

Facilitation of Performance on a Divergent Measure of Creativity: A Closer Look at Instructions to "Be Creative"

Ronald G. Evans and Gary B. Forbach
Washburn University

Research has suggested that instructing subjects to "be creative" enhances performance on alternate uses creativity tasks. However, previous research has confounded the effect of instructions to "be creative" with the effects of guidelines defining response quality criteria. These factors were separated by providing one of four instructions to 132 college students: creativity (C/NG), creativity plus guidelines (C/G), guidelines only (S/G), or standard alternate uses instructions (S/NG) that do not mention creativity or provide guidelines. Results indicated that males performed best in the C/G condition and females did so in the C/NG condition; standard instructions were least beneficial in eliciting maximal creative performance. The role of instructions in improving valid assessment of creativity is discussed. Suggestions for further research on the observed sex differences in response to instructions are presented.

In most areas of ability testing, special care is taken to provide examinees with instructions that clarify task demands and generally provide a set for performing as well as possible. An exception to this principle, however, is found in one of the most common methods for assessing creative thinking abilities: alternate uses tests. In this approach to measuring divergent thinking (Guilford, 1967), standard instructions stress only that uses other than the common one be listed for various objects. No explicit reference is made to the test as being a

creativity measure, nor are any criteria provided on which examinees can base their responding. In pointing out these factors, Harrington (1975) has argued that the standard alternate uses instructions may actually inhibit creative responding and thus obviously jeopardize the validity of this type of creativity index.

Support for this contention has come from Harrington (1975) and Katz and Poag (1979), who found that instructions to "be creative" when generating alternate uses increased the quality of responses given by males. Katz and Poag also reported an increase in quantity of responses under these instructions versus standard instructions for males, as well as an increase in quality accompanied by a *decrease* in quantity for females. In explaining their findings, Katz and Poag suggested that creativity instructions "disambiguate" the task for males, leading them to switch to a verbal problem-solving approach from what Kogan (1972) has proposed to be their habitually nonverbal style. The reasons for females' apparent decrement in quantity of ideas generated is less clear, since females were predicted to be less affected by creativity instructions, due to an already habitually verbal problem-solving style.

Although both Harrington (1975) and Katz and Poag (1979) discuss the facilitative effect of instructions to "be creative" (at least for males), a closer inspection of their instructions reveals that neither study clearly demonstrated that labelling the

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 7, No. 2, Spring 1983, pp. 181-187
© Copyright 1983 Applied Psychological Measurement Inc.
0146-6216/83/020181-07\$1.35

alternate uses task as a creativity measure is responsible for this effect. The principal goal of the present study involved clarifying this issue and thus, it is hoped, enhancing the sensitivity of alternate uses in measuring creativity.

In both of the above studies, creativity instructions differed from standard instructions in two potentially important ways: (1) subjects were told to "be creative" and (2) they were given what the present article will refer to as guidelines, i.e., "a creative use is one which is both unusual—a use which other people would be unlikely to think of—and worthwhile" (Katz & Poag, 1979, p. 522). Standard instruction subjects received no such guidelines concerning quality, nor were they told to be creative. Which factor, then, accounts for observed standard versus creativity instruction effects? For example, increments in males' performance could have been due to the provision of criteria for responding (the guidelines) rather than to instructions to "be creative" per se. Similarly, given results for females (Katz & Poag, 1979), it is of special interest to determine which instruction component was responsible for their decrement in response quantity under creativity instructions, as well as what instructions can enhance creative responding for these subjects.

To evaluate these questions, the present study systematically varied the two instructional elements of "be creative" and guidelines. This resulted in two conditions identical to those in previous research (Harrington, 1975; Katz & Poag, 1979)—standard instructions and creativity plus guidelines—and thus allowed a replication of these studies. In addition, one group of subjects received instructions only to "be creative" (without guidelines), and another group was given the guidelines without references to the term creativity.

Method

Subjects

A total of 132 introductory psychology students (61 males, 71 females) served as subjects in exchange for partial course credit.

Materials

Subjects completed a battery of three tests, including the divergent creativity measure (alternate uses) and two nondivergent tasks (Remote Associates and Digit Symbol).

Alternate Uses Test. This was composed of three problems, with subjects asked to provide uses for a newspaper, pencil, and a two-liter plastic bottle. The first two problems were used by Katz and Poag (1979), and the third has been found useful in previous research by Forbach and Evans (1981). Subjects wrote their responses on sheets of 8.5×11 -inch lined paper and were given 6 minutes per problem.

Remote Associates Tests. The Remote Associates Test (RAT; Mednick & Mednick, 1967) was included to evaluate possible generalization of divergent creativity instruction effects to nondivergent creativity tasks (see Katz & Poag, 1979). Each RAT item asks subjects to provide the correct associative element linking three disparate stimulus words (e.g., the correct associate for *rat*, *blue*, and *moon* is *cheese*). In the present study a 15-item RAT was used, which included items that pretesting revealed to have the highest correlations with total score for the original 30-item version.

Digit Symbol. This subtest from the Wechsler Adult Intelligence Scale (Wechsler, 1955) requires completing a series of symbol substitutions which match number-symbol relationships provided by a key. It was administered under the standard 90-second time limit, with total score defined as the number of correct substitutions. The Digit Symbol provides a measure of perceptual speed and freedom from distraction (Zimmerman & Woo-Sam, 1973) and was included to evaluate any nonspecific arousal effects engendered by the various instructional conditions for alternate uses (see Katz & Poag, 1979).

Procedure

Subjects were run in mixed gender groups of two through six, with each person seated alone at a table. Two undergraduates (a male and a female) served as experimenters and were randomly as-

signed the groups and instructional conditions which they ran. The experimenters were blind to the background and purposes of the study.

After subjects were seated, the experimenter explained that the study involved taking a battery of tests. He/she then switched on a tape recorder, which presented all further instructions. Each group of subjects was assigned, by a randomly predetermined order, to receive one of four instructional conditions for the alternate uses problems: *Standard* instructions with no guidelines (S/NG), $n = 32$; *Standard plus Guidelines* (S/G), $n = 34$; *Creativity* instructions with no guidelines (C/NG), $n = 33$; or *Creativity plus Guidelines* (C/G), $n = 33$. The basic instructions for the two Standard groups (S/NG and S/G) are presented below with the instruction differences for the creativity groups in parentheses:

The following is a test of your ability to think (creatively) about alternate uses for some common objects. Each of the objects you will be asked to think about has a common use with which you are familiar. Your task is to list as many other (creative) uses for which the object or parts of the object could serve. **Do not spend too much time on any one item. Write down those (creative) uses that occur to you. There are three parts to the test. You will have six minutes per problem.

**At this point, subjects in the appropriate groups were told that an alternate use (S/G

group only) or a creative use (C/G group only) is one "which is both unusual—that is, a use which other people would be unlikely to think of—and worthwhile."

Subjects in all experimental groups then completed the alternate uses problems, followed by the Digit Symbol and RAT.

Results

Alternate Uses Test

Ideational fluency. The quantity (total number of uses given) across the three problems was first obtained for each subject. A 2 (creativity vs. standard instruction) \times 2 (guidelines vs. no guidelines) \times 2 (gender of subject) unweighted means ANOVA revealed a main effect for the guidelines variable [$F(1, 123) = 10.70, p < .001$] and an interaction of creativity/standard instructions with guidelines [$F(1, 123) = 5.68, p < .02$]. Of greater interest, though, was the presence of the three-way interaction [$F(1, 123) = 62.49, p < .001$]. Means and standard deviations for all instruction groups by gender of subject are presented in Table 1.

Post-ANOVA comparisons (Newman-Keuls test, $\alpha = .05$) indicated no differences among males across the four instruction conditions. Instead, the creativity/standard \times guidelines/no guidelines interaction was accounted for within the female group. Here, the C/NG group gave significantly more responses than the C/G group. Although C/NG ex-

Table 1
Mean and Standard Deviation of Number of Alternate Uses Given by Males and Females for Creative and Standard Instruction Conditions

Condition	Males		Females	
	Creative	Standard	Creative	Standard
Guidelines				
Mean	31.00	30.47	25.11	32.32
S.D.	12.96	9.93	10.59	12.28
No Guidelines				
Mean	35.81	31.07	42.85	35.41
S.D.	11.08	10.20	15.69	11.22

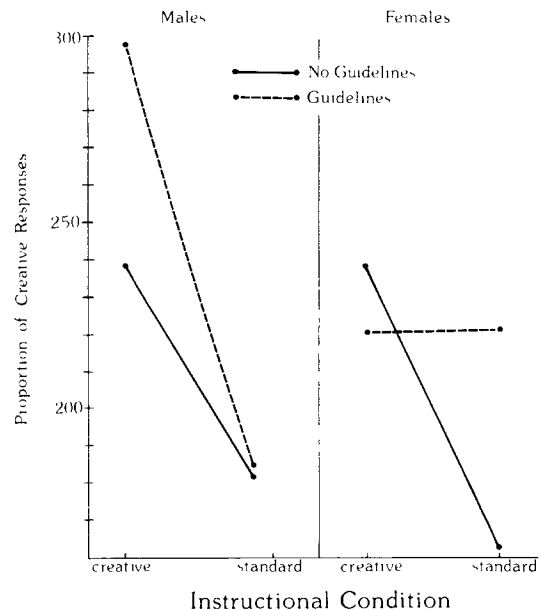
ceeded S/NG, and S/G exceeded C/G, these differences were not significant ($p < .25$). Thus, even though guidelines generally produced a decrement in quantity of responses over no guidelines for females, this was principally the case when the additional instruction to "be creative" was included.

Quality of responses. Each response was also scored for its creativity using a scoring procedure identical to Katz and Poag (1979). This procedure produces a creativity score which reflects the joint contribution of worthwhileness and infrequency. Briefly, two independent raters, blind to the subjects' experimental condition, rated the worthwhileness of each alternate use response on a 1 through 4 scale (4 = extremely worthwhile). This rating was then divided by the frequency of the response in the total sample. Frequency, like worthwhileness ratings, is not perfectly reliable, since similar responses are not always judged to be identical (e.g., use a newspaper "in the bottom of a birdcage" versus "to houstrain a dog" might be judged as separate responses by one rater and two instances of the same response by another). Consequently, interrater reliability was calculated for the derived worthwhileness/frequency ratio scores. These reliabilities ranged from .62 to .80 across problems.

The mean worthwhileness/frequency ratio was then calculated for each alternate-use problem, with individual responses falling above this mean designated as creative. Finally, the proportion of each subject's total responses thus deemed creative (i.e., both worthwhile and infrequent) was determined to provide the dependent qualitative measure of creative responding. Use of this proportion provides a control for individual differences in ideational fluency. Katz and Poag (1979) provide more discussion of this scoring procedure.

These qualitative proportions, collapsed across raters and uses problems, were analyzed in a three-factor ANOVA corresponding to that for ideational fluency. This analysis revealed a main effect for creativity versus standard instructions [$F(1,123) = 11.19, p < .01$], as well as the three-way interaction [$F(1,123) = 3.96, p < .05$]. Mean proportion of creative responses for all groups are presented in Figure 1.

Figure 1
Proportion of Total Responses Judged to be Creative as a Function of Gender of Subject and Test Instructions



As can be seen in Figure 1, there was a marked difference for instruction conditions across the gender groups. Among males, those in the C/G group gave a significantly higher proportion of creative responses (Newman-Keuls test, $\alpha < .05$) relative to the other three groups, with the C/NG group significantly higher than either standard group. For females, the S/NG group was significantly lower than the other three groups, which did not differ.

Overall, then, type of instruction did not affect the quantity of responses given by males, although the quality of responding was enhanced considerably by the combination of creativity plus guidelines instructions. For females, though, quantity was greatest when instructions referred to being creative but omitted guidelines (the C/NG group); quality was similar in all groups but S/NG, where quality scores were significantly lower. In other words, the combination of quality and quantity of alternate use response was greatest for males under CG instructions and females under C only instructions. Both gender groups gave a low proportion

of creative responses under the standard (S) instructions, which did not refer either to creativity or to guidelines for responding.

Remote Associates Test

The three-factor ANOVA produced no significant effects for this variable, supporting the Katz and Poag (1979) conclusion that differing instructions on the divergent creativity task (alternate uses) do not facilitate performance on a nondivergent verbal task.

Digit Symbol

The three-factor ANOVA on Digit Symbol total scores revealed only a significant interaction of the guidelines variable with gender of subject [$F(1,123) = 6.42, p < .02$]. Here, males had significantly (Newman-Keuls test $\alpha < = .05$) higher scores when given guidelines ($M = 66.10$) than when given no guidelines ($M = 58.32$), while the opposite pattern was found among females ($M = 68.93$ versus 63.90). Thus, whether or not the task was explicitly defined as requiring creativity, males seemed best able to focus their attention and effort when guidelines were provided for alternate uses, while the absence of guidelines had a similar effect upon females.

Discussion

The present findings provide clear support for Harrington's (1975) contention that standard instructions for alternate uses tasks do not elicit subjects' maximal performance. Of greater interest, though, is the finding that the role played by instructions to "be creative" is more complex than previously described. Before elaborating on this issue, some points of direct comparison between the present study and prior investigations should be noted.

It will be recalled that Harrington (1975) and Katz and Poag (1979) compared what has been termed creativity plus guidelines (C/G) with standard (S/NG) instructions. Results for those two groups in the present study are quite consistent with

both prior studies. In all three cases, males improved the quality of their alternate use responses under C/G versus S/NG instructions; while in the two studies which included females (Harrington did not), those subjects produce fewer, but higher quality, responses under C/G versus S/NG conditions.

The only inconsistency among these C/G versus S/NG comparisons occurs for ideational fluency among males, where Katz and Poag (1979) found greater fluency in the C/G group, while both Harrington (1975) and the present study found no differences. Katz and Poag (1979) suggested that this inconsistency may have been due to differences in their sample (introductory psychology students) from Harrington's broader sample of undergraduates. Since the present study also used introductory psychology subjects, this explanation does not appear to be adequate, and any potentiation of males' fluency by creativity instructions remains to be reliably demonstrated. When *quality* of responding is considered, though, there is now a consistent body of evidence that both sexes perform better on the alternate uses task when standard instructions are altered to the C/G format. The present findings go beyond this, however, in indicating several interesting features of this creativity plus guidelines combination.

Among the four combinations of instructions tested here, the C/G combination appears to be optimal in eliciting quality responses among males. For females, however, C/G instructions do not seriously hamper quality but do produce a marked decrement in quantity of alternate uses and, hence, a reduction in the absolute number of high quality ideas generated. Viewed another way, the facilitation effect of instructions to "be creative" discussed in previous investigations is only partly true for males. Although males did improve the quality of responses when instructed to "be creative" (C/NG versus S/NG and S/G), a key element for eliciting maximal performance in these individuals appears to be the addition of guidelines which define for them what a creative response is. Providing guidelines alone (S/G) or creativity instructions alone (C/NG) is not as effective as this C/G combination.

Perhaps most interestingly, the combination of creativity plus guidelines used by Katz and Poag

(1979) in effect masked the fact that instructions only to "be creative" are indeed the optimal strategy in enhancing females' task performance. The present study reveals that females produce more and better responses in the C/NG condition relative to other instructions, while, as noted above, adding guidelines to this creativity instruction hampers performance.

The present findings thus indicate that the value of alternate uses tasks as indices of creative thinking abilities can be enhanced when standard instructions are replaced by either C/G (for males) or C/NG (for females) instructions. Several questions for future research into the mechanisms responsible for these effects are also suggested.

For example, why do guidelines differentially affect the responding of males and females? Further, does this reflect the impact of performance guidelines in general, or is it due to differing interpretations of the specific criteria ("worthwhile" and "unusual") used in this and prior studies? Also, why do guidelines appear to influence attention and effort (as reflected by Digit Symbol scores) on the divergent uses problems but not on the more structured, nondivergent RAT? Pending data on these issues, the authors suggest a possible rationale for the different alternate use response patterns of males and females observed in the present study. For males, quality of responding in the C/NG condition improves over either standard instruction group, but not to the level of the C/G group. It may be that males are unsure what "creativity" means, and explicit guidelines clarify this. The general focusing of attention elicited by guidelines (higher Digit Symbol scores) versus no guidelines does suggest that such instructions may reduce task ambiguity, although they do not improve performance quality independently of creativity instructions (e.g., S/G quality is low). If creativity instructions do "disambiguate" the task for males by switching them to a verbal set (Katz & Poag, 1979), they thus do so only to a limited degree. Whether adding guidelines (CG instructions) completes this generalized "switch" to verbal responding or merely confirms that "unusual and worthwhile = creative" for males merits further research.

Females, on the other hand, seem to hold a definition of creativity which includes the guidelines criteria (C/NG quality is high). However, adding guidelines seems to slow them down, as reflected in lower fluency and poorer Digit Symbol scores relative to subjects with no guidelines. It may be, as Gall and Mendelsohn (1967) have suggested, that females seek to comply with directive instructions more than males, especially in a creativity task. If so, they might have taken more time in an effort to give fewer but better (e.g., meeting the guidelines) responses. The fact that quality under C/G or S/G did not exceed C/NG instructions suggests that this increased attention was not altogether successful, perhaps due to arousal of performance anxiety concomitant with attempts to meet the guidelines demands.

While the above speculations obviously require further study, the present investigation does indicate that seemingly subtle alterations in alternate uses instructions can improve quality of performance and, hence, the validity of this means of assessing creativity. If alternate uses, and perhaps other divergent measures of creative thinking, are to reflect subjects' maximal performance levels, more attention to the role played by task instructions is indicated.

References

- Forbach, G. B., & Evans, R. G. The Remote Associates Test as a predictor of productivity in brainstorming groups. *Applied Psychological Measurement*, 1981, 5, 333-339.
- Gall, M., & Mendelsohn, G. A. Effects of facilitating techniques and subject-experimenter interaction on creative problem solving. *Journal of Personality and Social Psychology*, 1967, 5, 211-216.
- Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.
- Harrington, D. M. Effects of the explicit instructions to "be creative" on the psychological meaning of divergent thinking test scores. *Journal of Personality*, 1975, 43, 434-454.
- Katz, A. N., & Poag, J. R. Sex differences in instructions to "be creative" on divergent and nondivergent

- test scores. *Journal of Personality*, 1979, 47, 518–530.
- Kogan, N. Creativity and sex differences. *Journal of Creative Behavior*, 1972, 6, 1–13.
- Mednick, S. A., & Mednick, M. T. *Examiner's manual: Remote Associates Test*. Boston: Houghton-Mifflin, 1967.
- Wechsler, D. *The Wechsler Adult Intelligence Scale: Manual*. New York: Psychological Corporation, 1955.
- Zimmerman, I. L., & Woo-Sam, J. M. *Clinical interpretation of the Wechsler Adult Intelligence Scale*. New York: Grune & Stratton, 1973.

Acknowledgments

The authors thank Tom Miller, Becky Hogue, and Cindy Hogue for serving as experimenters/raters and Chris Arheart for his assistance with data analysis.

Author's Address

Send requests for reprints or further information to Ronald G. Evans, Department of Psychology, Washburn University, Topeka KS 66621, U.S.A.