THE EFFECT OF PROGRAMMED INSTRUCTION ON CREATIVE PROBLEM SOLVING AND ATTITUDES *

DONALD J. TREFFINGER AND RICHARD E. RIPPLE
Purdue University Cornell University

This study examined the effectiveness of the Productive Thinking Program in developing creative thinking and problem solving abilities among 370 pupils in grades four through seven. At all four grade levels instructed pupils (n: 184) attained significantly higher scores than control pupils (n: 186) on a paper and pencil attitude measure. There were no significant indications of transfer from the programmed instructional materials to several measures of creative thinking and problem solving at any of the four grade levels. Criteria included: ten scores derived from Verbal Form A of the Torrance Tests of Creative Thinking; a General Problem Solving Test; an Arithmetic Puzzles Test and, an Arithmetic Problem Solving Test, Text Form, which represented text book number problems. Results were interpreted in terms of three general factors: the rapid administration of the programmed materials and lack of teacher involvement; differences between the format of the training materials and the criterion measures, which may have left the pupils unable to successfully apply the abilities developed; and, criterion difficulty.

Recent interest among psychologists and educators in creative thinking has led to the development of a number of programmes which seek to develop pupils' creative abilities. These have included eliciting uncommon responses (9, 15), 'brainstorming' (17) and workbook exercises (16). One of the most intriguing approaches has been the attempt to facilitate creative development through a programmed instructional sequence called the Productive Thinking Program (5). The materials for this programme, described in greater detail elsewhere (6, 8), consist of sixteen 'mystery problem' booklets. Pupils follow the progress of typical fifth and sixth grade youngsters through each of the problem sets. The materials are designed to develop a number of skills, complex abilities, and attitudes which are independent of traditional subject-matter content, but which should lead to positive transfer to a variety of problem tasks (7, 8).

In a study reported by Crutchfield and Covington (8), fifth and sixth grade pupils studied a thirteen lesson preliminary version of the pro-

*The research herein reported was supported by the United States Office of Education, Project No. 8-B-002, Grant No. OEG-0-8-08002-0220-010, and by the Finger Lakes Regional Supplementary Educational Center.
gramme. The instructed pupils markedly outperformed uninstructed pupils on several creative problem solving criteria presented in a format similar to that of the training materials. In a second study involving the full set of sixteen booklets, results again indicated marked superiority of instructed pupils over controls. The facilitative effects of the programme were found to be greater at the fifth than at the sixth grade level.

Ripple and Dacey (21), using a ten lesson version of the instructional programme with eight grade pupils, found no significant differences between instructed and control pupils on four measures of verbal creativity. Instructed pupils, however did solve the Maier Two-string problem significantly faster than controls. The authors concluded that their results supported 'non-specific transfer' from the instructional materials to a problem-solving criterion. They also noted that it might be necessary to provide more challenging materials to optimize the programme's effectiveness as grade level increases. Two further studies concern fifth grade pupils. In one of them, instructed pupils showed performance superior to uninstructed pupils on problem solving tasks presented in programmed booklet format. A large number of instructed pupils, however, did not show performance which differed from that of control pupils. The findings prompted the investigator to suggest the possibility of the need for additional supplementary practice or revision of part of the programme (3). In the other study (18), although instructed pupils' performance surpassed that of control pupils on 30 of the 40 criterion measures, only eleven of those differences reached significance and those tended to be on measures similar to the training materials rather than on more diverse tasks. There were no significant differences, for example, on verbal creativity test scores. The magnitude of the effects of instruction were considerably less than in the studies conducted by Crutchfield and Covington.

Thus, despite initial reports demonstrating pronounced effects of instruction on pupils' problem solving abilities, the findings of several recent studies have been less conclusive, and have failed to find transfer from the instructional materials to pupil performance on tests of creative thinking. The present study was addressed, therefore, to three main issues. First of all, no systematic investigation of the effectiveness of the programmed sequence at several grade levels has been undertaken, though two studies have suggested the possibility of decreasing effectiveness (8, 21). The present study included comparisons of pupils at four grade levels (four through seven). Secondly, the effect of the instructional materials on pupil performance on tests of creative thinking is not clearly understood.
The present study investigated transfer from the instructional programme to performance on the Torrance Tests of Creative Thinking with pretest scores and IQ covaried. Thirdly, there have been no reports of empirical support for the assumption that the generalized skills, abilities, and attitudes which the programme proposes to develop will transfer to problem solving tasks involving traditional subject-matter content. In addition, there has been only limited evidence suggesting that the instructional materials facilitate performance on problem solving tasks which are presented in a format dissimilar from that of the training materials. Thus, the present study employed a number of creative problem solving criteria, designed to test the occurrence and extent of non-specific transfer of learning.

**METHOD**

**Subjects**

Three hundred and seventy pupils from 16 classes in six public school systems were selected for the study. There were four classes at each grade level (four through seven). Two classes at each grade level were randomly assigned to the instructional condition; the remaining classes served as controls. The composition of all classes was similar with respect to measurements of socio-economic status, sex and intelligence distributions, and proportions of pupils with reading deficiencies.

**Procedure**

Classes in grades four through six were selected and assigned to conditions in October 1967. Seventh-grade classes were selected and assigned in February 1968. After assignment of classes to conditions, but prior to the instructional period, all pupils were pre-tested on verbal creativity and arithmetic computation skills. Immediately after the pre-testing, instructional condition classes began work with the *Productive Thinking Program*. The instructional materials were administered at the rate of one booklet per day for 16 consecutive school days. Pupils worked independently with the materials. Control classes received only their ordinary classroom instruction. Final testing of all pupils was conducted immediately following the last day of instruction for experimental classes. In grades four through six, post-tests were given during December 1967; in grade seven, during March 1968.
Variables investigated

Several tests were used to assess the effects of the instructional materials on verbal creativity, general problem solving, and arithmetic problem solving. A number of these measures were constructed by the writers for the study, and have been described in detail elsewhere (25).

Arithmetic Skills. The pretest consisted of 50 items, each scored on a right-wrong basis, emphasizing computational skills with fundamental arithmetic operations. The pupils' score was the total number of items correctly answered. The Kuder-Richardson Formula 20 reliability coefficient was .93.

Verbal Creativity. Six sub-tests for the Torrance Test of Creative Thinking, Research Edition, Verbal Form A (24) were given as pre- and post-tests. The battery was presented in paper-pencil format to classroom groups, by trained examiners. The six sub-tests used were: three Ask and Guess Tests (asking Questions. Guessing Causes, and Guessing Consequences), Unusual Use of cardboard boxes, Product Improvement of a toy elephant, and the Just Suppose task (that clouds had long strings). Inter-scorer reliabilities ranged from .73 to .97. Test-rating reliabilities for standardized total scores, at a three week interval, ranged from .52 to .69. Although the Torrance tests have been subject to a good deal of recent criticism (3, 27, 28), it was the authors' position that they represent the best available operational specifications of abilities which comprise creative thinking for pupils in elementary school. It is clear that divergent thinking, measured by Guilford or Torrance tests, does not comprehensively assess 'creativity,' and that these tests are accompanied by a number of technical, procedural difficulties. Nevertheless, the tests do seem to measure abilities which constitute an important sampling of creative thinking abilities or of the abilities included in creative problem solving (cf. 14). That is, the divergent thinking functions of fluency, flexibility, and originality are viewed by the present authors as important components of creative problem solving. Further, for use with classroom groups of elementary school pupils, the Torrance tests are the most practically useful instruments presently available.

General Problem Solving. The test consisted of eight problems, presented in a single session, in paper and pencil format. Problems included both Type O and C tasks, following the distinctions proposed by Davis (10). In Type O problems, outcomes of response alternatives are unknown to the subject and so he must engage in overt trial-and-error behaviour. In Type C problems, on the other hand, the subject can predict the outcomes of his
response alternatives, and so his trial-and-error behaviour may be covert. Each problem in the present study was scored independently, and total scores were derived only for Type O and C problems.

Arithmetic Puzzles. The test consisted of ten problems, each generally considered useful for ‘classroom enrichment’ in mathematics. Appropriateness of the tasks was assessed following the recommendations of Spitzer (23). The test was taken by half the pupils in each class, randomly selected: the remaining pupils took the Text Problems test, described below. The results from this test must be viewed with caution, since the Kuder-Richardson Formula 20 reliability coefficient was lower than desirable for research purposes (r = .49).

Arithmetic Problem Solving (text problems). The test consisted of sixteen word problems, selected to represent the kinds of number problems typically utilized in elementary arithmetic instruction. The Kuder-Richardson Formula 20 reliability coefficient was .83.

Attitude. The attitude inventory was a preliminary version of the Childhood Attitude Inventory for Problem Solving (2). In the first of two parts the pupil agreed-disagreed (yes-no) with 30 statements about creative thinking and problem solving. In the second part, 22 statements in the same format were presented to assess pupils’ self-confidence in engaging in creative problem solving. Reported test-retest reliabilities averaged .69 for Part One and .65 for Part Two, at five week intervals, The reported Kuder-Richardson Formula 20 reliability coefficients were .93 for Part One, and .86 for Part Two (2).

Treatment of the Data

Verbal Creativity. Each of the six sub-tests was scored for either fluency, flexibility, or originality; Asking Questions and Unusual Uses were scored for flexibility, Guessing Causes and Product Improvement for originality and Guessing Consequences and Just Suppose for fluency. For purposes of analysis, ten scores were used. First, each of the six sub-test scores was used separately. Next, three combined scores were used (fluency, flexibility, and originality). Finally, a composite total score was obtained. This was the average of the six sub-test scores, after each had been standardized by grade level with $M$: 50 and $SD$: 10. For each of these ten sets of post-test scores, instructed and control pupils’ scores were compared using one-way analysis of covariance (12). Two such analyses were conducted: in the first, pretest verbal creativity scores were covaried; in the second, both
pretest scores and IQ were used as covariates. Thus, twenty comparisons between instructed and control pupils' scores were made at each of the four grade levels.

General Problem Solving. For this test, no single total score was computed. Problems following the type C format or 'insight'-type problems, were scored on a pass/fail basis. Type O problems were scored continuously (one point for each appropriate response). Instructed and control pupils' scores were compared in five analyses; separate analyses were conducted at each of the four grade levels. In the first analysis, the proportions of pupils in each condition correctly solving each of the dichotomously-scored problems were compared, using a two-by-two Chi-square analysis (13). Next, a two-by-two Chi-square test was used to compare the proportions of pupils in each condition giving acceptable responses to an unsolvable problem (also dichotomously scored). Third, scores of pupils in both conditions on each of the continuously-scored problems were compared, using one-way analysis of covariance with IQ score as covariate. Fourth, pupils in both conditions were compared on the total number of solutions produced for dichotomously-scored problems, using one-way analysis of covariance. IQ was used as the covariate in this analysis. Finally, scores on the continuously-scored problems were standardized to a mean of 50 and a standard deviation of ten (at each grade level), then averaged to obtain a mean total score. The mean total scores for these problems of pupils in both conditions were then compared, using one-way analysis of covariance, covarying to equate the groups with respect to IQ.

Arithmetic Puzzles Test. Each of the ten items in this test was scored dichotomously (pass/fail). Two sets of analyses were conducted. First, the proportion of pupils in both conditions correctly solving each problem was compared using a two-by-two Chi-square test. Next, the total number of solutions achieved by pupils in both conditions was compared, using one-way analysis of covariance. Arithmetic pre-test scores and IQ were used as multiple covariates.

Arithmetic Problem Solving text problems. For this test, a single score was obtained, which indicated the total number of the sixteen problems which the pupil had correctly solved. Scores of pupils in both conditions were compared using one-way analysis of covariance, covarying both IQ and Arithmetic Pre-Test scores.

Attitude Inventory. At each grade level, scores of instructional and control pupils were compared using one-way analysis of variance. Three
analyses were conducted: first, total number of appropriate responses for Part I; next, total number of appropriate responses for Part II; and, a total score, based on the sum of scores from Parts I and II.

RESULTS

Verbal Creativity

Of the 80 analyses (20 at each grade level), there were five significant differences ($p < .05$) between the instructional and control groups' adjusted means. Of these, three favoured the instructed group. Of the remaining 75 comparisons, the adjusted means for instructed pupils exceeded the controls' in slightly fewer than half the comparisons. Because of the number of analyses performed, and in view of the fact that one might expect as many as four differences of eighty to reach significance as a chance occurrence, these data have not been summarized in tabular form.

General Problem Solving

Ten comparisons were made at each grade level. They were: (a) performance on each of the eight problems in the test; (b) total score on Type C problems; (c) total score on Type O problems. Of the total of 40 analyses, there were five significant differences between treatment groups, of which three favoured the instructed pupils. In grade four, instructed pupils’ mean scores were significantly greater than control pupils’ only on problem three (a paper-and pencil anagrams task). In grade five, instructional pupils’ mean scores were significantly greater than controls’ on problem one, a modified form of the Pea Problem (20), and on problem three. In grades six and seven, there were no significant differences between groups on any of the ten scores analyzed.

Arithmatic Puzzles Test

There were no significant differences between instructed and control pupils' scores on this test for comparisons by problem and by total score at any of the four grade levels.

Arithmetic Problems (text form)

There were no significant differences between the two conditions at any of the four grade levels.
Table 1 summarizes the results of one-way analysis of variance carried out on scores of instructional and control group pupils on the Childhood Attitude Inventory for Problem Solving. At all four grade levels instructed pupils' mean scores were significantly higher than control pupils' mean scores for Part I (general attitudes about creativity and problem solving) and for Total Score. There were no significant differences between treatment groups at any of the four levels on Part II (expressions of self-confidence in creative thinking and problem solving).

DISCUSSION

Our findings lead to the conclusion that the instructional materials did not influence pupils' verbal creativity scores to any appreciable extent. These results are consistent with previous findings (18, 21). Covington (3)
has argued that tests of verbal creativity such as those used in these
studies may not be adequate instruments for assessing complex cognitive
skills. Torrance (24) however, has argued, that these tasks measure a
variety of intellectual abilities which are fundamental to the creative
thinking process. If the test tasks are viewed as attempts to assess basic or
component abilities which are required for success in more complex
creative problem solving situations, they seem to be appropriate criteria
for inclusion in any evaluation of the effectiveness of the instructional
materials. The present results lend no support to the suggestion advanced
by Ripple and Dacey (21) that the instructional materials may become less
effective in developing pupils’ verbal creative thinking abilities as grade
level increases.

Findings from the general problem solving tasks provide no general
support for the effectiveness of the materials as they were used in this study.
Neither did there appear to be a systematic relationship between grade
level and the effectiveness of the instructional materials. It should be
noted, however, that at the fifth grade level, eight of ten comparisons
favoured instructed pupils, more than at any other grade level. In view of
the absence of significant differences, however, one cannot make inferences
from such observations with any confidence.

Several interpretations of these data merit consideration. First, it is
important to point out that, in accord with the description of these materials
as an ‘auto-instructional programme’ they were used in this study without
any supplementary teacher participation. In addition, the materials were
presented to pupils at a very rapid pace: one lesson on each of sixteen
consecutive days. Olton et al (18) have suggested that such ‘severe’ pro­
cedures may limit the potential effectiveness of the programmed materials.
There is also evidence from other sources that active teacher participation
may facilitate learning even from programmed instructional materials (1,
22). It may also be important to note that the problem solving criteria used
in this study were measures of products rather than of processes, and were
presented in a form not at all similar to the training materials. Even as a
test of non-specific transfer of learning, they must be considered a rigorous
test. In addition, at all grade levels there was a generally high level of
difficulty for all the problems which may have had the effect of ‘masking’
differences between treatment groups.

With regard to the differences between the format of the instructional
materials and the problem solving criteria, some recent research (26) is
suggestive. In this research, trained subjects became ‘notably unsuccess­
ful in finding the shortcut solution’ as criterion problems became increas-
ingly dissimilar from practice or training problems. It seems that 'limited' educational exposure to problem solving approaches may induce students to adopt the strategy to search when confronted with transfer situations, but leave them lacking the skill to successfully apply the strategy' (26).

There were no indications of significant effects of the instructional materials on pupil performance for either of the Arithmetic Problem Solving tests at any of the four grade levels. One factor which may be important in examining these results is the relevance of the problem solving criteria. One's first thought in examining the tests is that the problems presented are quite different in nature from the kinds of problems treated in the instructional sequence. But to dismiss the results on that basis will not hold under more critical examination. What are the skills and abilities required to engage in mathematical problem solving? Historically, at least, these abilities have been described in terms almost identical with those used to describe problem solving in other domains (cf. for example, 11 and 19). Further, the authors of the programmed materials proposed that it develops highly generalized abilities which will be transferable to a number of subject-matter content areas. Crutchfield (7) has argued specifically that the instructional materials should succeed in developing problem solving abilities which generalize to types of test materials quite different from the training problems. The arithmetic problems used in this study seemed to be appropriate samples of elementary level problem solving tasks in this content area. Thus, while the tests may have been rigorous or severe, they were not, in our judgment, irrelevant. Certainly, the same factors may have influenced our results on these measures as we considered for general problem solving: lack of teacher involvement, rapid presentation, dissimilarity of format between training materials and tests; and criterion difficulty. The generally low scores for all pupils on both Arithmetic Problem Solving Tests suggests that the difficulty of the problems may have been an important factor influencing the results. Also, the reliability of the Arithmetic Puzzles Test is lower than desirable for research purposes.

On the Childhood Attitude Inventory for Problem Solving there were consistent significant differences between instructed and control pupils. Instructed pupils' scores were significantly higher than Controls' on Part I, which dealt with general attitudes about creative thinking and problem solving, and on total score, at all grade levels. This result appears to lend support to the earlier interpretation that instructed pupils may have been influenced by the instructional materials, but were unable to apply what was learned when confronted with the transfer tests of problem solving and creative thinking which differed considerably from the training
materials. Although it may be tempting to suggest that the differences on the attitude inventory can be attributed to a tendency on the part of the pupils participating in the study to give 'desirable' responses, there is no evidence to warrant such an assertion. Since both instructed and control pupils were tested by trained personnel, rather than by classroom teachers, and pre- and post-tests were in no way associated with the delivery or utilization of the instructional materials, there seems to be no sound basis for predicting a differential effect on instructed and control pupils' responses to the attitude measures.

CONCLUSION

Three general conclusions emerge from the present study. Firstly, there was no indication of a general pattern of differential effectiveness of the instructional materials among pupils in grades four through seven. Secondly, there was no evidence for the effectiveness of the instructional materials, as here utilized, in developing pupils' verbal creative thinking abilities when IQ scores and pre-test scores were controlled by covariance procedures. Thirdly, with respect to other problem solving criteria, there was very little evidence for the effectiveness of the instructional materials as used in this study. There were significant differences favouring instructed pupils on Part One (General Attitudes) and Total Score on the Childhood Attitude Inventory for Problem Solving. There were no consistent indications that the instructional materials influenced pupils' scores on the several problems in the General Problem Solving Test. There was no support for the assumption of positive transfer from the programmed materials to Arithmetic Problem Solving for either of the two forms of the test at any of the four grade levels studied.

Two implications for future research emerge very clearly from these results. First, several procedural questions require systematic investigation. These include rate of presentation of the programmed lessons, nature and extent of teacher participation, and the utilization of supplementary exercises. Secondly, the relation between similarity of the instructional materials and the criterion measures in influencing the pupils' ability to apply what has been learned warrants systematic study. The results of this study, viewed in relation to previous reports, suggest, that as problem solving criteria become increasingly dissimilar from the format of the training materials, pupils may become notably less successful at applying their training.
REFERENCES


