment test in 10 minutes. All groups in the separate experiments had the same production and judgment tasks.

Procedures for both experiments were identical--with the above exception. Group C, a control group, produced and judged to Sentences 2 only. Group P produced and judged to both sentence problems, but without feedback. Group T produced and judged to both problems, but had judgment training after production, and before judgment, to the first problem. The training program enlisted the intellectual and motivational stimulation of Ss practicing in pairs.

Results

The solutions were typed, coded and rated independently by two expert judges. Interjudge agreement for Sentences 1 was .80 and .75 for Exps. M and S. For Sentences 2 it was .97 and .82. Since the solutions in each experiment were rated at different times and different subject samples were used, an equal number of solutions from Exp. M were randomly selected with the stimulation of some being at each quality level. To insure compatibility of the rating scales these were rated a second time with the solutions from Exp. S. The respective distributions were almost identical, and the separate ratings were highly correlated (Table 1). Notice also the similarity of the distributions for the two sentence problems.

TABLE 1. MEANS AND STANDARD DEVIATIONS FOR THE FIRST AND SECOND EVALUATION OF SOLUTIONS TO EXP. M.

	First Rating	Second Rating	Correlation
Sentence Problem 1.	7.89	7.66	.83
	2.26	2.18	
Sentence Problem 2.	7.89	7.57	.93
	3.41	3.20	

TABLE 2. MEANS AND STANDARD DEVIATIONS FOR NUMBER AND QUALITY OF SOLUTIONS PRODUCED BY EACH PROBLEM BY GROUPS OF EXPERIMENT M.

Group	Number Produced		Mean Quality	
	Sentence 1	Sentence 2	Sentence 1	Sentence 2
Training	7.12	4.90	7.30	8.45
	2.15	1.64	1.40	1.50
Practice	6.80	6.10	7.55	7.62
	2.18	1.74	1.35	1.56
Control		6.32		7.35
		2.20		1.09

Multiple solution experiment

Productivity and judgment training. The main emphasis of this study is the influence of judgment training on the production of solutions to a second problem. Figure 1 and Table 2 give productivity measures in terms of the number of solutions produced and their quality level. Although Groups P and T did not differ in productivity on Sentences 1, after the training of Group T, they did differ on Sentences 2.

Table 3 presents the analyses of variance with repeated measures on sentences which compares Groups P and T on the quality and number of solutions produced in Sentences 1 and 2. There was a significant decrease in the number of solutions produced between problems, but the decrease was greater for the trained group ($p < .01$). Similarly quality increased across problems, and that increase was greater for the trained group ($p < .05$).

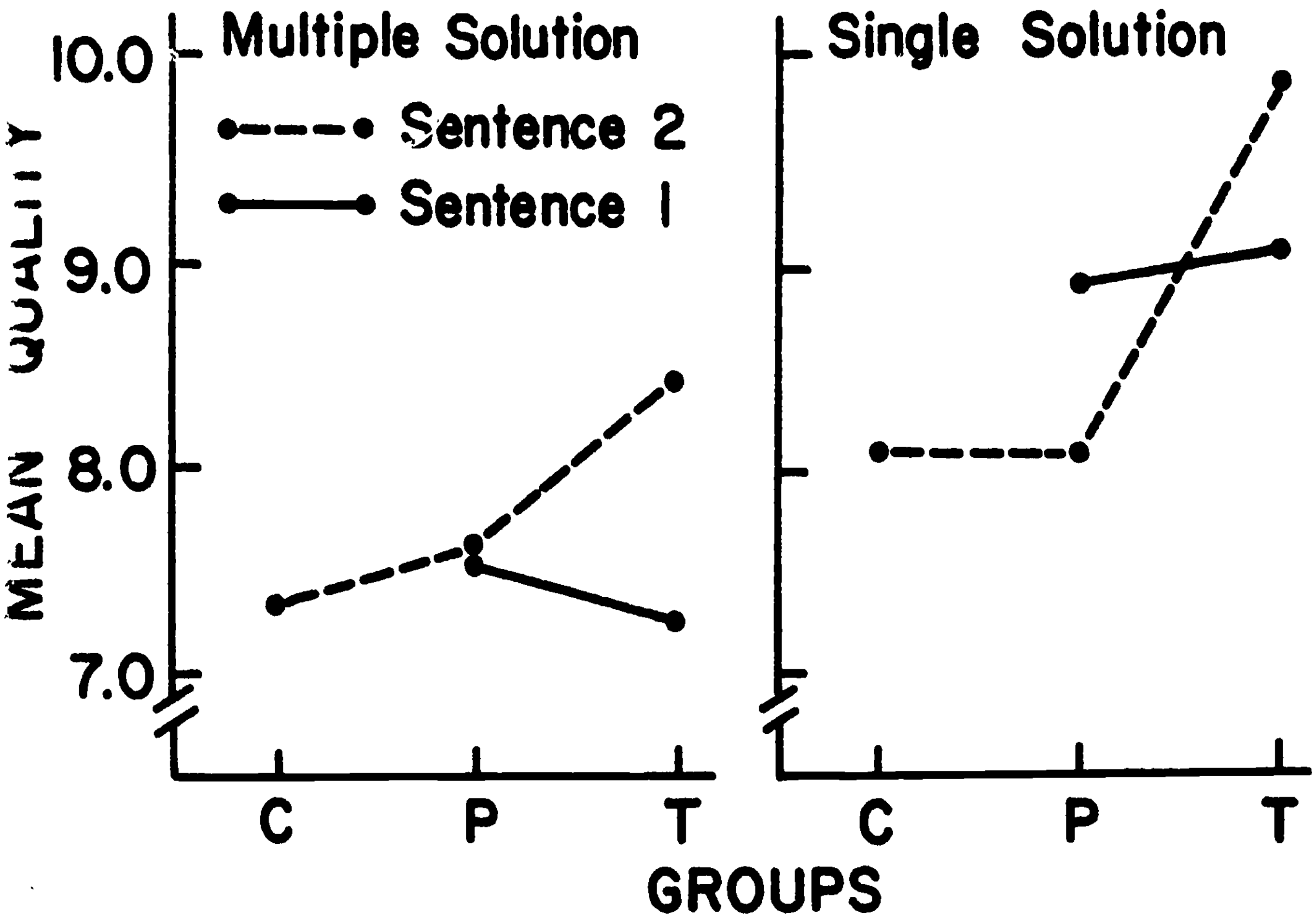
Table 4 presents the analysis of variance for production of all groups on Sentences 2 only. The number of solutions produced decreased across all groups with the control and practice groups being equal, and both producing more solutions than the trained group ($p < .01$) by Dunnett's t-statistic and Newman-Keul's comparisons. The same analyses were used on mean quality. Across groups quality increased, with the trained group producing better solutions than the practice or control groups ($p < .01$).

Table 5 presents data on the occurrence of superior solutions (above 90th percentile in quality) on Sentences 2. Although the small numbers preclude statistical comparisons, it is evident that in the trained group more people produced more superior solutions than in the other groups.

Judgment ability and judgment training. It is evident that judgment training did facilitate quality of production on all measures. But did the judgment training also increase Ss' ability to evaluate their own solutions? Table 6 and Figure 2 present the basic data on judgment performance on both sentence problems. A preferred solution is the one solution picked by S as his final "best", and the remainder are the nonpreferred solutions. Because the quality of the preferred and nonpreferred solutions is a function of the quality of the sample from which they are drawn, a difference score was calculated for each S by subtracting the average quality of the nonpreferred solutions from the quality of the preferred solution.

On the first sentence problem Ss produced their solutions, then judged them. The Ss of Group T were trained in judgment between the production and judgment of the first problem. Judgment ability on the first problem was evaluated by two-tailed t-tests

Fig. 1. Mean quality as a function of treatment on Sentences 1 and 2.



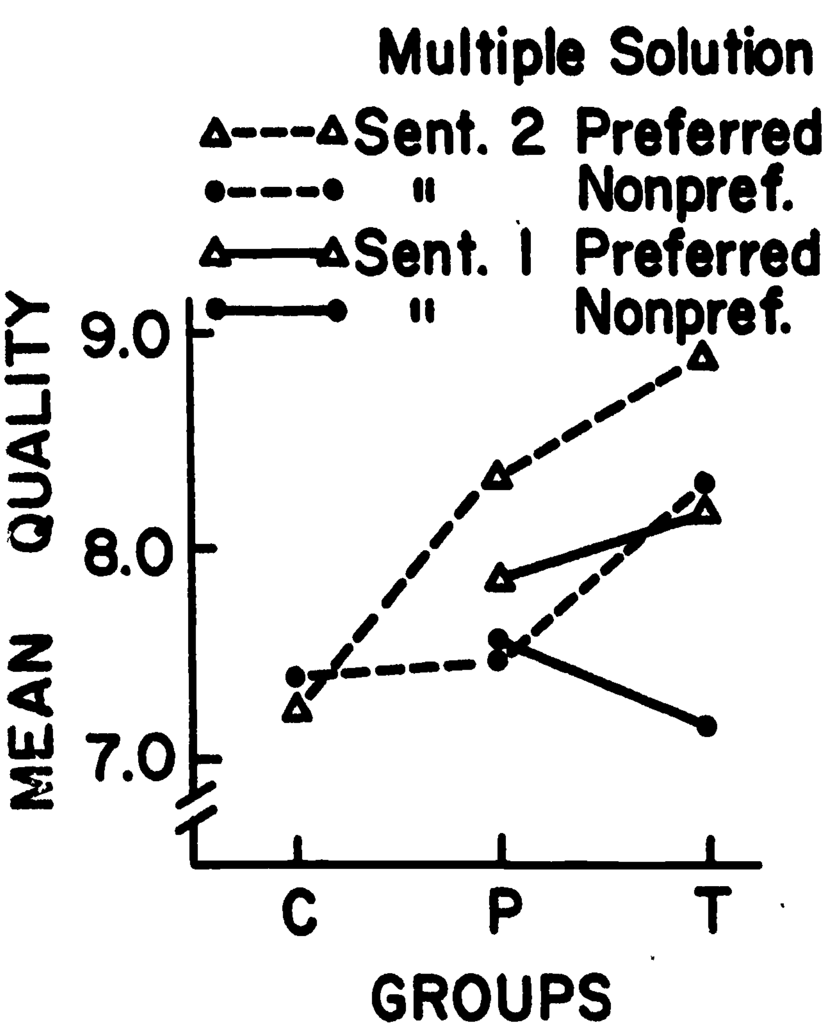


Fig. 2. Mean quality of preferred and nonpreferred solutions as a function of treatment.

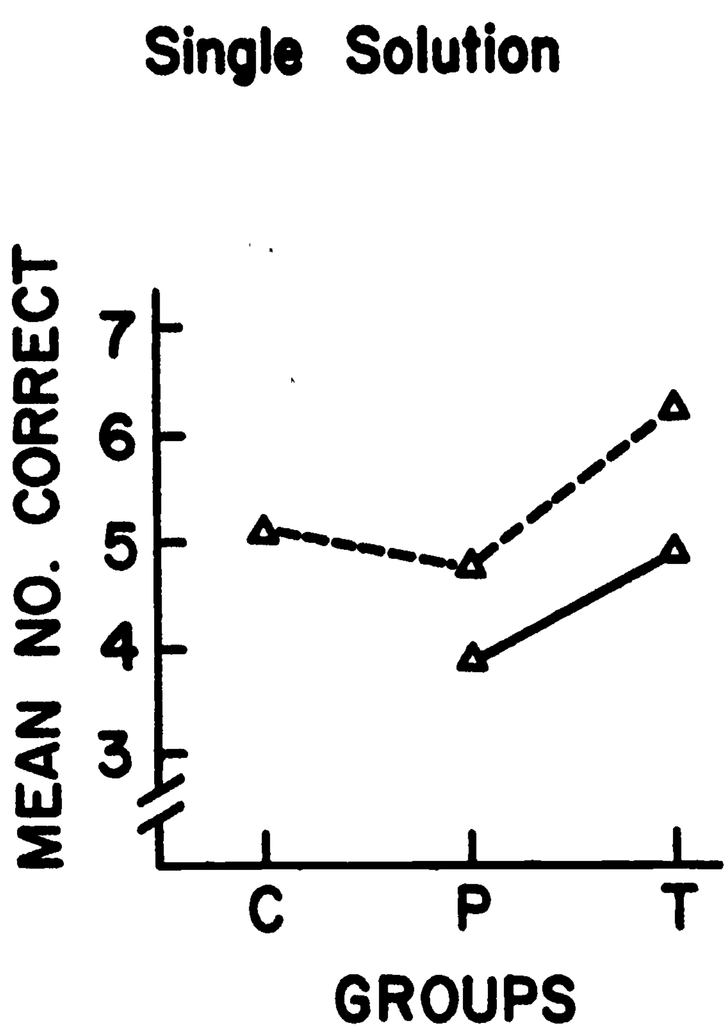


Fig. 3. Mean number correct on judgment tests as a function of treatment.

TABLE 3. ANALYSIS OF VARIANCE OF NUMBER AND QUALITY OF SOLUTIONS PRODUCED BY GROUPS T AND P ACROSS SENTENCES 1 AND 2 IN EXPERIMENT M.

		Number Produced		Mean Quality	
Source	df	MS	F	MS	F
Training (T)	1	7.66	1.38	3.33	1.31
Ss x T	78	5.54		2.53	
Sentences	1	85.56	38.87**	14.69	8.16**
T x S	1	23.26	10.57**	11.86	6.58**
Ss x T x S	78	2.20		1.80	

** $p < .01$.

TABLE 4. ANALYSIS OF VARIANCE OF NUMBER OF SOLUTIONS PRODUCED AND MEAN QUALITY ON SENTENCE 2 FOR ALL GROUPS IN EXPERIMENT M.

		Number Produced		Mean Quality	
Source	df	MS	F	MS	F
Groups	2	23.48	6.51**	13.19	6.59**
Within	117	3.61		2.00	

** $p < .01$.

TABLE 5. NUMBER OF SUPERIOR SOLUTIONS, PERCENT OF SOLUTIONS WHICH WERE SUPERIOR, AND NUMBER OF Ss HAVING ONE OR MORE SUPERIOR SOLUTIONS IN SENTENCE PROBLEM 2 IN EACH GROUP IN EXPERIMENTS M AND S.

	Group			
	Training	Practice	Control	Total
Multiple				
Number Superior	13	11	8	32
Percent Superior	11.4	9.1	6.4	8.9
Ss with one superior solution	10	8	7	25
Single				
Number Superior	18	5	4	27
Percent Superior	45	12.5	10	22.5
Ss with one superior solution	18	5	4	27

TABLE 6. MEANS AND STANDARD DEVIATIONS FOR QUALITY OF PREFERRED AND NONPREFERRED SOLUTIONS AND FOR DIFFERENCE SCORES OF SOLUTIONS PRODUCED TO EACH PROBLEM BY GROUPS OF EXPERIMENT M.

	Preferred Solutions		Nonpreferred Solutions		Difference Scores	
Group	Sent. 1	Sent. 2	Sent. 1	Sent. 2	Sent. 1	Sent. 1
Training	8.15	8.80	7.15	8.25	1.00	.55
	2.39	3.07	1.32	1.47	1.67	2.85
Practice	7.85	8.32	7.52	7.45	.33	.87
	2.84	2.94	1.22	1.49	2.47	2.58
Control		7.22		7.36		-.14
		2.48		1.23		2.76

TABLE 7. ANALYSIS OF VARIANCE OF MEAN PREFERRED, MEAN NONPREFERRED AND MEAN DIFFERENCE SCORES ON SENTENCE 2 FOR ALL GROUPS IN EXPERIMENT M.

		Preferred Solutions		Nonpreferred Solutions		Difference Scores	
Source	df	MS	F	MS	F	MS	F
Groups	2	26.11	3.16*	19.28	9.64**	10.64	1.39
Within	117	8.27		2.01		7.64	

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

for correlated scores. The trained group selected their best solutions at a better-than-chance level ($p < .025$); the practice group did not.

The transfer of judgment ability to the second problem was tested by analyses of variance, the results of which are given in Table 7. As one would expect from the production facilitation, differences between the groups occur on comparisons of preferred quality and of nonpreferred quality. On these comparisons the solutions of the trained group were better than both other groups (preferred, $p < .025$; nonpreferred, $p < .01$). But comparisons of difference scores showed no differences between groups. Apparently judgment training facilitated judgment on the first problem, but the increased judgment ability did not transfer to the second problem.

Single solution experiment

Productivity and judgment training. Every S produced only one solution for each problem in this experiment. The quality of these solutions is given in Figure 1 and Table 8. Groups T and P did not differ in performance on Sentences 1, but on Sentences 2 the trained group increased in quality and the practice group actually decreased in quality ($p < .01$) as shown by the analysis in Table 9. Comparing all groups on Sentences 2 the trained group produced better solutions than either untrained group ($p < .01$) by Dunnett's *t*-statistic and Newman-Keul's tests. (See Table 10).

In Table 5 it can be seen that the trained group also produced more than three times as many superior solutions as the untrained groups on Sentences 2.

Note in Table 1 that the distribution of the first and second ratings of Exp. M solutions were almost identical. This indicates that the solution quality in the two experiments may be compared with appropriate caution. Figure 1 presents the average quality of solutions produced in each experiment. Note that in all comparisons Ss producing only one solution produced better solutions than Ss producing several. Table 11 presents a more precise picture. Using the average quality as a standard of comparison, the single solution quality is higher ($p < .01$), but using the one preferred solution per subject as a comparison there is no difference in quality. Because of the caution with which this comparison is made, it cannot be said that the latter comparison approaches significance. In Table 5 the total number of superior solutions is larger for Ss in Exp. M, but the difference may be accounted for by some Ss in Exp. M producing more than one superior solution. More single solution Ss produced superior solutions, and a higher percentage of their solutions

TABLE 8. MEANS AND STANDARD DEVIATIONS, FOR SENTENCE 1 AND 2, OF QUALITY OF PRODUCED SOLUTION, AND JUDGMENT TEST SCORED FOR NUMBER CORRECT AND QUALITY OF SELECTED SOLUTION IN EXPERIMENT S.

		Group		
		Training	Practice	Control
Quality of Produced Solution				
	Sent. 1	9.08	8.95	
		2.73	1.63	
	Sent. 2	9.95	8.08	8.10
		2.48	2.61	2.44
Number Correct on Judgment Test				
	Sent. 1	4.80	3.80	
		1.38	1.78	
	Sent. 2	6.18	4.77	5.15
		1.79	1.59	1.64
Quality of Selected Solution on Judgment Test				
	Sent. 1	8.56	8.01	
		.56	.80	
	Sent. 2	8.97	8.40	8.60
		.69	.85	.76

TABLE 9. ANALYSIS OF VARIANCE OF SENTENCE 1 AND 2 QUALITY OF PRODUCED SOLUTION AND JUDGMENT TEST SCORES OF NUMBER CORRECT AND QUALITY OF SELECTED SOLUTION IN EXPERIMENT S.

		Production		Judgment Test			
		Quality		Number Correct		Quality	
Source	df	MS	F	MS	F	MS	F
Training (T)	1	39.01	6.31*	57.60	17.26**	12.79	19.01**
Ss x T	78	6.18		3.34		.67	
Sentences	1	.01	<1.0	55.22	25.17**	6.42	15.52**
T x S	1	29.76	5.23*	1.60	< 1.0	.01	<1.0
Ss x T x S	78	5.69		2.19		.41	

* p <.05 ** p <.01

TABLE 10. ANALYSIS OF VARIANCE OF SENTENCE 2 QUALITY OF PRODUCED SOLUTIONS AND JUDGMENT TEST SCORES OF NUMBER CORRECT AND QUALITY OF SELECTED SOLUTION IN EXPERIMENT S.

		Production		Judgment Test			
		Quality		Number Correct		Quality	
Source	df	MS	F	MS	F	MS	F
Group	2	46.26	7.16**	21.01	7.31**	3.43	5.74*
Within	117	6.46		2.87		.60	

* p <.05 ** p <.01

TABLE 11. ANALYSIS OF VARIANCE COMPARING TREATMENT CONDITIONS AND PRODUCTION MODES

		Total M vs S quality		Preferred M vs S quality	
Source	df	MS	F	MS	F
Experiment (E)	1	48.80	11.51**	21.20	2.88
Treatment (T)	2	53.60	12.64**	61.40	8.33**
E x T	2	5.80	1.37	11.00	1.49
Error	234	4.24		7.37	

** $p < .01$

were superior. In addition single solution Ss were more influenced by the training in that the differential effect of training on quality was much larger between the groups of Exp. S than in Exp. M.

Judgment ability and judgment training. The ability to evaluate solutions was ascertained through a 12-item multiple-choice judgment test where Ss had to select one of five solutions as the "best". Success was measured by the number of correct choices and the quality of the solutions selected even if it was not correct. Figure 3 and Table 8 present the basic data for these measures.

A repeated-measures analysis of variance compared Groups P and T on both problems. Table 9 indicates that both groups improved their judgment performance on the second test, but that the trained group were more accurate in their judgments on both problems ($p < .01$). This is true for both judgment measures.

Table 10 summarizes the analysis of variance comparing the judgment ability of all groups on the second problem. Again the trained group made more correct judgments than the untrained groups ($p < .01$). And the solutions selected as "best" by the trained groups were better than the untrained groups' ($p < .05$). Independent comparisons were confirmed by Dunnett's t-statistic and Newman-Keuls comparisons.

Since judgment ability was assessed independent of productivity, direct comparisons are possible between groups. It is apparent, then, that trained single solution Ss do improve their judgment ability through training and that improvement transfers to the second test. This was verified by two measures of judgment ability.

Discussion

It was hypothesized that an increased awareness of the criteria of solution evaluation would improve the accuracy of judgment, and would improve the quality of production on a transfer task. The data indicate that this was, in fact, the case. In Experiment M production quality on the second problem increased only in the group receiving judgment training, and judgment accuracy improved only for the trained groups on the first problem. In Experiment S judgment training produced the only gains in solution quality and judgment accuracy on both problems.

Previously, Johnson and associates (1964, 1969) have shown that solution quality may be increased by criteria-cued instructions and by instructions which ask Ss to write one solution rather than many. Stratton, Parrot, and Johnson (1969) extended this finding by using criteria-cued instruction for both single and multiple solution Ss. Criteria-cued instructions increased single solution quality, and compared to average multiple solution quality single solutions were better, even when produced in a fraction of the time. In the present experiments criteria-cued instructions were used for all groups, and it was found that judgment training increased production quality over whatever benefits were derived from criteria-cued instructions or practice on the first problem.

Looking for the optimum combination, the single and multiple production modes may now be compared. On average quality per S the single solution instructions are better, but no advantage is evident if the comparison includes preferred solutions produced under multiple solution instruction. In terms of the number of superior solutions the single solution instructions are better. If each trained multiple solution S selected his best solution as preferred, there would be almost twice as many superior solutions produced by trained single solution Ss. In terms of time and efficiency it appears that judgment training with single solution instructions represents the optimum problem-solving conditions. These allow the individual to concentrate his efforts on constructing one solution which meets the criteria of solution evaluation. In all probability it will be better than solutions produced under any other condition investigated in these studies.

It is evident that judgment training was effective in increasing the production and judgment ability in both single and multiple production modes. Now it is incumbent upon us to present some explanation as to just what did transfer from the judgment training to the production of solutions in another problem.

Specifically the judgment training consisted of (1) a specific statement of the criteria of solution evaluation, (2) examples of superior and inferior solutions with explanations of their placement, (3) practice in selecting the best solution with immediate feedback, and (4) an opportunity to state in their own words the criteria which differentiated between superior and inferior solutions. In addition Ss practiced in pairs.

The familiarity with the production and judgment tasks and with the specific examples given in training could transfer to production in the second task. If, for example, Ss were to remember only one unusual use, e.g., "Happy Lake," it could be applied in the second problem for an increase in quality, e.g., "Money Lake." Few instances of such direct transfer were noted. Furthermore, such superficial learning could not account for the increment in judgment ability.

If Ss learned the training materials perfectly, they would now be aware of the specific criteria by which their solutions were to be evaluated, and they would learn that their solutions would, in fact, be evaluated--which is implicit, but not explicit, in criteria-cued instructions. This knowledge would be applied and reinforced during training itself. Increases in judgment accuracy following training may be interpreted as the direct application of this criterial information. This suggests that Ss construct solutions which meet these criteria or that Ss produce more solutions than they record, discarding obviously inferior ones which normally would have been recorded.

Levy (1968) makes an important contribution to the analysis of training in productive thinking. He interprets originality as role-taking, or behaving in accordance with perceived situational demand characteristics. More uncommon responses were obtained on a word association test and an inkblot test from Ss who were given a role model and reinforced for behaving accordingly. Performance under these conditions was better than just being given a role model or reinforcement, and all conditions were better than a control. Thus, the judgment training of the present experiment may be interpreted as providing a definition of what was expected from Ss and reinforcement for the implementation of that role. The Ss would, then, have taken on a critical role, learned the criteria of solution evaluation, been reinforced for good judgments and finally produced solutions meeting a high criterion of excellence. Furthermore, such transfer would have been evidenced in an increased judgment ability. This nonspecific form of transfer would, then, be expected to transfer to very dissimilar problems, unlike the transfer of specific criterial information.

CONCLUSIONS AND RECOMMENDATIONS

This research with substantial problems has confirmed what previous research with trivial problems has shown, that an emphasis on quantity of production of solutions to problems increases quality--if quality is measured by the number of superior solutions produced. If quality is measured in terms of mean quality of all solutions, the quantity emphasis decreases quality. Detailed examination of all solutions produced (many thousands) demonstrates that most individuals can produce solutions of a wide range of quality; the quantity emphasis leads to an increase in number of solutions of all quality levels, including a few superior solutions as well as many inferior solutions. Special instructions to defer criticism or to ignore quality are not necessary.

When a person produces several solutions and has no criteria information for guidance, the production process is somewhat less controlled at first; quality improves with subsequent productions. Apparently, poor solutions are produced first, and, if only one solution is written down, the poor solutions are discarded immediately. Criteria information serves to improve the quality of the first solution produced, even if there are more solutions to come. The concepts of response hierarchy and stimulus (problem) control were helpful in interpreting the results.

Information about the criteria to be used in judging solutions, given before production, improves quality of solutions. This information is apparently effective at both ends of the quality distribution, cutting down the number of very poor solutions and increasing the number of superior solutions.

When college students write many solutions to a problem, they can select their best solutions moderately well, but practice in judgment of solutions written by others improves accuracy in selecting one's own, and some of the improvement transfers to a second problem. Judgment training also helps in the production of solutions to similar problems. In fact, solutions produced after such training are better than under any other instructional aid.

Hence improvement in problem solving can come from improving the production process and from improving the judgment process. In this project the best overall performance came from those students who were instructed to write one solution but were given the benefit of the criteria of good solutions and of training in the judgment of solutions.

The increase in understanding of problem-solving processes gained by this project justified the recommendation that similar projects be conducted with high-school and grade-school students.

Separate study and manipulation of different problem-solving processes is helpful, perhaps even necessary, for further progress. Emphasis on the production of solutions to problems should be complemented by an equal emphasis on evaluation of the solutions. Evaluation of solutions by teacher or researcher is not as realistic as evaluation by students.

These results suggest that teachers should encourage separate practice of production and judgment functions, that manipulation of instructions and working conditions may be as good as a longer training period, and that specific statement of the characteristics of a good solution will usually be helpful.

References

Berlyne, D. E. Structure and Direction in Thinking. New York: Wiley, 1965.

Bremer, B. A. Production and judgment processes in a word problem. Unpublished Ph.D. Dissertation, Michigan State University, 1968.

Christensen, P. R., Guilford, J. P. & Wilson, R. C. Relations of creative responses to working time and instructions, Journal of Experimental Psychology, 1957, 53, 82-88.

Duncan, C. P. Effect of word frequency on thinking of a word. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 434-440.

Duncan, C. P. Response hierarchies in problem solving. Presidential address to Midwestern Psychological Association, Chicago, Illinois. May, 1966. In Duncan, C. P. (Ed.) Thinking: current experimental studies. New York: Lippincott, 1967, pp. 18-40.

Gerlach, V. S., Schutz, R. E., Baker, R. L., & Mazer, G. E. Effects of variations in test directions on originality test responses. Journal of Educational Psychology, 1964, 55, 79-83.

Johnson, D. M. & Jennings, J. W. Serial analysis of three problem solving processes. Journal of Psychology, 1963, 56, 43-52.

Johnson, D. M., Parrott, G. L. & Stratton, R. P. Production and judgment of solutions to five problems. Journal of Educational Psychology, Monograph Supplement, 1968, in press.

Johnson, D. M., & Zerbolio, D. J. Relations between production and judgment of plot-titles. American Journal of Psychology, 1964, 77, 99-105.

Levy, L. H. Originality as role-defined behavior. Journal of Personality and Social Psychology, 1968, 9, 72-78.

Manske, M. E., & Davis, G. A. Effects of simple instructional biases upon performance in the unusual uses test. Journal of General Psychology, 1968, 79, 25-33.

MacKinnon, D. W. The nature and nurture of creative talent. American Psychologist, 1962, 17, 484-495.

Meadow, A. & Parnes, S. J. Evaluation of training in creative problem-solving. Journal of Applied Psychology, 1959, 43, 189-194.

Mednick, S. A. The associative basis of the creative process. Psychological Review, 1962, 69, 220-232.

Parnes, S. J. Effects of extended effort in creative problem solving. Journal of Educational Psychology, 1961, 52, 117-122.

Parnes, S. J. & Meadow, A. Evaluation of persistence of effects produced by a creative problem-solving course. Psychological Reports, 1960, 7, 357-361.

Stratton, R. P., Parrott, G. L., and Johnson, D. M. The effects of restricted time and instructions on the production of single solutions to four problems.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

APPENDIXES

Appendix A.

PROBLEMS

The following four words have been used in a problem. You will have one minute to familiarize yourself with these words.

MONEY

NOISE

BIG

ALONE

____

____

____

PROBLEMS

The following four words have been used in a problem. You will have one minute to familiarize yourself with these words.

BIG

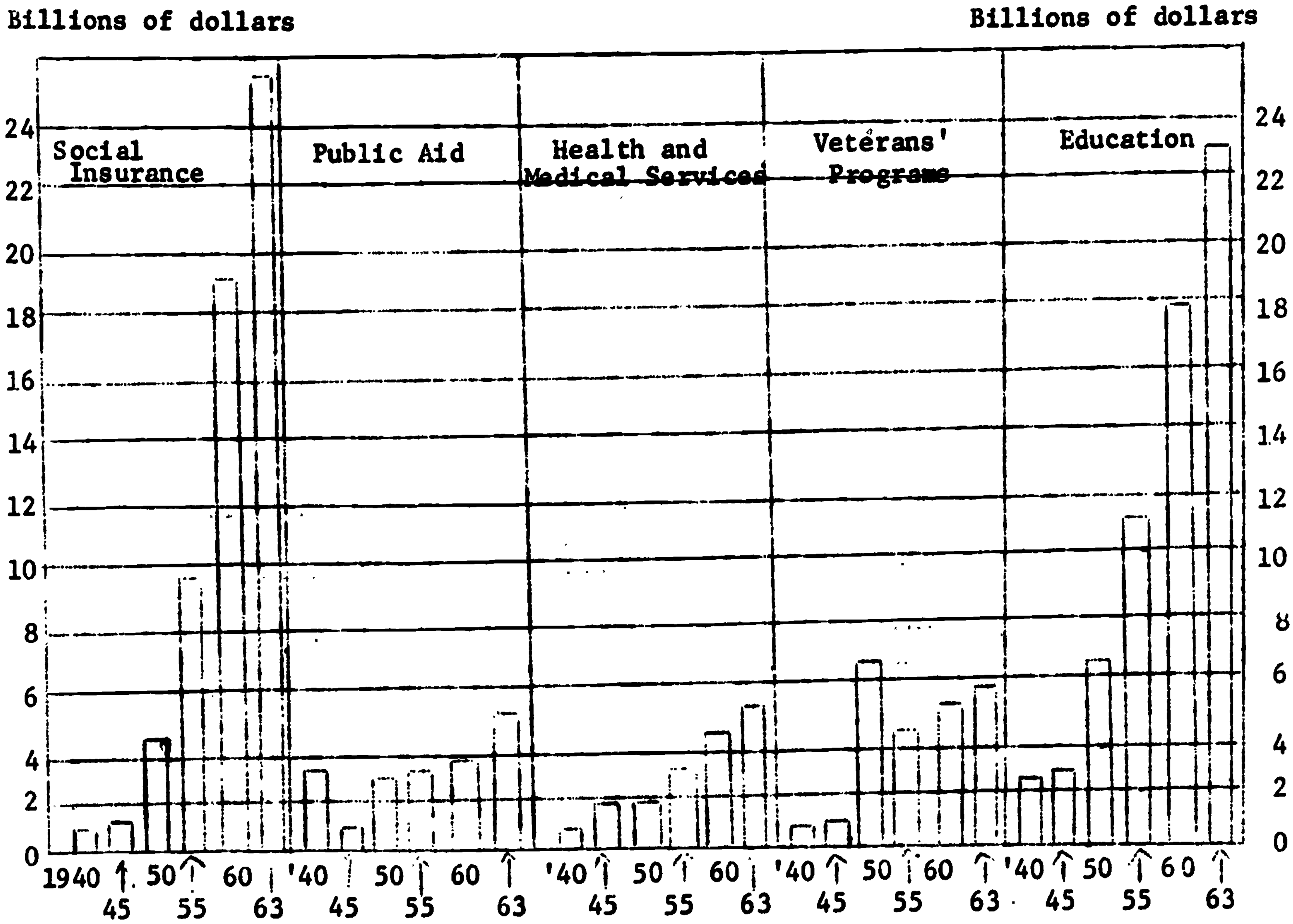
HORSE

LAKE

HAPPY

The following chart is taken from The Statistical Abstracts of the United States. What can you conclude from this table? Write as many good conclusions as you can on the lines below. A good conclusion would be a generalization which integrates all the information given in the table. It should be a valid conclusion which is drawn correctly from the given information and only from that information.

Social Welfare Expenditures Under Selected Public Programs: 1940 to 1963



STOP - Wait for further instructions

K3

Below is the plot for a novel, play or movie. Your task is to write as many clever titles for it as you can. By clever we mean an imaginative, creative, or unusual title for this plot. The title, however, must also be appropriate to the entire plot and the characters.

> Before the Gilsons moved to the little Connecticut town of Woodbridge, Stanley Gilson had lived a fast life, but his quiet, competent wife, Kay, had gradually toned him down and had achieved a degree of respectability for their family. When one of Stanley's old flames appeared in the Gilson house one afternoon and threatened to expose him, Kay quietly poisoned her and saved the family reputation.

1. ____________________
2. ____________________
3. ____________________
4. ____________________
5. ____________________
6. ____________________
7. ____________________
8. ____________________
9. ____________________
10. ____________________
11. ____________________
12. ____________________
13. ____________________
14. ____________________

STOP -- wait for further instructions.

Appendix B.

CRITERIA-CUED INSTRUCTIONS

I. Sentences:

"A good sentence reads smoothly, and the four words fit unobtrusively into the structure of the sentence. A good sentence must include the given four words."

II. Conclusions:

"A good conclusion would be a generalization which integrates all the information given in the table. It should be a valid conclusion which is drawn from the given information and only from that information; i.e., it is not an opinion."

III. Plot Title:

"A good title is one which is clever and appropriate. By clever we mean an imaginative, creative or unusual title for this plot. By appropriate we mean that it is true to the entire plot and characters."

IV. Cartoon:

"A good quote for this cartoon is one which is clever and appropriate. By clever we mean an imaginative idea. By appropriate we mean that it fits what Snoopy has said in the other squares and it fits Snoopy's posture in the last square."

Appendix C.

RATING GUIDES

A paragraph of general instructions is given to all subjects, which informs them of their next task, which will be to write one good sentence (conclusion, plot-title, cartoon caption) using the information provided in the problem.

Next, the evaluation procedure is explained:

"Your sentence (conclusion, plot-title, cartoon caption) will be evaluated on a seven-point scale. A good response would rate a "5", "6", or a "7". The scoring procedure is outlined below.

A statement of criteria-cued instructions follows for each problem, then the rating guides.

I. Sentences:

Refer to Appendix D.

II. Conclusions:

These are the basic considerations when evaluating conclusions. 1. A conclusion must be valid. It must be drawn correctly from the information given. 2. It must use all of the information from the chart. The more information used, the higher the score.

For a score of:	The conclusion must be:
1.	invalid (see above) or not a conclusion; e.g., an opinion.
2.	based on one or two programs on the chart over a one year period.
3.	based on one program for all year periods or all programs for one year period.
4.	based on two programs for all year periods or all programs for two year periods.
5.	based on 3 or 4 programs for all year periods or all programs for 3 to 5 year periods.
6.	a valid conclusion integrating all programs over all years, but not worded well.
7.	a "6" worded well; i.e., worded concisely and explicitly stated without ambiguity as to the meaning of the conclusion.

II. Conclusions (cont.)

Note: Conclusions receive a lower rating if they are ambiguous as to what is meant about which time period or programs are being considered.

III. Plot titles:

When a title is being evaluated the basic considerations are that it be a title and not, for instance, a sentence. Then cleverness and appropriateness are the relevant dimensions. By clever we mean imaginative, creative, or unusual. By appropriate we mean that it is true to the entire plot and to the characters.

For a score of:	The title must be:
1.	inappropriate or not a title.
2 to 3	simply the name of people or places.
2 to 4	statements of facts.
3 to 5	"How to..." titles. Using the idea of the old "flame" cliche's.
4.	the "murder" theme.
4 to 6	judgments, conclusions or rationalizations.

Note: Any of the above can be a "7" if it is humorous, is touched with irony or is otherwise outstanding; but it still must be a title which is appropriate to the plot and characters.

IV. Cartoon problem:

When a quote is being evaluated the basic considerations are that it be clever and appropriate. By clever we mean imaginative. By appropriate we mean it fits what Snoopy has said in the other squares and it fits his posture in the last square.

For a score of:	The quote must be:
1.	an inappropriate or incomprehensible quote.
2.	a trivial or common quote.
3 to 4	a comment about a "dogs life" or mention of a common excuse for not biting.
5 to 7	clever and/or unusual rationalization for not taking action--a philosophical comment on the management of inner tensions for the general social good.

Appendix D.

SENTENCE TRAINING

This task is to use four words in a single sentence.

horse lake expensive big

Write as many good sentences as you can, each of which includes all four words. A good sentence reads smoothly; the four words fit unobtrusively into the structure of the sentence.

1. ____________________

____

2. ____________________

____

3. ____________________

____

4. ____________________

____

5. ____________________

____

6. ____________________

____

7. ____________________

____

8. ____________________

____

9. ____________________

____

STOP - Wait for further instructions.

STOP -- WAIT FOR FURTHER INSTRUCTIONS!

Experiments in Production and Judgment

Introduction: As it has been pointed out, much previous research has been conducted with the problem you have just completed. As part of this earlier research, an accurate "Rating Guide" has been developed to grade the sentences.

In this phase of the experiment, you will practice using the Rating Guide in judging the sentences that others have written, and later this practice should help you in evaluating your own sentences.

Rating Guide: (Basic considerations--use of all the words, flowing smoothly and fitting unobtrusively into a sentence which must be reasonable, i.e., be believable.)

For a rating of:	The sentence must have these characteristics:
1......	does not use all four words
2......	lists the words; e.g., happy, expensive horse
3-5......	uses all words in sentence which flows well, but words stand out; a mediocre sentence
4-7......	words used in a different context; e.g., Horse Lake
6-7......	unusually good sentence without different context usage, but is well constructed, reads well, and is clever

- Additional considerations:

Complex sentences rated above compound sentences (e.g.,

complex--although expensive, the horse...;
compound-the horse ran, and it was expensive)

Rating increased if happy and expensive modify words other than horse or lake.

When you have finished examining the Rating Guide, go on to Practice A on the next page.

K4-b

Practice A

To illustrate how the Rating Guide has been used to judge sentences, we will present here some poor, some good, and some superior sentences selected from previous research.

The following are examples of poor solutions; we show you the poor ones first so that you can learn to eliminate them.

1. The expensive trip was not quite what they expected.

   Explanation: This is clearly a poor solution because not all the words have been used -- happy has been omitted.

2. Happiness is a serene lake and an expensive horse.

   Explanation: This is also a poor solution because the given words have not been used. Happy has been changed to "happiness" and this is just not acceptable.

3. The happy, expensive horse jumped into the lake.

   Explanation: This is a poor solution because it is a simple listing of the adjectives and they modify the given noun "horse". It does not take much imagination to produce this sentence, and it really doesn't read very smoothly. The words "stick out like a sore thumb".

4. The expensive horse made the boy near the lake happy.

   Explanation: This is fairly good, but "lake" is used as the subject and expensive modifies horse--this makes these words obvious. The given words should be better integrated into the sentence, so that you have to look twice to make sure they were all included.

* * *

The following are examples of good sentences, but they are not quite of a superior level.

1. A happy scene it was, the children splashing in the lake, a horse basking in the sun; their time not being expensive.

   Explanation: This is a good sentence, because it uses all the words unobtrusively and uses the adjectives other than to modify the given nouns. It falls short of being superior in that it is so complex it does not read smoothly.

2. Horse Lake is a happy resort for people with expensive tastes.

   Explanation: "Horse Lake" is a novel way of using "horse" and "lake". The sentence reads well, with the exception that a "happy resort" is rather unrealistic. This makes it good, but not superior.

3. I'm happy that our expensive weekend at Horse Lake is over with.

Explanation: The novel use of "Horse Lake", and the fact that it is so smooth reading and well-constructed, makes this a good sentence. A superior solution, however, would have a little something more. This one lacks imagination or humor.

4. When his expensive car broke down, John was happy to find a horse which he could ride to the lake.

Explanation: The idea is here for a superior sentence, but something is missing. It lacks the humor punch that this idea could carry. Otherwise, it reads smoothly and is well constructed.

* * *

The following are examples of superior sentences.

1. The only thing about the lake that we were not too happy about was spending that expensive time swatting horse flies.

Explanation: This is a clever sentence which reads smoothly and is well-constructed. The words fit into the structure unobtrusively, and "horse flies" is a novel use of the word "horse".

2. The mood of the happy crowd changed suddenly when Cole Roger's most expensive horse missed a jump and plunged into the lake.

Expalantion: In this sentence the emphasis is taken off the given words by using "crowd" as the subject. A clever sentence which is well-constructed and reads smoothly.

3. A beautiful, peaceful lake and a spirited horse, neither of which were expensive in comparison to how happy they make her, proved the best cure for Ann's dejected state of mind.

Explanation: Did you have to read this twice to find all four words? The best sentences will use the words so they are not obvious. In addition it is a compound sentence which reads exceptionally well, and the adjectives are used to modify words other than given nouns.

4. The horse ran straight for the lake, tearing through the patio with the expensive furnishings and disrupting the once-happy gathering.

Explanation: This sentence has allthe characteristics of a superior sentence. It is well-constructed grammatically and reads smoothly. Although the given words are more prominent than in other superior sentences, such a vivid, imaginative situation compensates for that shortcoming. Also, the given adjectives do not modify the given noun.

When you have finished reading these examples and you understand the characteristics of superior sentences, turn to the next page.

Practice B

Using what you now know about good and superior sentences, pick the best sentence in each set of three which follow. Place an "X" in the place at the left of your choice.

Best (Your Choice)

____ A. 1. The expensive trip was not quite what they expected.

____ 2. Many people can be happy on a vacation which is not expensive, but consists of a week on a horse ranch near a lake.

____ 3. The kids are happy with the lake; it would have been too expensive to see that horse -- White Stalion.

____ B.1. Happy was very pleased by his gifts; he received an expensive wallet, a toy horse, and a toy motor boat that he could use when he is at the lake.

____ 2. The expensive horse made the boy near the lake happy.

____ 3. It is very hard to think of many sentences using the words happy, expensive, lake, horse.

Now that you have made your choice of the best, refer to the next page and see which is the correct choice. To further your understanding of the selection process, an explanation is included of why the choice was made as it was.

A. Correct choice of best:

#2. "Many people can be happy on a vacation which is not expensive, but consists of a week on a horse ranch near a lake."

Explanation: This was rated a 6 by the judges, because it uses all four words in a compound sentence in which the adjectives do not modify lake and horse. It is readable and the words, although not completely unobtrusive, do not stick out like a sore thumb. In the other sentences #1 does not use all four words, and #3 is clumsy and does not really make much sense.

B. Correct choice of best:

#1. "Happy was very pleased by his gifts; he received an expensive wallet, a toy horse, and a toy motor boat that he could use when he is at the lake."

Explanation: This was rated a 6 by the judges, because it is a complex sentence using Happy in a different context with expensive and happy not modifying lake and horse. #3 is just a list of the words, and #2 is a choppy sentence which is almost a list, for the words do stick out.

Once you have finished this, turn to the next page.

As before, select the one best sentence of each set of three and record your choice with an "X" on the left.

____ C.1. "Expensive Lake", John's horse, was not happy about the fact that he was drugged before the race.
____ 2. The woman, in expensive riding-habit, led the horse to the lake.
____ 3. The treatments of salty lake water, which were needed to keep the old horse happy, were beginning to get rather expensive.

____ D. 1. Horses can be expensive and bring happiness, but a lake can be as cheap as a glance and still make one happy.
____ 2. Horse Lake has an expensive and happy resort.
____ 3. The happy owner of the lake prospered from the sale of its expensive tracts of land to many horse buyers.

____ E. 1. Happy and carefree was I, as I rode the expensive horse around the mountain lake.
____ 2. We drove to the horse ranch on the other end of the lake.
____ 3. The experimenter repeated the four words only once; happy, expensive, horse, lake; the subject was then induced to retain them for two weeks.

____ F. 1. "I'd be happy to buy you a horse," Jim told his son, "but after sending you to the lake all summer, I think you realize that it would be too expensive."

____ 2. The happy, expensive horse walked over the lake.
____ 3. An expensive, happy horse welcomes a lake.

____ G. 1. The happy jockey of the horse that won the race bought an expensive house on a lake.
____ 2. The boy was last seen riding a black horse with an expensive saddle, heading in the direction of Happy Moonlake.
____ 3. The happy horse rode across the lake in an expensive boat.

____ H.1. She wished the horse was not so expensive so she might be able to take it to the lake and make everyone happy.
____ 2. He was happy near the lake with the expensive horse.
____ 3. Having an expensive home on a lake is a lot more to be happy about than having a horse.

____ I. 1. Happy, expensive, horse, and lake are all English words.
____ 2. The lake people are a different breed, happy and living with their large,expensive boats and a horse in every stable.
____ 3. The sight of the lake and its surroundings during fall made her very happy.

____ J. 1. The happy, expensive horse skipped over the lake.
____ 2. The happy, expensive horse wandered close to the lake for years.
____ 3. He wasn't too happy about going horseback riding at the lake because it was quite expensive.

Total Correct ____

As before, select the one best sentence of each set of three and record your choice with an "X" on the left.

Best (Your Choice)

____K. 1. Horse Branch is an expensive resort located near Happy Lake.
____ 2. The happy child rode the horse around the lake and by the expensive houses.
____ 3. Happy Lake is the home of many an expensive horse.

____L. 1. The horse was happy in the expensive horse trailer until they went off the road into the lake.
____ 2. At Lake Park today a horse named Happy won an expensive purse.
____ 3. I was not happy; our horse fell into the lake and it was too expensive to get him out.

____M. 1. After the swim around the lake we rode horses.
____ 2. The owner of the horse was happy when he won at Lake Raceway, so he celebrated with an expensive meal.
____ 3. The lake catered to expensive, happy people who liked horses.

____N. 1. Happy is the lake frontage which has an expensive horse standing near.
____ 2. The horse in the lake was very happy and expensive.
____ 3. The lake was stormy and the happy man turned sour when his expensive horse bolted and ran.

Total Correct ________

Practice C

Here are some more examples of superior sentences. You should study these examples carefully. Note their common characteristics.

1. The treatments of salty lake water, which were needed to keep the old horse happy, were beginning to get rather expensive.

   Explanation: This was one of the best 7's we have. The words are all used in a complex sentence, and you almost have to look twice to see them. Although expensive and happy do modify lake and horse in the strict sense, it is done cleverly. Also, and not without consideration when rated, this is a novel sentence; as you will see from some of the other examples, they can get rather stereotyped.

2. The boy was last seen riding a black horse with an expensive saddle, heading in the direction of Happy Moonlake.

   Explanation: Not a complex sentence, but it reads smoothly, and happy and expensive modify other nouns. Happy Moonlake is a clever "different context" usage.

3. Suddenly an expensive, foreign sportscar came out of nowhere, startling me; but soon I resumed my leisurely walk around the lake, and thought how happy I was to be a horse with none of the problems people have.

   Explanation: If there were criteria for cognitive complexity, this one would be the best. It is enjoyable, readable and novel while using all four words in a complex sentence with the adjectives modifying other nouns. In this one you really have to read it again to notice that all four words are there. Note, though this most superior sentence is long, length is no criterion for superiority, as #4 indicates.

4. I was not happy; our horse fell into the lake and it was too expensive to get him out.

   Explanation: A complex, cute sentence using the adjectives to modify other nouns. Note, novelty can make a superior sentence, too.

5. Although expensive, "horse" is a drug which makes the user feel happy, swimming in a lake of ecstasy.

   Explanation: Complex, "horse" is certainly a different context, smooth reading, imaginative, and novel.

6. I'll be happy to train the horse out at the lake, but I must warn you...it will be expensive for you.

   Explanation: This was rated 6 by one judge and 5 by the other, because it reads so roughly. It is complex and does fall into the 3-5 category, since the adjectives modify other nouns. Thus, it is one step beyond mediocrity.

Now that you have seen the characteristics of some of the superior sentences, for the contrast here are seven <u>inferior sentences</u> and their characteristics.

1. After the swim around the lake we rode horses.

   <u>Explanation</u>: Does not use all four words. Rated 1.

2. The sight of the lake and its surroundings during fall make her very happy.

   <u>Explanation</u>: Does not use all words. Rated 1.

3. We drove to the horse ranch on the other end of the lake.

   <u>Explanation</u>: Does not use all words. Rated 1.

4. Happy is the lake frontage which has an expensive horse standing near.

   <u>Explanation</u>: Have you ever seen a "happy lake frontage?" Expensive modifies horse, and it does not read smoothly. Rated 2.

5. The lake catered to expensive, happy people who liked horses.

   <u>Explanation</u>: Unreasonable because of a catering lake and expensive people. In "expensive, happy people," almost a list. Rated 2.

6. Horse and lake are nouns, but happy and expensive are adjectives.

   <u>Explanation</u>: Rated 2. A very poor sentence. A lazy way to avoid writing an imaginative sentence.

7. Happiness is a horse and a serene lake with the expensive cottage which you own.

   <u>Explanation</u>: Rated 1. Does not use the four given words. "Happiness" is no substitute for "happy" in this task.

After studying these sentences and explanations, turn to the next page.

These are some more examples of superior sentences. What characteristics do they share? Write your opinion in the place provided.

1. The trip to the lake was fun and it made us happy, but it was so expensive we had to eat horse meat.
2. H orses can be expensive and bring happiness, but a lake can be as cheap as a glance and still make one happy.
3. She wished the horse was not so expensive so she might be able to take it to the lake and make everyone happy.
4. "I'd be happy to buy you a horse," Jim told his son,"but after sending you to the lake all summer, I think you realize it would be too expensive.
5. Happy and carefree was I, as I rode the expensive horse around the mountain lake.
6. Happy was very pleased by his gifts; he received an expensive wallet, a toy horse, and a toy motor boat that he could use when he is at the lake.
7. Many people can be happy on a vacation which is not expensive, but consists of a week on a horse ranch near a lake.

What characteristics do these superior sentences share? ____________

____

____

These are some more examples of inferior sentences. What characteristics do they share? Write your opinion in the place provided.

1. The expensive trip was not quite what they expected.
2. A lake is not complete without a big sailboat.
3. The happy, expensive horse wandered close tothe lake for years.
4. An expensive, happy horse welcomes a lake.
5. Happy, expensive, horse, and lake are all English words.
6. Lakes aren't expensive, but horses are and happy.
7. The words happy and expensive are adjectives modifying the nouns horse and lake.

What characteristics do these inferior sentences share? ________

____

____

After you have recorded your opinion, turn the page ......

These are some more examples of inferior sentences. What characteristics do they share? Record your opinion as before.

1. Diamonds are very expensive.

2. The kids are happy with the lake; it would have been too expensive to see that horse - White Stalion.

3. It is very hard to think of many sentences using the words happy, expensive, horse, lake.

4. The experimenter repeated four words only once; happy, expensive, horse, lake.; the subject was then induced to retain them for 2 weeks.

5. Expensive Lake had a happy horse present.

6. The happy, expensive horse is swimming in the lake.

What characteristics do these inferior sentences share? ______

____

____

These are some more examples of superior sentences. What characteristics do they share? Write your opinion as before.

1. I was very happy while I stayed at the expensive lake resort, except for the fact that the girls looked like cows, or should I say horses?

2. After it was all over, he wasn't very happy that he had eaten like a horse at that expensive restaurant overlooking the lake.

3. You have a right to be happy, but don't horse around with your boat on the lake, as it may turn out to be an expensive venture.

4. While riding a horse around a secluded lake, you couldn't imagine how happy I was to find this expensive watch I'm wearing.

5. He wasn't too happy about going horseback riding at the lake because it was quite expensive.

6. The lake people are a different breed, happy and living with their large, expensive boats and a horse in every stable.

What characteristics do these superior sentences share? ________

____

____

After you have finished these, STOP -- wait for further instructions.

Up to now you have been practicing the evaluation of sentences which were written by <u>other</u> students. Now the real test of your judgment ability comes. Go back to the sentences which you wrote on the first page. Read each sentence carefully and select the <u>three</u> (3) best according to the criteria of superior sentences which you have been using. Put an "X" by these three.

Then, reread these three sentences and put a line completely around the <u>one</u> best.

Appendix E.
Judgment Test: Sentences

Pick the one best sentence from each set or group of sentences which are given below.

Record your answers on the IBM answer sheet you have been given. Work rapidly for you have 10 minutes to complete this task.

Make sure you do not make any errors in recording your answers on the IBM answer sheet. DO NOT MAKE ANY MARKS ON THE TEST BOOKLET.

1. 1. The happy people lived by an expensive lake and each had a horse.
   2. The people by the lake had expensive horses and were happy.
   3. The lake was calm until a happy, expensive horse jumped in.
   4. We threw an expensive party at our cottage near the lake, and everyone was happy to ride by horse once.
   5. The expensive horse walked happily down to the lake for a drink.

2. 1. The formerly happy family with the expensive car lost all their money, forcing them to buy a family horse and shack on a dirty lake.
   2. The beautiful, but expensive horse down by the lake looked happy and content.
   3. As the horse looked into the lake water he saw a happy looking face staring back at him with an expensive looking watch tied to his pocket.
   4. The lake is the home of expensive but happy horses.
   5. Happy again was I while riding my horse around the lake in my expensive new outfit.

3. 1. The horse was not happy because his expensive house was lost when the lake overflowed.
   2. An old-fashioned horse and buggy ride around the lake is fun, but an expensive auto makes everyone happier.
   3. The horse happily drank water from the lake, by which expensive cottages were built.
   4. The horse was not happy because he lost his expensive saddle in a lake.
   5. The happy man with an expensive hat on rode his horse up to the lake.

4. 1. I wasn't very happy to find out how expensive a horse was because I was looking forward to riding up at the lake this summer.
   2. A lake of perspiration ran off the expensive horse after the happy owner had put him through his paces.
   3. Lakes and horses usually make children happy but they are also two very expensive items for children's parents to buy.
   4. Horses are most happy when they are away from expensive surroundings and near a lake.
   5. The lake, which is an expensive place to keep a horse, makes many people happy each year.

5. 1. Although I like the horse, I am happy near the expensive lake.
   2. The lake seemed to be a happy medium for the horse but it was expensive.
   3. An expensive horse seems more happy living by a lake.
   4. The happy man who bet on the winning horse spent the weekend at Louisville's most expensive lake resort.
   5. The huge lake was surrounded by happy people living in big, expensive houses with horses and other objects of luxury.

6. 1. The horse monster which inhabited the lake was expensive to the resort in the area but it made the sadists happy.
   2. Although the people watching the horse show wore obviously expensive clothing, they did not seem as happy as the penniless beach-bums on the other side of the lake.
   3. To be happy is priceless, yet not as expensive as a racing horse or cottage on a lake.
   4. The expensive horse was happy because he lived by a lake.
   5. I had a happy time at the lake, but it turned out to be rather expensive because I was kicked by a horse.

7. 1. An expensive horse grazes in a field near a lake which is owned by happy people.
   2. An expensive hobby is keeping a horse happy and near a lake.
   3. The happy child, who fished at the lake with an expensive pole, rode the horse.
   4. We are happy when we go to our expensive cottage on the lake where we have our horse.
   5. A horse is most happy when it is near an expensive lake.

8. 1. Lakes contain many happy animals like the sea horse, plus many expensive stones like the pearl.
   2. The horse was happy as long as he stayed in an expensive stable near a lake.
   3. An expensive horse was purchased by happy people who have a lake.
   4. The happy owners keep their horse in an expensive stable overlooking a lake.
   5. The woman often became happy while riding the expensive horse near the lake.

9. 1. The lake atmosphere is calming to the horse, whose expensive tastes in feed do not bring happy smiles to Len's face.
   2. The race horse, "Happy", was not expecting the murky lake as he cleared a jump which proved expensive.
   3. The lake was cold as the horse plunged in and threw its rider, happy and laughing, ruining his expensive leather boo
   4. Bill rode his horse, with its expensive gear, down the trail to Happy Lake.
   5. I was happy until I lost my expensive ring while walking the horse around the lake.

10. 1. The horse had an expensive barn by a lake, and he was very happy.
   2. The expensive horse looked happy as he drank from the lake.
   3. To make the child happy, they had an expensive lunch, rode horses and swam in the lake.
   4. I bought an expensive horse which galloped through the lake giving me happy satisfaction.
   5. Horse tails grew along the lake where my horse happily grazed.

Judgment Test: Sentences 2

Pick the one best sentence from each set or group of sentences which are given below.

Record your answers on the IBM answer sheet you have been given. Work rapidly for you have 10 minutes to complete this task.

Make sure you do not make any errors in recording your answers on the IBM answer sheet. DO NOT MAKE ANY MARKS ON THE TEST BOOKLET.

1.
   1. As I sat alone, I heard a big noise which turned out to be money, as it rolled off the table.
   2. If you are alone, it is no fun to make a big noise, and celebrate, even if you have the money.
   3. Alone in a big desert, one has no need for money or the noise of society.
   4. I was alone making a noise by a big bass drum, when, upon breaking its head, I found a lot of money.
   5. The big bear made a lot of noise and I was afraid, but I still shot my rifle; I knew I had to get my money for him, I, now being alone in this world.

2.
   1. I was alone in the bank and the money I dropped made a big noise.
   2. Big Bob had money to blow but when left alone without friends his blustering and noisy ways did not make him good company in the rush and noise of the big city life.
   3. The money alone will make an impressive noise on the big executives.
   4. The money made a big noise when it hit the floor being pushed off a table by a boy who was alone and wanted someone to come.
   5. Alone, money can make a big noise.

3.
   1. All the money, and the noise of the big city could not turn the newly divorced man away from his feeling of aloneness.
   2. Although the boy was not very big, he alone at the sound of the loud noise, ran into the collapsing building and saved all of the money.
   3. The big Pierce Arrow made a lot of noise and cost a lot of money, yet somehow it knew it was like a dinosaur-the last of its kind and very much alone.
   4. Big money lake is a nice place to escape the noise of the city and to be alone.
   5. The noise alone in the arena of the stock exchange could prevent someone from wanting to get big money.

4.
   1. The big dog made a lot of noise when left alone, so we spent more money to buy him a companion.
   2. Many people are so alone in this world that they have to make a lot of noise about being big and having a lot of money.
   3. One balloon which had cost a lot of money, made a big noise, and was thrown away and left alone.
   4. The noise of the big cities, and having money to spend, can still make a person feel alone.
   5. Alone and without money can be more frightening than a big noise.

- Cont. on next page -

Judgment test (cont.)

5. 1. Money alone cannot make a big man but it can cause more noise out of him.
   2. The big star appeared to be standing alone among the noise of those exchanging money.
   3. Big and alone, is the noisy money man.
   4. When you are alone amidst a great deal of clamor noise, money does not make a noise or have any effect.
   5. I was alone counting my money when I heard a noise at the door.

6. 1. Money, noise, and big ambiguity make me feel alone many times.
   2. The big money alone man is noise.
   3. The big noise of the construction meant for the boss that he was not alone in the world of big business.
   4. My money made me a big man but there was no noise from the crowd.
   5. Big was the only word for the safe which contained alone his money in the noise-proof room.

7. 1. Alone in the street I made a big noise so no one would come and take the money.
   2. A burglar with a big gun traveling alone demanded money and told me not to make noise.
   3. I spent a lot of money in the big city with the noise and crowds and felt alone.
   4. To go to a big noisy party will cost you less money if you go alone.
   5. He paid money to be alone yet a big noise would be heard every hour.

8. 1. Professional trumpet players make big money and a lot of noise, but I think they must feel all alone in front of their audiences.
   2. How much noise can a big football star make when his money demand alone is not met?
   3. Dick, being alone, went for a walk and found some money for a big bash with lots of noise.
   4. Alone, except for the money I needed, I built a big room where no noise could enter.
   5. The noise was loud and all I wanted to do was to take my big bag of money and be alone.

9. 1. The big boy stood alone in a room without noise.
   2. The big noisy money man is alone.
   3. Money alone is often big enough to speak for itself, or at least create a lot of commotion and noise.
   4. The money made a big noise as it was alone in the collection plate.
   5. Jane sat alone counting money in the cashier booth of Big Water Amusement Park until the noise from the crowd died down and she realized it was time to close.

10. 1. The noise from Big Lake Amusement park alone kept them from making money.
   2. Money alone can make you seem big, but balloons make most noise when they break.
   3. The big noise was nothing but the sound of Jimmy alone playing with his money.

- Cont. on next page -

Judgment test cont.

4. Why can't I be alone -- I'm big, can make lots of noise, and have money to get anything.
5. Mr. Big made lots of noise when he heard he had won "Alone" the great money jackpot.

11. 1. The money made noise as I jingled it standing alone in a big park.
2. I was alone in the big vault, when the loud noise caused me to run and drop the bags of money.
3. Because I wasn't very big the noise from all that money in my pocket aroused me to feel alone.
4. The babysitter really earned her money, for the child put up a big fuss and made a lot of noise when he was left alone with her.
5. The noisy big bear at money lake frightened us because we were alone.

12. 1. The noise of the silver money made the worker feel very big next to the teller who was alone.
2. Money and noise are symbols of being big and great, but often they only make one feel alone.
3. A big person but alone and the money making noise in my pocket is of little comfort.
4. With big sums of money he was never alone, but could always be with the crowds and the noise.
5. The big dog which had cost a great amount of money, made much noise when it barked.

STOP - DO NOT TURN THE PAGE! WAIT FOR FURTHER INSTRUCTIONS.