LINDA S. BUYER

Creative Problem Solving: A Comparison Of Performance Under Different Instructions*

There are too few examples of experiments which directly compare competing theories of creativity (Treffinger, Isaksen, & Firestien, 1983). The present study makes such a comparison and raises the issue of what measures of creative production are to be used. Newell, Shaw, and Simon (1962) argued that creative problem solving is essentially the same as ordinary problem solving. They see the process of solving a problem, be it creative or otherwise, as the construction of a problem space (the solver searches for and evaluates solutions. From this it follows that instructions which constrain the subject to construct a problem space encompassing the critical components of the solution should increase the quality of the responses generated. In contrast, Osborn (1957) argued that creative problem solving (which he terms "ideational") operates differently from ordinary ("analytic") problem solving. Osborn felt that ideas which are generated later in the response sequence would be more creative. He therefore saw the process of creative problem solving as the unrestricted generation of ideas which are only subsequently judged as to quality. Hence, one would predict that instructions to produce as many alternatives as possible while reserving judgment should result in greater creativity.

One of the critical questions to be considered here is what constitutes "greater creativity". Osborn (1957) and Newell, Shaw, & Simon (1962) appear to consider different measures of creative output. Osborn seems to prefer a count of the number of creative solutions, without regard to the overall quality of the responses while Newell, Shaw, & Simon would seem to advocate a measure of either the mean creativity of responses or the proportion of superior responses, as both measures take the entire

*This article is based on a thesis submitted in partial fulfillment of the requirements for a Master of Arts degree at the University of Illinois at Chicago.

response set into account. Where one is interested in the creativity of a "best" single response, creativity ratings of either a subject's best (as measured by judges) or preferred (best as assessed by the subject) responses are other possible measures of creative output, again disregarding the composition of the entire response set.

This study was designed to assess the performance of subjects working under different instructions on all five of the aforementioned measures of creative production. Subjects were given either Brainstorming, Criteria-cued, or Standard instructions (Gerlach, Shutz, Baker, & Mazer, 1964) and asked to apply them to five different problems.

SUBJECTS

Subjects were 75 students registered in introductory psychology classes who participated in the experiment for class credit. The subjects were randomly assigned in equal numbers to the three instructional conditions.

PROCEDURE

Each subject received a booklet with one set of instructions based on those used by Gerlach, Schutz, Baker & Mazer (1964) and five problems similar to those used by Johnson, Parrott, & Stratton (1968). The problems were chosen to represent verbal, numerical, and pictorial areas of endeavor. Similar requirements were imposed by all five tasks: all responses were written, and none required computation or drawn responses. The problems were as follows:

1. Produce a title for a short story.
2. Construct a sentence using all of the words in a word set.
3. Complete the verbal portion of the last panel of a cartoon.
4. Write a title for a table of data.
5. Write a summary for a graph.

The tasks were arranged so that, across subjects, each appeared equally often in each sequential position. The general instructions were on page 1 of the booklet and, appropriately modified, appeared above each problem. The subjects were allowed to work for ten minutes on each task. For all conditions, the last page of the booklet instructed subjects to look over all their responses and choose the best one for each task.

All responses were individually typed onto 3″ x 5″ index cards with an eight digit code to designate subject, task, type of instruction, and response number to insure that the raters were blind to the conditions and subject characteristics.

Raters of creativity were two psychology graduate students. The raters rank ordered the responses according to degree of creativity which was defined as a combination of originality and practicality. Each rater first rank-ordered the responses (tied ranks were allowed) within each of the tasks and then, maintaining their within task orderings, rank ordered the responses across the tasks. Each response was assigned a creativity rating corresponding to its rank order in the across-task ordering, e.g., all responses in the lowest ranked category received a creativity rating of 1. The inter-rater reliability (Spearman rank order correlation) ranged from +.60 (Sentence Completions) to +.80 (Table Titles). The average of the two ratings was used in all subsequent analyses.

RESULTS

The group working under the brainstorming instructions produced more responses (44% of the total number) than either of the other two groups. Subjects in the Criteria-cued and Standard instruction conditions produced approximately equivalent numbers of responses (28% and 29% of the total number of responses respectively). See Table 1.

The number and percent of superior responses (those in the upper 10% of the rating distribution) are shown in Table 1. Subjects in the Brainstorming and Criteria-cued conditions produced approximately the same number of superior responses (120 and 123 respectively) and both groups produced substantially more superior responses than subjects in the Standard instruction condition (71). Almost 15% of the responses produced by subjects in the Criteria-cued condition were superior as compared with less than 10% for both of the other conditions.

TABLE 1 Total Number, Number Superior, and Percent Superior Responses Under Different Instructions

	Response Measure		
Instruction	Total Number of Responses	Number of Superior Responses	Percent of Superior Responses
Criteria-cued	835	123	14.7
Brainstorming	1317	120	9.1
Standard	865	71	8.2
Total	3017	314	10.4

The overall mean rating, the rating of preferred responses (those identified by the subjects as their best responses), and ratings of best responses (those identified by the raters as each subject's single best response to each task) were computed for each task/instruction combination for each subject. Three repeated measures analyses of variance with three classifications of instruction and five task variables were performed, one using average creativity ratings, one using preferred creativity ratings (two subjects were excluded from this analysis because they failed to identify their preferred responses), and one using best creativity ratings as the dependent measures. Only the analysis of average creativity resulted in a significant effect of instruction. The effect supported the Information Processing prediction that criteria-cued instructions would result in better performance, e.g., mean creativity was highest with Criteria-cued instructions; $F(2,72)=5.57$, $p \leq .02$, MSE=141.33. All three analyses resulted in significant effects of task: $F(4,288)=86.40$, $p \leq .001$, MSE=81.86 for mean creativity of responses, $F(4,280)=31.94$, $p \leq .001$, MSE=264.29 for creativity rating of preferred responses, and $F(4,288)=48.83$, $p \leq .001$, MSE=185.81 for creativity rating of best responses. There were no significant interactions. The summary statistics are shown in Table 2.

TABLE 2 Average Creativity Ratings for Mean, Preferred, and Best Responses

	Response Measure		
Instruction	Mean	Preferred	Best
Criteria-Cued	44.38	46.56	58.98
Brainstorming	40.71	42.68	59.02
Standard	39.59	43.81	53.52
	Task		
Story	50.73	52.73	64.95
Sentence	47.56	51.60	67.23
Cartoon	44.73	46.47	57.80
Table	39.37	44.20	56.67
Graph	25.84	26.76	39.24

To test Osborn's assumption that better responses should occur later in the response sequence an analysis was performed to test whether there was a preponderance of best responses (as determined by judges) in the second half of the subjects' response

sequences. The position of the best responses was determined by dividing each subject's responses for each task into halves based on the order of occurrence and assigning a value of 1 to best responses which occurred in the first half of the sequence, a value of 2 to responses which occurred in the second half, and a value of 1.5 to responses which were either in the middle of a sequence of uneven length or single responses. For subjects who had more than one "best" response, the average position of best responses was computed for each subject by adding the position values and dividing by the number of best responses. The position scores were then averaged over subjects within each instructional condition. A repeated measures analysis of variance resulted in no significant effects. A t-test was performed to determine whether the grand mean (1.53) was significantly different than 1.5 which is the value of the mean of a population in which best responses are randomly distributed throughout a series of responses. The t-test was not significant, which suggests that best responses are indeed randomly distributed throughout the responses.

A further consideration of the quality/quantity relationship involved the overall quality of responses generated in the first and second halves of the response sequences. Subjects' responses were divided into first and second halves of their response sequence (for subjects who only wrote one response to a task it was counted in both the first and second halves) and the mean ratings for each half were computed. Mean creativity of responses generated in the first half of the production sequence was slightly higher than for those generated in the second. This does not, however, hold true for the Sentence Construction task. Under all three instructions responses written in the second half of the response sequence for the Sentence task have higher mean ratings than those written in the first half. There was a half by problem interaction; $F(4,288)=2.79$, $p \leq .05$, $MSE=75.13$. Otherwise there doesn't appear to be any effect of the number of responses generated (time) on the creativity of the responses. The means are shown in Table 3.

TABLE 3 Means of First and Second Half Responses

	Tasks					
Half	Cartoon Completions	Table Titles	Sentence Construction	Graph Summary	Story Titles	Mean
First	45.39	40.39	45.12	26.84	52.30	42.01
Second	44.78	39.13	49.08	26.38	50.17	41.91

DISCUSSION

Creativity was measured in a variety of ways. The measure of creativity which Osborn (1957) seems to endorse is the number of superior responses. The total number of superior responses was approximately equal for the Brainstorming and Criteria-cued instruction conditions. Both instructions resulted in a greater number of superior responses than Standard instructions. Relative to Standard instructions, Osborn is correct in predicting that Brainstorming instructions will result in a greater *number* of highly creative responses. However, the Criteria-cued condition resulted in a greater *percentage* of superior responses than either the Brainstorming or Standard instructions. It appears that Brainstorming instructions increased the number of responses at all levels of quality while the Criteria-cued instructions increased the number of superior responses without promoting the production of responses of lesser quality. While it is true that Brainstorming instructions increased the number of superior responses, as Osborn would predict, it appears that the Criteria-cued instructions did so in a more efficient fashion.

Brainstorming theory depends, in addition, on the assumption that as problem solvers continue to produce possible solutions their productions will increase in creativity. The analyses of the position of best responses and mean creativity of first and second halves of the response sequences failed to confirm this assumption.

Of the additional analyses of creativity, only the analysis of average creativity of responses resulted in an instruction effect. This effect is most likely beacuse the Criteria-cued instructions inhibited subjects from expressing responses not meeting the solution constraints. This interpretation is supported by a re-analysis of the average creativity ratings with all irrelevant responses excluded. Irrelevant responses were those which did not conform to the task demands, e.g., a cartoon completion which had a person speaking who was not present in the final panel, as judged by two independent raters. With irrelevant responses excluded, there was no effect of instructions on mean creativity.

The differences among tasks on several creativity measures may reflect the subjects' familiarity with the content. A premise which underlies much of the research on creativity (MacKinnon, 1970) is that one must have some degree of expertise or familiarity with a content area before one can be creative in that domain. Support for this interpretation is offered by the fact that 57% of the responses to the graph summary task (which had the lowest creativity ratings) were excluded as irrelevant, which suggests that perhaps this task was one for which the subjects did not have the requisite prior knowledge of how to read a graph. The failure to find

any task by instruction interactions show that whatever is causing the task effect is independent of the instruction effect on the creativity of the responses.

Taken as a whole these results suggest that, when one considers the composition of the entire response set (average creativity and percentage of superior responses), Criteria-cued instructions seem to improve creative performance. When one considers a subset of the entire response set (number of superior responses, creativity ratings of best and preferred responses), only the measure of the number of superior responses differentiated among the instructions. Brainstorming and Criteria-cued instructions were both superior to the Standard instructions but were indistinguishable from one another.

Thus, the resolution of the dispute over which type of instruction produces the most creative results may turn on the type of measure of creativity used. For two of the measures used (percentage of superior responses and mean creativity rating) Criteria-cued instructions produced the most creative results. For one of the measures (number of superior responses) Brainstorming and Criteria-cued instrcutions affected subjects' performance similarly. On the final two measures (creativity ratings of best responses and creativity ratings of preferred responses) there were no differences among any of the instructions. Therefore, it seems that using Criteria-cued instructions will result in improved creative performance wherever it is possible to increase creativity. It will also produce this result efficiently, i.e., getting the largest number of creative responses for the smallest total response output.

REFERENCES

GERLACH, V.S., SCHUTZ, R.E., BAKER, R.L., & MAZER, G.E. Effects of variations in test directions on originality test response. *Journal of Educational Psychology*, 1964, *35*, p. 79-83.

JOHNSON, D.M., PARROTT, G.L., & STRATTON, R.P. Production and judgment of solutions to five problems. *Journal of Educational Psychology*, 1968, *59*, (6, Pt. 2).

MACKINNON, D.W. Creativity: A multi-faceted phenomenon. In Roslansky, J.D. (ed.), *Creativity: A discussion at the Nobel conference*. Amsterdam: North-Holland Publishing Co., 1970.

NEWELL, A., SHAW, J.C., & SIMON, H.A. The process of creative thinking. In Gruber, H.E., Terrell, G., & Wertheimer, M. (eds.) *Contemporary approaches to creative thinking*. NYC: Atherton Press, 1962.

OSBORN, A. F. *Applied imagination: Principles and procedures of creative thinking* (revised edition) NYC: Charles Scribner's Sons, 1957.

TREFFINGER, D. J., ISAKSEN, S. G., & FIRESTIEN, R. L. Theoretical perspectives on creative learning and its facilitation: An overview. *Journal of Creative Behavior*, 1983, *17*, pp. 9-15.

Linda S. Buyer.
Address: University of Illinois at Chicago, Department of Psychology, Box 4348, Chicago, IL 60680.