DOCUMENT RESUME

ED 021 312 24 EA 001 341

By- Olton, Robert M.; And Others

THE DEVELOPMENT OF PRODUCTIVE THINKING SKILLS IN FIFTH-GRADE CHILDREN.

Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
Spons Agency- Office of Education (DHEW), Washington, D.C. Bureau of Research.
Report No- TR-34
Bureau No- BR-5-0216
Pub Date Nov 67
Contract- OEC-5-10-154
Note- 47p.
EDRS Price MF-$0.25 HC-$1.96
Descriptors- CLASSROOM ENVIRONMENT, COGNITIVE DEVELOPMENT, *CREATIVE THINKING, EDUCATIONAL EXPERIMENTS, *GRADE 5, INSTRUCTIONAL MATERIALS, INTELLIGENCE QUOTIENT, PROBLEM SOLVING, *PRODUCTIVE THINKING, *PROGRAMED INSTRUCTION, *PROGRAMED MATERIALS, SEX (CHARACTERISTICS), STATISTICAL ANALYSIS, STUDENTS, TEACHER PARTICIPATION

This study investigated (1) the extent to which thinking and problem-solving of fifth-grade students could be improved by the use of self-instructional programed lessons (the productive thinking program), (2) the relationship between productive thinking abilities and IQ-sex characteristics of the learner, and (3) the relationship between level of productive thinking performance and the extent to which overall classroom environment was judged to facilitate creative thinking. Results from 44 fifth-grade classes in Racine, Wisconsin, involving a total of 704 students, showed statistically significant increments in thinking and problem-solving performance on a wide variety of productive thinking measures. These instructional benefits occurred for virtually all types of students, regardless of sex or general IQ level, and were especially marked for students in classrooms having environments which were judged to provide relatively little support and encouragement for the development of productive thinking. Apart from the effects of the instructional materials, performance on the productive thinking measures used in this study was significantly related to sex (girls generally scored higher than boys) and showed a strong and positive relation to IQ. EA 001 340 is a related document. (Authors/JK)

THE DEVELOPMENT OF PRODUCTIVE THINKING SKILLS IN FIFTH-GRADE CHILDREN

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR COGNITIVE LEARNING

RESEARCH
COGNITIVE LEARNING
DEVELOPMENT

ERIC

Technical Report No. 34

THE DEVELOPMENT OF PRODUCTIVE THINKING SKILLS IN FIFTH-GRADE CHILDREN

Robert M. Olton, James L. Wardrop,
Martin V. Covington, William L. Goodwin,
Richard S. Crutchfield, Herbert J. Klausmeier, and Teckla Ronda

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1967

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Martin Covington, Richard Crutchfield, and Lillian Davies TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education and Welfare, under the provisions of the Cooperative Research Program Center No. C–03/Contract OE 5–10–154), and with additional support of a grant from the Carnegie Corporation of New York to the Creative Thinking Project, Institute of Personality Assessment and Research, University of California. Berkeley.

GPO 806–459–2

PREFACE

High priority is now being given to the improvement of children's learning in the cognitive domain, particularly in connection with productive thinking skills. The federal government, various profit and non-profit agencies, and many individuals are attempting to extend knowledge about this aspect of human learning and to improve related educational practices through programs of research and development. Behavioral scientists and school people are joining together in an attempt to learn more about the nature of productive thinking skills and the conditions associated with more efficient learning of them. This report of a large-scale controlled experiment involving personnel from three different organizations shows how specialized knowledge and resources can be brought to bear upon extending knowledge and improving educational practice through research and development activities. Personnel participating in the study were from the Wisconsin Research and Development Center for Cognitive Learning, the Creative Thinking Project of the Institute of Personality Assessment and Research at the University of California, Berkeley, and the Racine (Wisconsin) Unified School District, No. 1.

The instructional materials used in the experiment—The Productive Thinking Program, Series One: General Problem Solving—were developed by Professor Martin Covington, Professor Richard Crutchfield, and Mrs. Lillian Davies of the Creative Thinking Project. Dr. Robert Olton, of the same project, was chiefly responsible for the development and subsequent scoring of the large battery of productive thinking tests used in the experiment and wrote the first draft of the major portion of this report. Both the instructional and test materials are copyrighted; the instructional materials are now available for regular school use. Professor Covington and Dr. Olton, in addition to their role in the development of the new materials, worked directly with the staff of the Wisconsin R & D Center in planning, executing, and reporting the experiment.

Dr. William Goodwin of the Wisconsin R & D Center took primary responsibility in working out the design of the experiment with Dr. Olton. Dr. Goodwin coordinated the various components of the experiment, making sure that the pretesting, the instruction, and the posttesting were carried out according to schedule during the second semester of the 1965—1966 school year. Dr. James Wardrop of the Wisconsin R & D Center was responsible for the computer analysis of the data during the 1966—1967 school year, and for writing the first draft of the results during the spring and summer of 1967. Professor Herbert J. Klausmeier of the Wisconsin R & D Center initially consulted with Professor Crutchfield regarding the experiment in the summer of 1965 and subsequently developed and directed the strategy for bringing the resources of the three organizations to bear most effectively upon execution of the experiment. He consulted with relevant personnel on all matters regarding the design and conduct of the experiment, per se. He and Professor Crutchfield defined the more precise roles of the staff and the contribution of their two respective organizations; he and Mr. Harris Russell of the Racine schools similarly defined the roles and contributions of their two respective organizations.

Mr. Harris Russell, Director of Instruction, Racine Unified School District No. 1, made the necessary arrangements for his central staff, building principals, and teachers first to decide whether to participate in the experiment and, subsequently, to carry it out. Miss Teckla Ronda of the Racine schools estimated the extent to which each participating classroom represented a favorable creative environment, made sure that materials were received by teachers and used according to the plan outlined, and secured relevant information regarding each child, including IQ score, educational achievement test scores, socioeconomic status, and sex. The elementary school principals and fifth-grade teachers gave generously of their time, and also offered many useful ideas concerning the experiment.

Through this union of personnel and resources from the three organizations, answers to the following questions were secured.

1. To what extent can specified productive thinking abilities be taught to elementary school children through the use of programed instructional material designed to teach thinking skills independently of the content of any specific subject field ?
2. Are some productive thinking skills more amenable to instruction than others ?
3. Can productive thinking skills be learned by children of all IQ levels ?
4. Are there consistent sex differences in productive thinking skills ?
5. How are productive thinking skills affected by a classroom environment judged to facilitate creative expression by the students ?

The reader will find this report interesting from two viewpoints. First, it is a clear account of the development of productive thinking skills in school children; answers to the preceding questions were obtained and are more definitive than in most experiments. Second, the execution of the experiment, including the placement of classes into experimental and control groups, the type of information gathered, the execution of the treatments, and the analysis of the data, shows that useful information can be obtained through carefully conducted research in school settings.

Herbert J. Klausmeier
Professor of Educational Psychology
Director, Wisconsin Research and Development Center for Cognitive Learning

Richard S. Crutchfield
Professor of Psychology
Director, Creative Thinking Project of the Institute of Personality Assessment and Research, University of California, Berkeley.

ERIC

CONTENTS

LIST OF TABLES

List of Tables (cont'd)

ABSTRACT

The major purpose of this study was to investigate the extent to which increments in the thinking and problem-solving performance of fifth-grade students could be produced by the use of self-instructional programed lessons (The Productive Thinking Program, Series One: General Problem Solving) which were designed to teach skills and strategies of creative thinking independent of any specific subject field. The study also investigated the relationship between productive thinking abilities and certain characteristics of the learner (IQ and sex), and the relationship between level of productive thinking performance and the extent to which overall classroom "environment" was judged to facilitate creative thinking.

Results from 44 fifth-grade classes showed that The Productive Thinking Program produced statistically significant increments in thinking and problem-solving performance on a wide variety of productive thinking measures. These instructional benefits occurred for virtually all types of students (regardless of sex or general level of IQ), and were especially marked for students in classrooms having environments which were judged to provide relatively little support and encouragement for the development of productive thinking. These effects were obtained when the materials were used as an entirely self-contained, self-instructional program; considerably greater educational benefits could be expected under conditions where the materials are reinforced by active teacher participation.

Quite apart from the effects of the instructional materials, performance on the productive thinking measures used in this study was significantly related to sex (girls generally scoring higher than boys) and showed a strong and positive relation to IQ.

I

RATIONALE OF THE STUDY

BACKGROUND

During the past five years, the Creative Thinking Project has been developing a set of special instructional materials designed to develop and facilitate creative thinking in students of the upper grades of elementary school. These materials (described in Section II) are a series of self-administering programed lessons which require a total of about 16 classroom hours for completion. The materials have undergone extensive pilot development and revision and were used in two recent experimental studies involving several schools in the San Francisco Bay area. In both of these studies, the materials produced a substantial increase in the students' creative thinking performance (Covington and Crutchfield, 1965; Crutchfield, 1966) and were very favorably received by both students and teachers. After the second San Francisco Bay area study, the materials underwent a complete revision, and the desirability of a large-scale experimental study with the revised materials (titled The Productive Thinking Program, Series One: General Problem Solving) became apparent.

Researchers at the University of Wisconsin have also studied creativity in school children, particularly in gifted children (Klausmeier, Harris, and Ethnathios, 1962; Klausmeier and Wiersma, 1964), and have organized programs for the gifted which included education for creativity (Klausmeier, 1963; Klausmeier and Teel, 1964). In addition, the focus of the Wisconsin Research and Development Center for Cognitive Learning is on improving education through a better understanding of cognitive learning, including both creativity and problem-solving processes.

Thus, the directors of the Creative Thinking Project and the Wisconsin R & D Center agreed that an outstanding research opportunity could be realized by pooling some of the specialized resources and capabilities of each group in a large-scale experimental study. The officials of the Racine (Wisconsin) Unified School District endorsed the study enthusiastically since it offered the possibility of extending the program of creativity to a large segment of the entire school population.

AIMS OF THE STUDY

The major purpose of this study was to investigate the extent to which increments in the thinking and problem-solving performance of fifth-grade students could be produced by the use of programed lessons (Series One of The Productive Thinking Program) which were designed to teach skills and strategies of creative thinking independent of any specific subject field.

Within this framework, several specific aims were incorporated:

(1) To investigate further the extent to which productive thinking (i.e., creative thinking and problem solving) abilities can be taught to elementary school children independent of any specific subject field. Currently, there is strong emphasis in the schools upon specific subject matter areas, with little attention being given to teaching general skills which transcend specific curricula. If such cognitive skills can be taught successfully, they may be applicable in a variety of subject areas and should enhance performance in these areas.

(2) To discover whether some of these skills or behaviors are more readily taught than others. For example, one aspect of problem solving is the asking of appropriate questions. Can students be taught this behavior more or less successfully than they can some other skill, such as generating a set of possible solutions to a problem?

(3) To investigate the relationship between certain learner characteristics (intelligence and sex) and performance on a broad diversity of productive thinking measures. It is generally reported that girls perform slightly better than boys on most measures of academic achievement. Is this also true of measures of productive

thinking? It has also been reported that intelligence and "creativity" are relatively independent. Does this independence persist throughout several different aspects of productive thinking and on several different types of problems? In addition, is there a relation between intelligence or sex and the ability to learn productive thinking skills?

(4) To investigate extent to which overall classroom "environment" facilitates productive thinking in students, and the extent to which students in different types of classroom environments profit from instruction in productive thinking. To what extent does merely providing an encouraging environment promote creative thinking? Are instructional materials more (or less) effective in such environments?

(5) To probe some of the instructional limits of Series One of The Productive Thinking Program by using these materials as an entirely self-contained program of instruction, with all forms of teacher participation purposely held to a minimum—in much the same way that a teacher might assign a program of independent reading which a particular student or group of students would be expected to do on their own, with a minimum of supervision. Moreover, the entire program was used in a compressed period of time at the rate of virtually one lesson per day. Thus, this study investigated the effectiveness of these materials when they were used under the most demanding conditions likely to be encountered in a school setting. (In addition to being valuable in a practical sense, the results would also become an important component of a systematic research effort to explore the effectiveness of materials such as these under a variety of instructional conditions.)

II
METHODS

SUBJECTS AND CONDITIONS

Subjects for this experiment were the students in 44 of the 47 fifth-grade classrooms in the Racine, Wisconsin, Unified School District No. 1. All students in these 44 classes participated in the experiment, but analysis is based upon a sample of eight males and eight females selected from each class. Selection was accomplished by rank-ordering all male students from a given classroom on the basis of IQ, and then randomly selecting one student from each eighth of the distribution. Eight females were then selected from the same class in the same way. This procedure was continued for all 44 classes. The entire analysis involved a total of 704 students. Selection was made after the experiment had been completed but before any scoring of tests was begun.

This large sample reflected, in proper proportion, the wide variety of abilities and backgrounds characteristic of students in an urban school system (a city of about 100,000 population) and hence provided a thoroughly representative base for analyzing both overall and internal effects. The size of the sample also permitted a very stable estimate of error and made possible the use of classroom means (rather than individual student scores) as the unit of analysis, thus reducing sporadic variation and also meeting the statistical requirements of random assignment.

Several weeks before the experiment began, the teaching supervisor of the district observed each class while the teacher and students were having an extended discussion about an interesting event in social studies, an event which provided an opportunity for free, creative interchange between teacher and students. All classes discussed the same event, and on the basis of her observations, the supervisor recorded the extent to which the overall classroom atmosphere (including the teacher) seemed to provide an environment that facilitated and fostered creative thinking on the part of the students. The supervisor reported that she felt these ratings were representative of the normal state of affairs that existed in the various classrooms and were not an artifact of the particular discussion situation in which they were obtained. The final, single, Environment rating assigned to each class was a composite score derived from the several objective ratings the supervisor made.

Classrooms were rank-ordered according to the composite environment rating. The twenty-two classrooms with the highest ratings were designated as Facilitative environment classrooms; likewise, those 22 classrooms whose composite rating was below the median were classified as Nonfacilitative environment groups.

Within each of these Environment blocks, half the classrooms were randomly assigned to the Treatment group (which would use *The Productive Thinking Program*) and the remaining half constituted the Control group. Thus, the classroom population was subdivided as follows:

		Number of Classes
Facilitative	Treatment	11
	Control	11
Nonfacilitative	Treatment	11
	Control	11

INSTRUCTIONAL MATERIALS

The Productive Thinking Program (Covington, Crutchfield, and Davies, 1966) is a set of instructional materials designed to develop and strengthen the productive thinking skills of students in the upper grades of elementary school. The Program undertakes to develop the student's ability to use important *skills and strategies* of thinking and problem solving, seeks to enhance his *understanding* of the processes of his own thinking, and aims to promote favorable *attitudes and motivations* toward activities which involve the use of the mind.

Series One of this Program, which deals with General Problem Solving, served as the instructional materials in this study.

The skills taught by Series One of The Productive Thinking Program are those which are essential for the kind of complex and original thinking in which an individual strives to produce new ideas, fresh insights and understandings, and his own solutions to problems. Hence the Program teaches the student how to generate many ideas, especially clever and unusual ones, how to look at a problem in a different and more fruitful way if he gets "stuck," and how to use the implications of crucial facts or events to gain new perspective on a problem which seems to resist solution. Moreover, although productive thinking is not entirely a matter of logical reasoning, it is obvious that reasoning and planfulness play an important role in this kind of thinking. Consequently, the Program also shows the student how to proceed in an organized and systematic way when attacking a problem, how to clarify the essentials of a problem, and how to pay attention to relevant facts and conditions of the problem in evaluating one's ideas.

In addition to teaching certain essential thinking skills, the Program seeks to promote those motivations and attitudes that are required for effective performance in productive thinking tasks. It attempts to increase the student's interest in and liking for activities which involve the use of the mind, to build up his readiness to work on thinking tasks in a persistent and concentrated way, and to strengthen his self-confidence in his ability to think.

Description of the Materials

Series One of The Productive Thinking Program consists of 16 programed lessons, each of which is an individual booklet about 40 pages long. The lessons have a cartoon-text format which serves to heighten interest and at the same time permits clarity and emphasis in the presentation of the material. (Four sample pages from one of the lessons are presented in the Appendix.) The lessons are written primarily for fifth- and sixth-grade students and can be used by students with a wide range of intelligence within these grades. Each lesson is self-administering, permitting the student to work at his own pace; virtually all students complete a lesson within 50 minutes.

Each lesson is centered about a complex and engaging problem, presented in story form, which the student attempts to solve. The problem features a mysterious occurrence, and as the student works on this mystery, he is called upon to generate ideas, to look for and evaluate the meaning of puzzling facts or discrepancies in information, to ask questions, and to practice other productive thinking skills.

On certain pages, the student writes out his questions and ideas. Feedback to his responses is then provided in following pages, where he finds a range of illustrative ideas appropriate to the problem at this point. In this way, he is led to understand what constitute relevant, fruitful, and original ideas, and his readiness to generate such ideas is enhanced. Each lesson is so designed that the student works gradually toward a solution to the problem and eventually is brought to discover it by himself. Thus, he experiences the pleasurable excitement which accompanies successful problem solution and acquires a growing sense of competence in coping with difficult and complex problems.

In addition to giving the student guided practice in using productive thinking skills, the lessons also give some direct, didactic instruction about thinking and about each of the skills which the student is being taught. This direct instruction is incorporated into each lesson as the central problem develops and as the student works toward solution. In addition, the major points taught in each lesson are summarized for the student at the end of that lesson.

The human element in this series of lessons is provided by a continuous story-line built around the Cannon family of Elmtown, USA. Lila and Jim Cannon are in the fifth grade and sixth grade, respectively, and their Uncle John is a high school science teacher who occasionally does some high-level detective work for the Elmtown police. Uncle John teaches Jim and Lila how to become better thinkers as they work on various "cases" together.

Jim and Lila are intended as counterparts of the students who work on these lessons. They are typical children with whom the student can identify—likeable and human. Their thinking provides examples of how to attack problems, and—in some cases—how not to attack them. Jim and Lila are intended to serve as stimulating "companions" for the reader as he pursues these intellectual adventures, and in many cases the student works on the problems in concert with Jim and Lila—first the student generating his own ideas and questions, then Jim and Lila responding with theirs.

Administration of the Materials

The 16-lesson series requires a total of 16 classroom hours of instruction time. In the present study, the lessons were administered one per day for 4 days of each week; the remaining school day in each week was used for makeups. Generally, each student worked individually at his own pace and had an assignment to which he turned when he had completed the day's lesson.

It was the intention of this study that the teacher's role in the use of these materials be held to a minimum so that any posttest superiority of students who had used the materials (as opposed to the controls, who had not) would result from the effectiveness of the materials themselves, rather than from whatever effects might occur when they served as raw material for imaginative treatment by capable teachers. Thus, the teachers were instructed only to distribute the lessons and answer procedural questions. They were not to conduct class discussions based on the materials or otherwise participate in the way they were used. One exception to this policy was permitted in classes with a number of poor readers. In these classes, teachers often read portions of the lessons aloud with small groups, or otherwise assisted students with substantial reading problems.

MEASURING INSTRUMENTS AND SCORING PROCEDURES

Most of the measuring instruments used in this study were developed by the Creative Thinking Project at Berkeley.[1] Since they differ from traditional educational tests in both content and structure, it is helpful to present a brief description of their underlying rationale before describing the instruments themselves. A more complete discussion of this rationale may be found in Covington (1968).

Briefly, the instruments are based on the proposition that creative thinking can be conceptualized as an intricate and highly complex problem-solving process; one which requires the individual to use and to coordinate a number of different cognitive skills as he seeks to reach his goal, whether this goal is to create an insightful explanation, to design an effective course of action, or to produce a tangible product. This complex problem-solving process involves: (1) the use of a <u>number and variety of cognitive skills</u>, such as generating ideas, avoiding premature judgment, and breaking mental "sets" in order to look at a problem in a new and different way; (2) the effective <u>management</u> and deployment of the various cognitive skills by what might be called a "master" thinking skill. For example, this Master Thinking Skill would be used to determine, at a particular point in the problem-solving sequence, whether it would be more fruitful to suspend critical judgment and give free reign to speculation and fantasy in a search for entirely new ideas, or whether, in contrast, it would be more fruitful to examine systematically and critically the ideas one already has; and (3) a high degree of <u>involvement</u> or dedication on the part of the thinker—the individual is highly engrossed with the creative task and finds it meaningful and engaging.

Hence the measuring instruments used in this study attempt to present the student with thinking tasks which he will find meaningful and engaging, and which require him to use and to coordinate a number of different cognitive skills. These tasks do not involve the use of specialized knowledge or technical skills, nor do they require the mastery of any particular curriculum material.

Because of the complex thinking which these tasks involve, they are not "pure" in any factorial sense and cannot be neatly classified as measuring thinking of any particular "type" (e.g., "convergent" or "divergent"). Instead, each test typically requires the interrelated use of <u>several</u> types of thinking. Thus the classification of tests which is given below merely indicates the relative emphasis given to a certain kind of thinking by the tests in each category; it does <u>not</u> mean that the tests in a certain category involve only one kind of thinking.

A complete list and classification of the tests of productive thinking used in this study is presented in Table 1. In this classification, the terms "convergent" and "divergent" thinking are used in the traditional way to denote thinking which, on the one hand, involves processing information so as to "converge" upon a single solution to a problem or, on the other hand, involves the use of information as a

[1] Four tests in the battery (Circles, Squares, Truck Improvement, and Dog Improvement) were developed by Dr. E. Paul Torrance while at the University of Minnesota. His permission to use those tests as part of this study is gratefully acknowledged.

Table 1

Classification of the Tests of Productive Thinking Used in this Study

Name of test	Use	Performance indicators
1. Complex Extended Problems with an Emphasis on Convergent Thinking (all verbal)		
Pit problem	pretest	number of ideas; quality of ideas (two indices); achievement of solution
Jewel problem (short form)	pretest	number of ideas, quality of ideas (two indices); achievement of solution
Rare Coins problem	internal test	achievement of solution
Jewel problem (long form)	posttest	number of ideas; quality of ideas (two indices); achievement of solution
Black House problem	posttest	number of ideas; quality of ideas (two indices); achievement of solution; number of odd facts or discrepancies in information that are noted
X-ray problem	posttest	number of ideas; quality of ideas (two indices); achievement of solution; number of relevant questions asked or preparatory steps taken before work on problem was begun
1. Complex Extended Problems with an Emphasis on Divergent Thinking (all verbal)		
Elevator problem	pretest	number of ideas
The Deep Sea Dive	internal test	number of odd or puzzling facts within the problem that are noted; achievement of solution
The Ancient City	posttest	number of ideas; quality of ideas (two indices); number of relevant questions asked or preparatory steps taken before work on the problem was begun
A Visit to Karam	posttest	number of puzzling facts which are noted; intellectual persistence
3. Brief Problems with an Emphasis on Convergent Thinking (all verbal)		
Ingenuity Items (Form A)	pretest	achievement of solution (number of solutions achieved: 3 possible)
Ingenuity Items (Form B)	posttest	achievement of solution (number of solutions achieved: 10 possible)
4. Brief Problems with an Emphasis on Divergent Thinking		
Squares	pretest nonverbal	number of ideas; quality of ideas (two indices)
Truck Improvement	pretest verbal	number of ideas; quality of ideas (two indices)

Table 1 (cont.)

Name of Test	Use	Performance indicators
Suppose People Could Read Minds	pretest verbal	number of ideas; number of inventions suggested
A Summer Carnival (Form A)	pretest verbal	quality of ideas (two indices); number of available elements used when creating a product
Circles	posttest nonverbal	number of ideas; quality of ideas (two indices)
Dog Improvement	posttest verbal	number of ideas; quality of ideas (two indices)
Suppose There was an Intelligence Pill	posttest verbal	number of ideas; number of inventions suggested
A Summer Carnival (Form B)	posttest verbal	quality of ideas (two indices); number of available elements used when creating a product
5. Test of the Master Thinking Skill (verbal)		
The Missing Statue	posttest	total number of "best" or "second best" alternatives to which the student gives a rank of one; number of occasions on which the student chooses the "give up" alternative

point of departure from which one's thinking "diverges" into a variety of new mental products. "Complex Extended Problems" in the classification are those which feature a gradual development of the problem and require the student to use a number of different cognitive skills as he proceeds. The student responds at several points and is presented with feedback in the form of more information about the problem after each response opportunity. (A problem which uses these response-feedback sequences is illustrated and discussed below in the section on Performance Indicators and Scoring.) Those tests labelled "Brief Problems" require the use of relatively few cognitive strategies and do not employ response-feedback sequences; instead, they present the entire problem to the student at once, and he then makes his response(s).

All the problems are further classified as verbal or non-verbal, and whether they were used as pre- or posttests or as internal tests in this study. The name of each test and the indices of thinking proficiency on which it was scored are also listed.

Performance Indicators and Scoring

On each of the various tests, several measures of performance were obtained; the particular measures employed depended on the nature of the problem. The following example both illustrates the general nature of a Complex, Extended Problem and presents a detailed discussion of the rationale and procedures by which performance was scored.

The problem, titled The Mystery of the Old Black House, is presented to the student in the form of a brief written story. The story begins as a detective drives out to the country to investigate an old black house in which gold is reported to be hidden. The detective finds the house in the late afternoon and begins his search, but he stops work just before sunset and goes to a nearby white house where he has dinner and spends the night. When he awakens the next morning, the black house has completely disappeared, without leaving any trace whatsoever. The problem is to explain how the black house could have disappeared. Embedded in the story are several puzzling discrepancies (e.g., the detective saw the sunset through his bedroom window but saw the sunrise through the same window when he awoke the next morning) which remain unaccounted for if one assumes that the black house must somehow have been torn down or moved away. However, the problem can be solved in a way that neatly accounts for these discrepancies by concluding that it was actually the detective who was moved, in his sleep, to a similar white house some distance away. This, then, is the principal solution—one which meets all the requirements of the problem and accounts for all the puzzling discrepancies as well.

The story, containing all essentials of the problem, is presented to the student on the first page of a short booklet. On the next page, he is asked to write down all the ideas he has for explaining how the black house could have disappeared. He is encouraged to write as many ideas as possible, especially unusual ones. After writing his ideas, he turns the page and then is asked to write down any odd or puzzling facts he has noticed in the story (besides the disappearance of the black house itself). The following page provides feedback, focusing the student's attention on the several odd or discrepant facts in the story. Then he is given another opportunity to write down any new ideas he has for explaining how the black house could have disappeared. Next, a succession of question-response-feedback units gradually provides the student with more and more information about the problem, giving him additional opportunities to write down any new ideas he has for explaining how the black house could have disappeared and leading him step-by-step toward the principal solution. Finally, he is given a last opportunity to write down ideas for explaining the disappearance of the black house.

A number of indices of thinking proficiency can be applied to performance on such a problem, but because of the massive amount of test data which this study produced, it was decided to use only those indices which were felt to be essential for a reasonably comprehensive yet manageable analysis of performance. In this particular problem, the following performance indicators were used: whether or not the student achieved the principal solution to the problem; the number of odd facts or discrepancies in the problem which he noticed; the number of ideas he wrote for explaining how the black house might have disappeared (regardless of the quality or adequacy of these ideas); and, finally, two indices of the quality of the ideas he produced.

The quality of an idea was judged on the basis of (1) the degree of imaginativeness exhibited and (2) the extent to which it accounted for the various facts without violating the constraints of the problem. Based on previous research with this problem, a normative scale of quality which incorporated these two criteria (presented in Table 2) was prepared. Each idea which the student produced in the course of working on the problem was rated with respect to this scale. As in the case of all quality ratings done in this study, the ratings were made by trained scorers, each of whom first "practice rated" 50 previously obtained protocols and then exchanged these protocols with another scorer who rated them independently. The different sets of quality ratings were then compared for reliability, and when a high degree of reliability between scorers was evident the regular ratings were begun. Because of the large number of protocols in this study, the task of rating the protocols from any given test was divided among five scorers. In general, each protocol was rated only once, so no interjudge reliabilities could be computed. However, a systematic program of rescoring sets of randomly selected protocols resulted in high interjudge agreement. Moreover, because the protocols were divided among the scorers on a semirandom basis, the effect of a scoring bias by any particular scorer should not affect the outcomes in a systematic way. In addition, a "blind" scoring procedure was used, such that the scorers had no way of knowing from which experimental condition a given protocol had been drawn. All quality ratings in the study were made by the same five scorers.

When a scorer had assigned a quality rating to each of a student's ideas in a problem, two indices of quality were generally recorded: the number of "high quality" ideas (in the case of the Black House problem, those receiving a rating of three or higher), and the quality rating of the best idea which the student produced.

This general procedure for scoring and recording the quality of ideas was used for all measures in which quality was scored as a performance indicator.

Test of the Master Thinking Skill

One problem in the test battery requires special mention because of its unique construction and scoring. This problem, a test of the Master Thinking Skill, was designed specifically to measure the student's ability to manage and deploy the various cognitive operations involved in solving a complex extended problem. The student is first presented with a statement of the basic problem. He then confronts a "choice point" at which he is presented with a list of five alternative courses of action, any of which might be taken at that point in the problem. His task is to indicate which he thinks is the "best" course of action, then the second best, and so on, by ranking these

Table 2

Normative Quality Ratings for Ideas in The Mystery of the Old Black House

Rating	Type of ideas	Examples
0	Ideas which are irrelevant, impossible, or contrary to fact	The black house was never there at all; a magician destroyed it
1	Ideas which explain the apparent disappearance of the black house, but which account neither for the fact that no trace of the house was found nor for the discrepancies in the story	The black house blew up; it was torn down
2	Ideas which account for both the apparent disappearance of the house and for the fact that no trace of the house was found, but still do not explain the discrepancies in the story	It was removed by a helicopter; it was moved by a truck and then the tracks were covered; it was carefully camouflaged during the night
3	Ideas which explain the apparent disappearance of the black house, and which account for all the facts and discrepancies in the story, but can only do so by denying the reality of the problem	The detective was drugged during supper, so he was confused when he woke up and only thought the black house had disappeared; the drugs made him see things strangely; it was all a dream
4	Elegant, feasible, ideas which account for all the facts and events	The principal solution: The detective was moved to another highly similar white house during the night

alternatives from one to five. Next, he turns the page, and the development of the problem continues. After sufficient new information has been presented, the student encounters a second "choice point" with its accompanying list of alternative courses of action. Again, he indicates which of these courses he thinks would be best, second best, and so on. There is a total of six choice points, each with its own list of possible courses of action.

Each list of choices always contains the following alternatives (in a random order, of course): a "best" course of action (the action that would be most effective at this point), a "second best" course of action (an action that is reasonable, though not as effective as the "best" course), a "contrary to fact" course of action (one that ignores or violates an already established fact), an "appealing but irrelevant" action (one that is attractive but would not aid in solving the problem), and finally, in the first three choice points, a "jumping to conclusions" course of action (one which "forces" a conclusion to the problem, but does so on the basis of insufficient information). For the last three choice points, this "jumping to conclusions" alternative is replaced by a "give up" alternative (in which the student indicates that he wants to stop working on the problem and simply wants to find out what the solution is).

Two different scores were used as performance indicators: the number of times (summed over the six choice points) that the student gave a rank of one to either the "best" or the "second best" course of action, and the number of times he gave a rank of one to the "give up" alternative presented in the last three choice points. The first score provides a measure of the Master Thinking Skill per se, while the second score constitutes a reciprocal measure of intellectual persistence (i.e., the extent to which the student keeps on working on the problem when he is given several opportunities to give up). Another measure of intellectual persistence was also obtained by the same general means in one other test (A Visit to Karam).

Measures of Understanding Thinking

In addition to measuring the student's performance on problem-solving tasks, an attempt was made to probe his understanding of the nature of problem solving by presenting him with three open-ended sentences which he was asked to complete. The sentences were: "Making mistakes in working on a problem is__________"; "Wild or silly ideas are__________"; and, "In order to solve problems you must__________."

The student's answers were rated by scorers on a three point scale, indicating the degree of insight or understanding which was shown by the answer. For example, in the first sentence above, a rating of zero was given to an answer such as "Making mistakes in working on a problem is bad." A rating of one was given to answers such as "Making mistakes in working on a problem is something that happens even to good thinkers." A rating of two was given to an answer such as "Making mistakes in working on a problem can show you where you went wrong and lead you to new ideas."

Each of the three sentences was scored as a separate variable and is referred to as Sentence I, II, or III, respectively, in the analysis.

EXPERIMENTAL DESIGN AND STATISTICAL ANALYSES

The basic analyses for this experiment were done with a 2 × 2 × 2 factorial design. The factors represented were: Instruction (Treatment vs. Control); Environment (Facilitative vs. Nonfacilitative); and Sex. There were 11 classrooms within each of the cells of this design, and standard analysis of variance procedures were used for the analyses.

Because the investigators also wished to look at the relationships between IQ and the Treatment and Environment factors, a secondary analysis was performed, using a 3 × 2 × 2 factorial design. Three levels of IQ were represented in this design—high (above 115), middle (101 through 115), and low (100 and below). The other two factors were Treatment-Control and Facilitative-Nonfacilitative Environment. After all subjects had been blocked into three groups on the basis of IQ, a random sample of 20 was selected at each IQ level within each Treatment × Environment combination in the design. Thus, a total of 240 subjects was used for this secondary analysis. (In order to obtain such a sample, it was necessary to ignore classifications according to sex and classroom units involved in the experiment.)

For both the main analysis (the 2 × 2 × 2 factorial design) and the secondary analysis (the 3 × 2 × 2 factorial design involving IQ), separate analyses of variance were performed on each of five stratification variables—socioeconomic status, overall mean score on the Stanford Achievement Battery, IQ, performance on the Sequential Test of Educational Progress (Social Studies, Form 4A) when given as a pretest, and scores on this test when given as a posttest (using Form 4B). Similar analyses were performed on all pretest, internal test, and posttest measures.

In the second design, the analysis was performed using scores of individual students as the basic unit, in contrast to the main design where the classroom mean score for each variable was the unit of analysis.

III
RESULTS

Tables 3, 4 and 5 present the means and standard deviations of all the variables analyzed in this experiment. Values for the stratification variables are in Table 3, for the pretest measures in Table 4, and for the internal and posttest measures in Table 5.

The results of the analyses of variance for the stratification variables are given in Table 6. Note that there were no significant Treatment-Control differences on any of these variables. The Environment factor, on the other hand, represented a significant source of variance on all five of these variables. Hence the supervisor's ratings of what constituted a "facilitative" classroom environment reflected, to some extent, the calibre of students in the class being rated. However, with the exception of socioeconomic status, the mean differences between Facilitative and Nonfacilitative classes on these stratification variables were fairly small, although statistically significant. Means for the Treatment-Control and Facilitative-Nonfacilitative Environment groups on the stratification variables are presented in Table 7.

The data in Table 6 also indicate significant sex differences in achievement and IQ. These differences arise from the fact that girls are superior to boys on these two variables (Table 8). (Although differences on the two forms of the STEP test were not significant, they were in the same direction.) These differences are consistent with other findings concerning sex differences in aptitude and achievement at this level of development.

Table 3

Means and Standard Deviations of Stratification Variables

Variable	Mean	Standard Deviation
Socioeconomic status	2.98	1.49
Stanford achievement	5.09	.65
Kuhlmann-Anderson IQ	107.39	7.68
STEP Test (Pretest: Form 4A)	26.44	3.22
STEP Test (Posttest: Form 4B)	25.48	3.15

Table 4

Means and Standard Deviations of Pretest Variables

Variable	Mean	Standard Deviation
Extended problems in convergent thinking[a]		
Pit problem:		
Achievement of solution	.42	.20
Number of ideas	3.63	.74
Number high quality ideas	1.08	.37
Quality rating, best idea	3.02	.43
Jewel problem:		
Achievement of solution	.06	.09
Number of ideas	3.53	1.06
Number high quality ideas	.37	.21
Quality rating, best idea	1.10	.35
Brief problem in convergent thinking[a]		
Ingenuity test:		
Achievement of solution	.48	.26
Extended problem in divergent thinking[a]		
Elevator problem:		
Number of ideas	2.92	1.08
Brief problems in divergent thinking		
Squares problem (nonverbal):		
Number of ideas	8.45	2.49
Number high quality ideas	1.92	.83
Quality rating, best idea	1.65	.26
Truck improvement problem (verbal):		
Number of ideas	8.04	2.30
Number high quality ideas	1.42	.62
Quality rating, best idea	1.64	.22
Mind-reading problem (verbal):		
Number of ideas	3.94	1.43
Inventiveness	.20	.25
Summer carnival problem (verbal):		
Number high quality ideas	.40	.29
Quality rating, best idea	6.10	.30
Ability to synthesize elements	2.10	.59

[a]Problems in this group are all verbal.

ERIC

Table 5

Means and Standard Deviations of Internal and Posttest Variables

Variable	Mean	Standard Deviation
Internal test measures[a]		
Rare coins problem (extended, convergent):		
Achievement of solution	.21	.19
Deep Sea Dive problem (extended, divergent):		
Achievement of solution	.27	.19
Sensitivity to incongruous information	1.46	.57
Extended problems in convergent thinking[a]		
Jewel problem:		
Achievement of solution	.51	.25
Number of ideas	6.18	1.57
Number high quality ideas	1.29	.35
Quality rating, best idea	2.23	.45
House problem:		
Achievement of solution	.65	.20
Number of ideas	4.18	.79
Number high quality ideas	1.17	.38
Quality rating, best idea	4.09	.58
Sensitivity to incongruous information	.53	.36
X-ray problem:		
Achievement of solution	.15	.11
Number of ideas	2.79	.83
Number high quality ideas	.82	.41
Quality rating, best idea	2.58	.45
Questioning and clarifying ability	.63	.44
Brief problem in convergent thinking[a]		
Ingenuity:		
Achievement of solution	3.27	.99
Extended problems in divergent thinking[a]		
Ancient city problem:		
Number of ideas	4.79	1.54
Number high quality ideas	.35	.20
Quality rating, best idea	1.00	.33
Questioning and clarifying ability	.87	.58
A Visit to Karam:		
Sensitivity to puzzling facts	.74	.78
Intellectual persistence	2.46	.84

[a]Problems in this group are all verbal.

(Continued on next page)

ERIC

Table 5 (cont'd)

Variable	Mean	Standard Deviation
Brief problems in divergent thinking		
Circles (nonverbal):		
Number of ideas	14.08	2.52
Number high quality ideas	2.08	.75
Quality rating, best idea	1.76	.18
Dog improvement (verbal):		
Number of ideas	7.92	2.18
Number high quality ideas	1.19	.60
Quality rating, best idea	1.54	.24
Intelligence pill (verbal):		
Number of ideas	2.81	.87
Inventiveness	.06	.09
Summer carnival (verbal):		
Number high quality ideas	.47	.26
Quality rating, best idea	1.16	.32
Ability to synthesize elements	1.80	.61
Special tests		
Statue problem:		
Effectiveness of Master Thinking Skill	3.53	.66
Intellectual persistence	1.59	.80
Understanding thinking:		
Sentence I	.66	.37
Sentence II	1.00	.28
Sentence III	1.24	.23

[a]Problems in this group are all verbal.

Table 6

Summary of Analyses of Variance for Stratification Variables
(df = 1, 80)

Variable	Treatment		Environment		Sex	
	MS	F	MS	F	MS	F
Socioeconomic status	.05	<1	24.04	11.58**	.00	< 1
Stanford achievement	.01	<1	2.09	5.64*	3.01	7.81**
Kuhlmann-Anderson IQ	.75	<1	484.02	9.37**	454.70	8.81**
STEP Pretest	.17	<1	47.39	4.55*	6.60	< 1
STEP Posttest	6.36	<1	42.62	4.30*	3.95	< 1

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

ERIC

Table 7

Mean Level of Performance of Students in Treatment and Control Groups and in Facilitative and Nonfacilitative Environments: Stratification Variables

Variable	Group		Environment	
	Treatment	Control	Facilitative	Nonfacilitative
Socioeconomic status**	2.95	3.00	2.45	3.50
Stanford achievement*	5.11	5.09	5.25	4.93
Kuhlmann-Anderson IQ**	107.49	107.31	109.74	105.05
STEP Pretest*	26.40	26.49	27.18	25.71
STEP Posttest*	25.22	25.77	26.18	24.79

*"Environment" difference significant at $p < .05$.
**"Environment" difference significant at $p < .01$.

Table 8

Mean Level of Performance of Males and Females: Stratification Variables

Variable	Males	Females
Socioeconomic status	2.98	2.98
Stanford achievement*	4.91	5.28
K-A IQ*	105.12	109.67
STEP Pretest	26.17	26.72
STEP Posttest	25.27	25.70

*Difference significant at $p < .01$

PRETEST RESULTS

In Table 9 is a summary of the analyses of the 21 pretest variables. In the 21 analyses, there was only one Treatment effect significant at the .05 level. It seems likely that this significant F is simply an artifact of the number of analyses carried out. The means for treatment and control groups are presented in Table 10.

The Environment factor produced significant differences on five pretest measures, all of which concerned the two brief problems in divergent thinking—the nonverbal Squares problem and the verbal Truck problem. Inspection of the means in Table 11 reveals that these differences are all in the same direction, and all reflect the superior performance of students in the Facilitative environment.

Tables 12 and 13 present the means for males and females on the pretest measures. Except for the Pit problem, where males were superior to females on three of the four performance indicators, females performed better than males for all measures on which statistically significant sex differences were found.

These, then, were the differences which existed among fifth-grade children on various measures of productive thinking before any special training was administered. Let us turn now to the results obtained after Series One of The Productive Thinking Program had been administered.

Table 9

Summary of Analyses of Variance of Pretest Variables
(1 and 80 degrees of freedom)

	Main effects					
	Treatment		Environment		Sex	
Variables	MS	F	MS	F	MS	F
Extended problems in convergent thinking[a]						
Pit problem:						
Achievement of solution	.00	< 1	.00	< 1	.41	11.59**
Number of ideas	1.34	2.59	1.93	3.74	2.21	4.27*
Number high quality ideas	.06	< 1	.07	< 1	1.18	9.31**
Quality rating, best idea	.02	< 1	.00	< 1	2.16	13.07**
Jewel problem:						
Achievement of solution	.04	5.46*	.00	< 1	.01	1.78
Number of ideas	.54	< 1	2.42	2.11	.49	< 1
Number high quality ideas	.03	< 1	.11	< 1	.07	< 1
Quality rating, best idea	.38	3.17	.11	< 1	.07	< 1
Brief problem in convergent thinking[a]						
Ingenuity test:						
Achievement of solution[b]	.00	< 1	.25	3.68	.22	3.34
Extended problem in divergent thinking[a]						
Elevator problem:						
Number of ideas	.57	< 1	.07	< 1	3.87	3.19
Brief problems in divergent thinking						
Squares problem (nonverbal):						
Number of ideas	22.76	3.92	23.66	4.08*	19.64	3.38
Number high quality ideas	.26	< 1	5.73	8.90**	.80	1.24
Quality rating, best idea	.00	< 1	.32	4.70*	.15	2.21
Truck improvement problem (verbal):						
Number of ideas	.00	< 1	.98	< 1	4.91	< 1
Number high quality ideas	.00	< 1	2.04	5.37*	.24	< 1
Quality rating, best idea	.03	< 1	.27	5.56*	.02	< 1
Mind-reading (verbal):						
Number of ideas	.05	< 1	.07	< 1	15.54	8.38**
Inventiveness	.00	< 1	.01	< 1	.15	2.45
Summer carnival (verbal):						
Number high quality ideas	.01	< 1	.03	< 1	.47	5.67*
Quality rating, best idea	.01	< 1	.01	< 1	.47	5.10*
Ability to synthesize elements	.05	< 1	.01	< 1	5.77	19.11**

[a]Problems in this group are all verbal.

[b]Number of solutions out of a possible three.

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

Table 10

Mean Level of Performance of Treatment and Control Groups: Pretest Variables

Variable	Treatment	Control
Pit problem:		
Achievement of solution	.43	.41
Number of ideas	3.75	3.51
Number high quality ideas	1.05	1.10
Quality rating, best idea	3.00	3.03
Jewel problem:		
Achievement of solution	.08	.04
Number of ideas	3.61	3.45
Number high quality ideas	.39	.36
Quality rating, best idea	1.16	1.04
Ingenuity test:		
Achievement of solution	.48	.48
Elevator problem:		
Number of ideas	3.00	2.84
Squares problem:		
Number of ideas	8.96	7.94
Number high quality ideas	1.98	1.87
Quality rating, best idea	1.66	1.64
Truck improvement problem:		
Number of ideas	8.03	8.04
Number high quality ideas	1.43	1.42
Quality rating, best ideas	1.65	1.62
Mind-reading problem:		
Number of ideas	3.97	3.92
Inventiveness	.20	.21
Summer carnival:		
Number high quality ideas	.42	.39
Quality rating, best idea	1.11	1.09
Ability to synthesize elements	2.12	2.08

Table 11

Comparison of Ideas Generated by Students in Facilitative (F) and Nonfacilitative (N) Environments: Pretest Variables

Problem	Fluency		No. high quality		Quality of best	
	F	N	F	N	F	N
Pit	3.48	3.78	1.05	1.11	3.01	3.02
Jewel	3.36	3.70	.40	.35	1.13	1.06
Elevator	2.94	2.89	-	-	-	-
Squares	8.97*	7.93*	2.18**	1.67**	1.71*	1.59*
Truck	8.14	7.93	1.58*	1.27*	1.69*	1.58*
Mind: No. of Ideas	3.97	3.91	-	-	-	-
Mind: Inventiveness	.21	.19	-	-	-	-
Carnival	2.11[a]	2.09[a]	.42	.38	1.01	1.11

[a]Not a measure of fluency. This entry represents the number of available elements used when creating a "product."

*Difference significant at $p < .05$.
**Difference significant at $p < .01$.

Table 12

Mean Level of Achievement of Solution for Males and Females: Pretest Variables

Test	Males	Females
Pit*	.49	.35
Jewel	.07	.05
Ingenuity	.53	.43

*Difference significant at $p < .01$

Table 13

Comparison of Ideas Generated by Males (M) and Females (F): Pretest Variables

	Fluency		No. high quality		Quality of Best	
	M	F	M	F	M	F
Pit	3.47*	3.79*	1.19**	.96**	3.17**	2.86**
Jewel	3.46	3.61	.39	.36	1.13	1.07
Elevator	2.71	3.12	-	-	-	-
Squares	7.97	8.92	1.83	2.02	1.61	1.69
Truck	7.80	8.28	1.48	1.37	1.65	1.62
Mind: No. of Ideas	3.52**	4.36**	-	-	-	-
Mind: Inventiveness	.16	.24	-	-	-	-
Carnival	1.84[a]**	2.36[a]**	.33*	.48*	1.03*	1.17*

[a]Not a measure of fluency. This entry represents the number of available elements used when creating a "product."

*Difference significant at $p < .05$.
**Difference significant at $p < .01$.

ERIC

Table 14

Summary of Analyses of Variance:
Main Effects for Internal and Posttest Variables
(1 df for effects)

Variable	Source					
	Treatment		Environment		Sex	
	MS	F	MS	F	MS	F
Internal Measures[a]						
Rare coins problem:						
Achievement of solution	.02	< 1	.07	2.15	.00	< 1
Deep sea dive problem:						
Achievement of solution	.05	1.46	.00	< 1	.00	< 1
Sensitivity to incongruous information	2.25	7.41**	.00	< 1	1.25	4.11*
Extended problems in convergent thinking[a]						
Jewel problem:[c]						
Achievement of solution	.11	1.61	.01	< 1	.01	< 1
Number of ideas	13.46	6.59*	.05	< 1	4.78	2.34
Number high quality ideas	.25	2.37	.02	< 1	.37	3.50
Quality rating, best idea	.62	2.80	.16	< 1	.04	< 1
House problem:[d]						
Achievement of solution	.23	6.47*	.11	2.98	.04	1.15
Number of ideas	.54	< 1	.14	< 1	1.61	2.52
Number high quality ideas	2.05	16.99**	.03	< 1	.57	4.71*
Quality rating, best idea	1.84	5.98*	1.13	3.67	.52	1.68
Sensitivity to incongruous information	1.50	15.65**	.06	< 1	.67	7.00**
X-ray problem:[d]						
Achievement of solution	.16	6.20*	.03	1.04	.00	< 1
Number of ideas	.00	< 1	.01	< 1	1.05	1.45
Number high quality ideas	.45	2.73	.01	< 1	.01	< 1
Quality rating, best idea	.99	5.22*	.00	< 1	.02	< 1
Sensitivity to incongruous information	.50	2.83	.00	< 1	1.70	9.59**
Brief problem in convergent thinking[a]						
Ingenuity:[d]						
Achievement of solution[b]	.45	< 1	4.84	4.89*	.92	< 1
Extended problems in divergent thinking[a]						
Ancient city problem:[c]						
Number of ideas	22.66	10.97**	1.59	< 1	.11	< 1
Number high quality ideas	.01	< 1	.00	< 1	.08	1.88
Quality rating, best idea	.16	1.39	.00	< 1	.18	1.60
Questioning and clarifying ability	.06	< 1	.49	1.47	1.15	3.44

[a]Problems in this group are all verbal.
[b]Total number of solutions out of a possible ten.
[c]36 df for error.
[d]80 df for error.
*$p < .05$
**$p < .01$

(Continued on next page)

Table 14 (cont'd)

Variable	Treatment MS	Treatment F	Environment MS	Environment F	Sex MS	Sex F
A Visit to Karam:[c]						
Sensitivity to puzzling facts	1.32	2.12	.00	< 1	.00	< 1
Intellectual persistence	2.74	4.15*	1.78	2.69	.23	< 1
Brief problems in divergent thinking						
Circles (nonverbal):[d]						
Number of ideas	10.00	1.68	5.02	< 1	55.63	9.37**
Number high quality ideas	.08	< 1	.53	< 1	.37	< 1
Quality rating, best idea	.04	1.39	.01	< 1	.00	< 1
Dog improvement (verbal):[d]						
Number of ideas	10.98	2.72	6.20	1.53	57.05	14.13**
Number high quality ideas	.34	1.03	.57	1.71	1.52	4.55*
Quality rating, best idea	.00	< 1	.08	1.38	.34	1.09
Intelligence pill (verbal):[d]						
Number of ideas	.54	< 1	.79	1.20	6.61	9.99**
Inventiveness	.38	4.33*	.00	< 1	.00	< 1
Summer carnival (verbal):[d]						
Number high quality ideas	.02	< 1	.02	< 1	.98	16.97**
Quality rating, best idea	.01	< 1	.01	< 1	1.70	20.49**
Ability to synthesize elements	.22	< 1	.01	< 1	4.99	15.41**
Special tests						
Statue problem:[d]						
Effectiveness of Master Thinking Skill	1.97	5.06*	.14	< 1	1.59	4.08
Intellectual persistence	.04	< 1	.09	< 1	1.50	< 1
Understanding thinking:[c]						
Sentence I	.68	6.79*	.33	3.27	.81	8.12**
Sentence II	.13	1.58	.01	< 1	.28	3.42
Sentence III	.08	1.56	.06	1.17	.18	3.45

[a]Problems in this group are all verbal.
[b]Total number of solutions out of a possible ten.
[c]36 df for error.
[d]80 df for error.
*p < .05.
**p < .01.

INTERNAL AND POSTTEST RESULTS

Table 14 summarizes the results of the analyses of variance of all internal and posttest measures. Of particular interest are the Treatment-Control differences, thirteen of which are statistically significant (as opposed to only one such difference that reached significance on the pretest measures). Table 15 presents the means for the Treatment and Control groups on incidence of achievement of solution for problems which have an emphasis on convergent thinking (the type of problem which permits only a small set of "best" solutions).[2] Two of the six measures show significant differences, and on five of the six measures performance of the Treatment group exceeds that of the Controls.

The sixth measure, involving the X-ray problem, requires special comment since this problem was the only one in the entire posttest

[2] The Deep Sea Dive problem, although it emphasizes divergent thinking, was included in this Table because it, too, permits only a small set of "best" solutions.

Table 15

Mean Level of Achievement of Solution for Treatment and Control Groups: Internal and Posttest Variables

Problem	Treatment	Control
Rare coins	.22	.19
Deep sea dive	.29	.24
Jewel	.56	.45
House*	.70	.59
X-ray*	.11	.19
Ingenuity	3.34	3.20

*Difference significant at $p < .05$.

battery on which the Control group consistently outperformed the Treatment group by a statistically significant margin. This unexpected reversal of the general pattern of Treatment group superiority appears to be an artifact of this particular problem, most likely due to the fact that the task was so difficult (only 15% of all students ever achieved solution) that an outstanding performance by relatively few Control students could produce significant difference in favor of that group.

On measures dealing with the number and quality of ideas (Table 16), six significant Treatment-Control differences were found. Five of these differences favored the Treatment group, while the sixth (again involving the X-ray problem) favored the Controls. The means for Treatment-Control performance on all remaining internal and posttest measures are presented in Table 17, and all significant differences in this table were the result of superior performance by the Treatment group.

In brief, then, Series One of The Productive Thinking Program produced significant Treatment-Control differences on one-third (13 out of 40) of the internal and posttest measures, when only one significant difference (out of 21 measures) existed before training began. Eleven of the significant posttest differences indicated superior performance by the Treatment group, while the two differences which did not both involved the maverick X-ray problem.

Turning now to differences resulting from the two types of classroom environment, Table 14 shows that only one such difference (out of the 40 posttest analyses) was statistically significant. This is a considerable contrast to the pretest results in which five significant differences (out of 21 analyses) were found, all of which indicated the superiority of a Facilitative environment. The notable lack of posttest effects may partly be accounted for by the presence of several significant Treatment × Environment interactions (see below), all of which indicate that treatment was more effective with students in Nonfacilitative environments, thus reducing the initial differences between Facilitative and Nonfacilitative groups. Several additional Treatment × Environment interactions,

Table 16

Comparison of Ideas Generated by Students in Treatment (T) and Control (C) Groups: Posttest Variables

Problem	Fluency		No. high quality		Quality of best	
	T	C	T	C	T	C
Jewel	6.77*	5.58*	1.37	1.21	2.35	2.11
House	4.26	4.10	1.32**	1.01**	4.23*	3.94*
X-ray	2.79	2.79	.75	.89	2.48*	2.69*
Ancient city	5.22**	4.06**	.36	.34	1.06	.95
Circles	14.42	13.75	2.05	2.11	1.74	1.78
Dog	8.28	7.57	1.26	1.13	1.54	1.53
Pill: No. of ideas	2.73	2.88	-	-	-	-
Pill: Inventiveness	.08*	.03*	-	-	-	-
Carnival	1.75[a]	1.85[a]	.45	.48	1.15	1.17

[a]Not a measure of fluency. This entry represents the number of available elements used when creating a "product."

*Difference significant at $p < .05$.

**Difference significant at $p < .01$.

Table 17

Means for Treatment and Control Groups on Remaining Internal and Posttest Measures

Variable	Treatment	Control
Deep sea dive problem:		
Sensitivity to incongruous information	1.63*	1.29**
House problem:		
Sensitivity to incongruous information	.66*	.40**
X-ray problem:		
Sensitivity to incongruous information	.55	.71
Ancient city problem:		
Questioning and clarifying ability	.90	.84
A Visit to Karam:		
Sensitivity to puzzling facts	.90	.58
Intellectual persistence	2.70*	2.22*
Statue problem:		
Effectiveness of Master Thinking Skill	3.73*	3.32*
Intellectual persistence	1.56	1.61
Insight and understanding of thinking:		
Sentence I	.78*	.53*
Sentence II	1.05	.95
Sentence III	1.28	1.20

*Difference significant at $p < .05$.
**Difference significant at $p < .01$.

though they failed to reach significance, also showed this pattern.

Each of the significant sex differences reported in Table 14 (12 in 40 analyses) reflects the superior performance of females. Seven of the 12 differences occurred on variables dealing with the production of ideas (see Table 18), three differences reflected the greater sensitivity of females to discrepant or puzzling information embedded in the tests, and the remaining two differences occurred on a variety of measures. With respect to the verbal fluency measures, the data are consistent with the numerous findings that girls are verbally more proficient than boys.

In the analyses of all these measures—pretest, internal test, and posttest—there was a total of seven significant interactions, all in the internal and posttests. Two of these, three-way interactions of Treatment × Environment × Sex, were uninterpretable. The other five were Treatment × Environment interactions, for which the means are presented in Table 19. The interactions for these five measures all have the same pattern: within the Treatment condition those students in Nonfacilitative environments performed better than students in Facilitative environments, while for the Control group, students in the Facilitative environments performed better than those in the Nonfacilitative environments. Thus, it would seem that one important effect of the training materials was to enable students from less favorable classroom climates to overcome the detrimental effects that such environments would otherwise have produced on their performance. This helps explain the fact (noted earlier) that classroom environment was a significant source of variance on 5 out of 21 pretest measures, but on only 1 out of 40 posttest measures.

EFFECTS OF IQ

The second major analysis, carried out with the 3 × 2 × 2 design, involved three levels of IQ (above 115, 101-115, and 100 or below), the two Treatment conditions, and the two types of Environment. The purpose of this analysis was to determine the effect of IQ on performance, both as a main effect and in its interactions with Treatment or Environment. Since the main effects of the Treatment and Environment factors have already been discussed, they will not be considered in connection with this analysis.

Table 18

Comparison of Ideas Generated by Males (M) and Females (F): Posttest Variables

Problem	Fluency		No. high quality		Quality of best	
	M	F	M	F	M	F
Jewel	5.84	6.51	1.20	1.38	2.20	2.27
House	4.04	4.31	1.08*	1.25*	4.01	4.16
X-ray	2.68	2.90	.81	.83	2.60	2.57
Ancient city	4.74	4.84	.31	.39	.94	1.06
Circles	13.29**	14.88**	2.14	2.01	1.76	1.76
Dog	7.12**	8.73**	1.06*	1.32*	1.47	1.60
Pill	2.53**	3.08**	-	-	-	-
Carnival[a]	1.56**	2.03**	.36**	.57**	1.02**	1.30**

[a]"Fluency" means reflect the number of elements combined to generate a "product."

*Difference significant at $p < .05$.

**Difference significant at $p < .01$.

Table 19

Means for Significant Treatment × Environment Interactions
(1 df for Interaction)

Variable	df for error		Environment: Facilitative	Environment: Nonfacilitative
Rare coins problem (internal test):	80			
Achievement of solution		Treatment	.23	.22
		Control	.28	.12
Jewel problem (posttest):	36			
Number of ideas		Treatment	6.25	7.21
		Control	6.03	5.21
Dog problem (posttest):	80			
Number of ideas		Treatment	8.10	8.46
		Control	8.28	6.86
Number high quality ideas		Treatment	1.20	1.31
		Control	1.34	.92
House problem (posttest):	80			
Sensitivity to incongruous information		Treatment	.54	.79
		Control	.48	.33

Table 20 presents the results of this IQ analysis for the stratification variables. With the exception of socioeconomic status, performance on each of these variables was significantly related to IQ ($p < .01$). Table 21 indicates that the mean differences for these variables are all in the predicted direction—high IQ students perform better on the Stanford and STEP tests than do middle IQ students, who in turn perform better than those with low IQ.

With respect to the pretest measures, Table 22 indicates that IQ represented a significant source of variance for 17 of the 21 measures (15 of these beyond the .01 level). The mean levels for achievement of solution to the three convergent thinking problems used in the pretest are presented in Table 23. In each of these problems, the high IQ group outperformed the others, and the middle IQ group, with one small reversal, was consistently superior to the low IQ group. For measures of the quantity and quality of ideas produced on the various tests, this same rank-order relation between IQ and performance was found (Table 24).

Table 20

Summary of Analyses of IQ as a Main Effect: Stratifying Variables (df for Effect = 2)

Variable	df for Error	IQ	
		MS	F
Socioeconomic status	228	.11	< 1
Stanford achievement	225	108.66	154.58*
STEP Pretest	215	922.08	46.34*
STEP Posttest	215	1049.46	55.66*

*p < .01

Table 21

Means of IQ Groups on Stratifying Variables

Variable	IQ		
	High	Middle	Low
Socioeconomic status	2.84	2.88	2.91
Stanford achievement*	6.32	5.13	3.96
STEP Pretest*	29.73	27.79	22.92
STEP Posttest*	29.18	26.75	21.78

*Differences significant at p < .01.

GPO 806-459-5

Table 22

Summary of Analyses of Main Effect of IQ:
Pretest Variables
(df for Effect = 2)

Variable	df for error	IQ effect MS	IQ effect F
Extended problems in convergent thinking[a]			
Pit problem:	221		
Achievement of solution		1.32	5.67**
Number of ideas		5.68	2.40
Number high quality ideas		5.15	5.48**
Quality rating, best idea		9.62	9.16**
Jewel problem:	225		
Achievement of solution		.46	6.92**
Number of ideas		41.25	10.82**
Number high quality ideas		2.47	7.10**
Quality rating, best idea		9.51	12.52**
Brief problem in convergent thinking[a]			
Ingenuity test:	224		
Achievement of solution[b]		6.68	17.54**
Extended problem in divergent thinking[a]			
Elevator problem:	221		
Number of ideas		38.39	10.80**
Brief problems in divergent thinking			
Squares (nonverbal):	221		
Number of ideas		489.51	26.76**
Number high quality ideas		47.83	15.92**
Quality rating, best idea		3.62	12.70**
Truck improvement (verbal)	221		
Number of ideas		226.61	13.52**
Number high quality ideas		5.34	2.52
Quality rating, best idea		.46	1.67
Mind-reading (verbal):	221		
Number of ideas		142.75	26.39**
Inventiveness		.68	2.38
Summer carnival (verbal):	225		
Number high quality ideas		1.17	3.99*
Quality rating, best idea		2.24	4.31*
Ability to synthesize elements		7.62	5.26**

[a]Problems in this group are all verbal.

[b]Total number of solutions out of a possible three.

*$p < .05$

**$p < .01$.

Table 23

Mean Level of Achievement of Solution for IQ Groups: Pretest Variables

Problem	IQ High	IQ Middle	IQ Low
Pit*	.55	.41	.28
Jewel*	.16	.03	.04
Ingenuity*	.86	.46	.30

*Differences significant at $p < .01$

Turning now to the effect of IQ on the posttest measures (IQ effects on the three internal test measures were not analyzed), Table 25 shows significant differences on 27 of the 38 measures (19 of these beyond the .01 level). For problems emphasizing convergent thinking, significant differences occurred on 13 of the 16 measures; for problems emphasizing divergent thinking, significant differences were found on 13 of the 17 measures; and on 1 of the remaining 5 measures a significant difference was found. These means are presented in Tables 26, 27 and 28; in all cases, the high and middle IQ groups performed substantially better than the low IQ group. However, performance of the high and middle IQ groups on these posttest measures was much more alike (relative to the low IQ group) than it had been in the pretest. The absence of any systematic IQ × Treatment interaction (see below) suggests that this change is not of sufficient magnitude to be considered an important result of exposure to the program. As in the pretest, there was a general rank-order relation between IQ and performance on most of the posttest measures.

A total of nine significant interactions occurred in this second analysis, four involving pretest measures and the other five involving the posttest. The means for these significant interactions are presented in Tables 29 and 30. No consistent pattern appears to be present in these interactions. (Two posttest interactions are omitted, one because it is an IQ × Treatment × Environment interaction which yields no meaningful interpretation, the other because it is a Treatment × Environment interaction which does not involve IQ and about which information is best obtained from the first analysis.)

Table 24

Comparison of Ideas Generated by IQ Groups: Pretest Variables

Problem	Fluency H	Fluency M	Fluency L	No. high quality H	No. high quality M	No. high quality L	Quality of best H	Quality of best M	Quality of best L
Pit	3.90	3.88	3.41	1.35*	1.09*	.83*	3.29*	3.05*	2.59*
Jewel	4.31*	3.49*	2.86*	.58*	.33*	.23*	1.48*	1.08*	.78*
Elevator	3.64*	3.24*	2.27*	-	-	-	-	-	-
Squares	10.97*	8.30*	5.87*	2.62*	1.76*	1.04*	1.86*	1.69*	1.42*
Truck	9.21*	8.73*	6.04*	1.60	1.41	1.08	1.69	1.66	1.54
Mind	4.99*	4.45*	2.37*	-	-	-	-	-	-
Carnival[a]	2.10*	1.92*	1.46*	.48*	.34*	.23*	1.21*	1.11*	.88*

[a]"Fluency" means reflect the number of elements combined to generate a "product."

*Differences significant at $p < .05$.

ERIC

Table 25

Summary of Analyses of Main Effect of IQ:
Posttest Variables
(df for Effect = 2)

Variable	df for Error	IQ effect MS	IQ effect F
Extended problems in convergent thinking[a]			
Jewel problem:	95		
Achievement of solution		.57	2.50
Number of ideas		20.37	3.50*
Number high quality ideas		3.59	3.85*
Quality rating, best idea		2.41	3.07
House problem:	216		
Achievement of solution		3.85	22.72**
Number of ideas		20.67	5.70**
Number high quality ideas		7.14	11.93**
Quality rating, best idea		38.56	24.79**
Sensitivity to incongruous information		5.91	8.45**
X-ray problem:	215		
Achievement of solution I[b]		.05	1.18
Achievement of solution II[b]		1.24	10.41**
Number of ideas		42.18	18.07**
Number high quality ideas		9.89	13.45**
Quality rating, best idea		12.44	15.04**
Questioning and clarifying ability		4.75	9.10**
Brief problem in convergent thinking[a]			
Ingenuity test:	204		
Achievement of solution		100.50	35.43**
Extended problems in divergent thinking[a]			
Ancient city problem	95		
Number of ideas		42.64	7.14**
Number high quality ideas		1.49	4.68*
Quality rating, best idea		3.50	5.26**
Questioning and clarifying ability		5.21	3.44*
A Visit to Karam	98		
Sensitivity to puzzling facts		.32	<1
Intellectual persistence		13.00	6.09**

[a]Problems in this group are all verbal.

[b]These two variables represent a breakdown into separate solutions of the measure of achievement of solution reported for this problem in the earlier analysis.

*$p < .05$

**$p < .01$

(Continued on next page)

Table 25 (cont'd)

Variable	df for Error	IQ effect MS	F
Brief problems in divergent thinking			
Circles (nonverbal):	222		
Number of ideas		198.30	10.07*
Number high quality ideas		16.62	5.47**
Quality rating, best idea		.47	2.49
Dog improvement (verbal):	222		
Number of ideas		87.04	5.01**
Number high quality ideas		4.96	2.78
Quality rating, best idea		.75	2.49
Intelligence pill (verbal):	223		
Number of ideas		20.90	7.96**
Inventiveness		.23	3.44*
Summer carnival (verbal):	223		
Number high quality ideas		1.80	4.23*
Quality rating, best idea		2.11	3.85*
Ability to synthesize		8.57	6.56**
Special tests[a]			
Statue problem:	110		
Effectiveness of Master Thinking Skill		6.37	2.50
Intellectual persistence		1.11	<1
Insight and understanding of thinking:			
Sentence I	74	.94	9.37**
Sentence II	85	.14	2.11
Sentence III	81	.17	3.01

[a]Problems in this group are all verbal.

[b]These two variables represent a breakdown into separate solutions of the measure of achievement of solution reported for this problem in the earlier analysis.

*p < .05

**p < .01

Table 26

Mean Level of Achievement of Solution for IQ Groups: Posttest Variables

Problem	IQ High	IQ Middle	IQ Low
Jewel	.67	.49	.34
House	.86*	.78*	.43*
X-ray[a]	.07, .29*	.03, .12*	.03, .04*
Ingenuity	4.37*	3.59*	2.03*

[a]These two variables represent a breakdown into separate solutions of the measure of achievement of solution reported for this problem in the earlier analysis.

*Differences significant at p = .01.

Table 27

Comparison of Ideas Generated by IQ Groups: Posttest Variables

Problem	Fluency			No. high quality			Quality of best		
	H	M	L	H	M	L	H	M	L
Jewel	7.00*	6.23*	5.49*	1.52*	1.46*	.89*	2.52	2.28	1.89
House	4.23*	4.71*	3.68*	1.33*	1.38*	.83*	4.68*	4.49*	3.35*
X-ray	3.41*	3.13*	2.00*	1.19*	.99*	.48*	3.11*	2.65*	2.24*
Ancient city	5.88*	5.03*	3.86*	.61*	.41*	.17*	1.30*	1.10*	.69*
Circles	15.13*	14.36*	12.06*	2.37*	2.18*	1.49*	1.82	1.78	1.68
Dog	8.65*	7.68*	6.55*	1.29	1.24	.83	1.62	1.58	1.44
Pill: No. of ideas	3.24*	2.79*	2.22*	-	-	-	-	-	-
Pill: Inventiveness	.11*	.04*	.01*	-	-	-	-	-	-
Carnival[a]	2.10*	1.92*	1.46*	.62*	.47*	.32*	1.30*	1.24*	1.00*

[a]"Fluency" means reflect the number of elements combined to generate a "product."
*Differences significant at $p < .05$.

Table 28

Means for Remaining Posttest Variables for which Significant IQ Effect was Found

Variable	IQ H	IQ M	IQ L
House problem: Sensitivity to incongruous information	.92	.55	.38
X-ray problem: Questioning and clarifying ability	.87	.51	.39
Ancient city problem: Questioning and clarifying ability	1.30	1.10	.57
A Visit to Karam Intellectual persistence	3.16	2.84	1.77
Insight and understanding Sentence I	.94	.76	.29

Table 29

Means for Significant Interactions Involving IQ and Treatment

Variable		IQ H	IQ M	IQ L
Jewel problem:				
No. high quality ideas	Treatment	.78	.32	.21
	Control	.38	.32	.26
Squares problem:				
Number of ideas	Treatment	11.40	7.20	6.57
	Control	10.51	9.40	5.20
Pill problem:				
Inventiveness	Treatment	.23	.05	.03
	Control	.00	.03	.00
Summer carnival problem:				
No. high quality ideas	Treatment	.70	.38	.23
	Control	.25	.30	.24

Table 30

Means for Significant Interactions Involving IQ and Environment

Variable		IQ H	IQ M	IQ L
Truck problem:				
Number of ideas	Facil. Env.	8.05	8.60	6.47
	Nonfac. Env.	10.39	8.85	5.60
Summer carnival problem:				
Quality rating, best idea	Facil. Env.	1.22	1.02	1.10
	Nonfac. Env.	1.38	1.47	0.89
X-ray problem:				
Number of ideas	Facil. Env.	3.03	2.82	2.18
	Nonfac. Env.	3.79	3.44	1.81

IV
DISCUSSION

EFFECTS OF TREATMENT

Mean performance of the group which received Series One of The Productive Thinking Program surpassed that of the Control group on 30 of the 40 internal and posttest measures. Of the 13 significant differences, 11 favored the Treatment group. Clearly, then, use of the program led to an increase in the level of productive thinking in this large and representative group of fifth-grade students. Moreover, this superiority was found at all three levels of IQ. It was evident on a broad variety of the measures, including tasks which were quite different from the problems used in the lessons themselves. The practical and theoretical implications of these findings merit further discussion here.

First, while the performance of the Treatment group generally exceeded that of the Controls, it did so only by a modest margin—rarely did mean scores for the Treatment group exceed those of the Control group by a factor of more than 20%. Differences of this magnitude, although they may be both statistically and educationally significant, are considerably smaller than the dramatic differences obtained in two previous stu s (Covington and Crutchfield, 1965) using a.. earlier version of Series One. In those studies the Treatment group sometimes scored more than twice as high as the Control group on certain measures. The present study, however, differs from those previous ones in several important respects, and failure to find such large differences here is readily understandable. One important difference concerned the amount and kind of teacher participation during training. In the earlier studies, the classroom teachers were deliberately instructed to supplement the lessons of Series One by featuring them in class discussions, by assisting and encouraging their students to transfer the thinking principles to other curriculum areas, and by making the lessons a regular component of the overall educational effort. Recent evidence (Blount, Klausmeier, Johnson, Fredrick, and Ramsay, 1967) indicates that even a moderate degree of teacher participation (e.g., writing informative comments on student papers, providing encouragement, etc.) can increase the effectiveness of programed material by as much as 50%. This facilitative effect would probably be heightened even more by the greater extent of teacher participation which occurred in the earlier experiments.

In contrast, the present study was designed to probe the instructional limits of the materials by using them as an entirely self-contained program, with all forms of teacher participation purposely held to a minimum. Moreover, the materials were administered at the rate of nearly one lesson per day—the fastest pace feasible for regular school use. This fast-paced teacherless condition represents a very severe test of any instructional program; indeed, it is difficult to imagine a program normally being used under such restrictive conditions as these. By way of comparison, the most recent experimental study with Series One featured individualized supplementary activities and distributed practice, in which the Program was used over a much longer period of time. Results of this study are similar to those of the earlier studies, with performance of the Treatment group surpassing that of the Control group by a substantial margin.

The results of the present experiment, then, in conjunction with those from the other studies just mentioned, indicate that some degree of teacher participation (even if only modest) results in much greater educational benefit from the lessons, even though the materials alone do produce significant differences in thinking and problem-solving performance.

A second noteworthy aspect of this treatment effect is its generality. Superior performance of the treatment group was evident on a broad variety of measures and was not restricted to any one aspect of performance. A diverse set of performance indicators, each reflecting a different aspect of the total problem-solving process, showed consistent benefits as a result of training. These included achievement of solution to problems, number and quality of ideas produced, intellectual persistence, sensitivity to discrepant or puzzling facts, use of a Master Thinking Skill, and an understanding of the process of thinking itself.

However, the pattern of results did reveal some limits to the generality of training. These limits appear to be a function of the kinds of problems, rather than the kinds of performance indicators, employed. Specifically, there was a lack of treatment effects on nearly all performance measures for the brief problems, most of which emphasized divergent thinking. These brief tests, such as those developed by Torrance, are structured so as to elicit a large number of rather short, discrete responses on relatively uncomplicated tasks. Under these conditions, it seems likely that a student's level of performance would be affected considerably by his disposition or "set" to respond in a worthwhile manner. Presumably, a positive set was present for the instructed students in this study as a result of training, but performance of the Control group may have been equally facilitated by the presence of a positive set which resulted from the novel and engaging quality of the tests and from the "escape" they provided from normal classroom routine. Thus, the effect of set per se might not lead to differential performance in the two groups.

Regardless of the effects of set, however, one would still predict that the training of relevant cognitive skills would result in superior performance by the Treatment group. One possible explanation for the lack of such results focuses on the question of the relevance of such skills. If it is assumed that the brief tests require less complex cognitive processes for adequate performance than those processes taught for in the complex problem settings of Series One, then what is learned in Series One is only marginally relevant to performance on these brief problems. Thus the Treatment group might not show any marked superiority on such tasks. Another possibility which must be considered is that most children of this age, whether trained or not, can demonstrate these relatively simple cognitive skills on demand (under proper conditions of short-term motivation), even though they may rarely make use of them in their regular school work. The more complex tasks, on the other hand, presumably require the use of advanced skills which are not yet well developed, but which are learned and practiced in Series One. If this is true, training provided by the program would produce superior performance on complex thinking tasks, where the use of simple cognitive skills and a positive set is not sufficient for successful performance, but no obvious benefit from training would be apparent on the brief tests employed here.

Clearly, there is a need for further research into (1) what cognitive skills are already well developed in children this age, skills which would require practice rather than special training or instruction which was designed to create and establish them in the child's repertoire in the first place, and (2) what skills do not yet exist in the child's repertoire and will require specific and perhaps extensive training before the student will be able to make use of them. Further efforts devoted to the development of criterion measures should incorporate such research and reflect this distinction, if it is found to be valid.

Yet another important aspect of the treatment differences is the occurrence of a gradual increase in the magnitude of these differences as training proceeds. The means presented in Table 15 show a very small difference in favor of the Treatment group on the first internal test (the Rare Coins problem), which was administered just before the half-way point in the training program. On the second internal test (the Deep Sea Dive problem), administered near the end of training, performance of the Treatment group exceeded that of the Controls by a greater margin, even though it did not reach statistical significance. However, on the posttest measures, administered when all training had been completed, performance of the Treatment group surpassed that of the Control group by a statistically significant margin on two of the three relevant measures (i.e., measures comparable to those used in the internal tests). There was, then, a steady increase in the effectiveness of training as the student proceeded through the lessons.

This result differs from previous findings (Covington and Crutchfield, 1965), where the treatment effects were nearly as great on the first internal test as they were on the posttest. However, the conditions of teacher participation in that study (discussed above) may have accelerated the effects of training. Thus, an important effect of teacher participation may be to increase the rate as well as the amount by which cognitive skills are developed when materials such as these are used.

EFFECTS OF ENVIRONMENT

The five significant Environment effects on the pretest measures all resulted from the fact

GPO 806-459-4

that classes with Facilitative environments performed better than Nonfacilitative classes on the two brief divergent thinking problems developed by Torrance. The probable sensitivity of these problems to the effects of set has already been mentioned. If this interpretation is extended to the present data, the significant Environment effects may possibly have resulted from the positive or negative set induced by the two types of classroom environment. Indeed, in this study a Facilitative environment was defined as one in which an overall positive set seemed to be present while the class was working on a divergent thinking task. Hence these pretest findings support the validity of the Environment classifications. On the complex problem-solving tasks, in contrast, an increase in the student's "readiness" to think should not be sufficient to raise the level of performance, and in fact these problems did not show a significant Environment effect.

On the posttest and internal test measures, there was only one significant Environment effect out of 40 comparisons. This virtual absence of posttraining Environment effects seems to occur in large part because the treatment was particularly effective with students in Nonfacilitative environments, differentially raising the performance of this group and thus reducing any initial differences favoring the group having a Facilitative environment. Nearly a dozen Treatment × Environment interactions (five of which are significant and are presented in detail in Table 19) support this conclusion. It should be noted that the special benefits of training for students in Nonfacilitative classes amounted to more than simply a change of "set," since some of the significant Treatment × Environment interactions involved the kind of complex problems in which no main effect of Environment was found. Thus it appears that an important effect of the training materials was to improve the cognitive skills and general attitudes of students from the less favorable classroom environments, even in the absence of direct teacher intervention. Further research is needed to determine precisely what teachers of varying capabilities and attitudes can do, in conjunction with the training materials, to produce varying amounts of pupil gain or change.

EFFECTS OF SEX

Significant sex differences occurred on 21 of the total of 61 pretest, posttest, and internal test measures used in this study. With the exception of three measures in the pretest Pit problem, all significant sex effects resulted from the superior performance of females. This female superiority occurred on all types of problems (both verbal and nonverbal, convergent and divergent), and on all performance indicators except measures of intellectual persistence and of the achievement of solution to problems. This finding of generally superior performance by females is consistent with other findings (e.g., Klausmeier and Wiersma, 1964) concerning the verbal performance of students of this age.

The absence of any significant Treatment × Sex interactions is particularly noteworthy, indicating as it does that Series One of The Productive Thinking Program was equally effective with both boys and girls (although the girls managed to maintain their initial advantage throughout the training period).

EFFECTS OF IQ

The general level of intellectual ability had a clear and significant effect upon performance for most of the measures used in this study. In most cases, the predictable rank-order relation between IQ and performance was found, with the High IQ group obtaining scores superior to those obtained by the Middle IQ group, which were in turn superior to those obtained by the Low group. The effect of IQ was significant on all types of performance measures, and on all types of problems except the special test of the Master Thinking Skill.

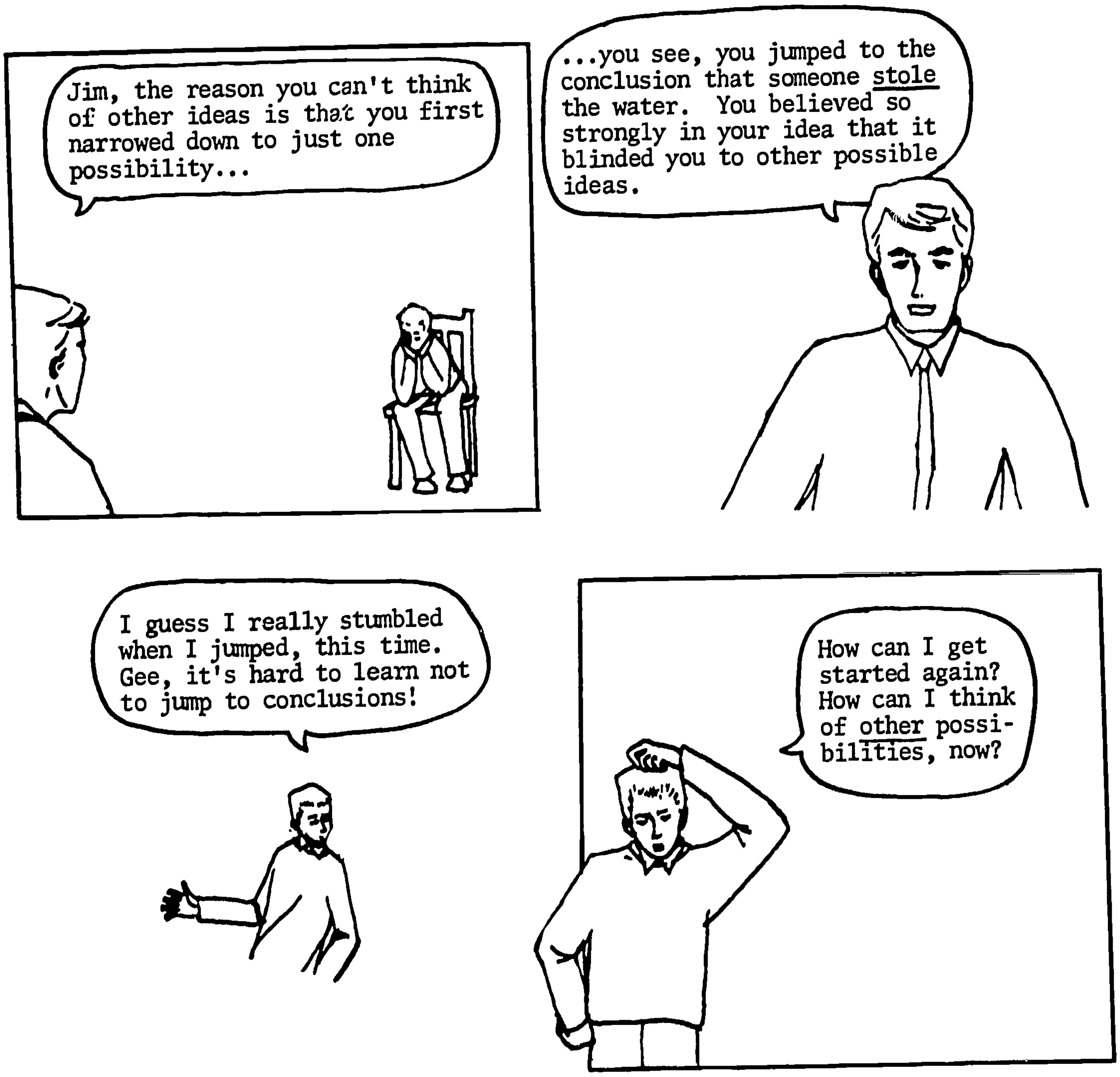
Large and significant IQ effects were found on all of the brief divergent thinking problems, and this effect was especially marked on the Squares problem, which is one of Torrance's nonverbal divergent thinking tests. Significant IQ effects were also found on the other three Torrance tests, and these effects were often large in an absolute sense (Tables 25 and 27) as well as significant statistically. This set of findings suggests that early reports of a "virtual lack of relationship between measures of creative thinking and IQ" (Torrance, 1962) are perhaps something of an overstatement and that such findings may have resulted from sampling restrictions. Since all IQ analyses in the present study are based upon the carefully selected subsample of 240 students (described on page 3) which should be representative of the entire distribution of ability in a total of 44 classrooms, the highly significant IQ effects obtained cannot easily be dismissed as an artifact of the particular sample used.

Indeed, it seems reasonable to assume that there is a moderate correlation between IQ and creative thinking when the entire distribution of intellectual ability is considered. Most studies which have reported the absence of such a correlation (e.g., Getzels and Jackson, 1962; Wallach and Kogan, 1965) involved highly selective, nonrepresentative samples, while studies which reported a positive relation between IQ and creative thinking (e.g., Ripple and May, 1962; Guilford and Hoepfner, 1966) have used far more representative samples. In any case, IQ had highly significant effects upon performance for the vast majority of measures used in the present study.

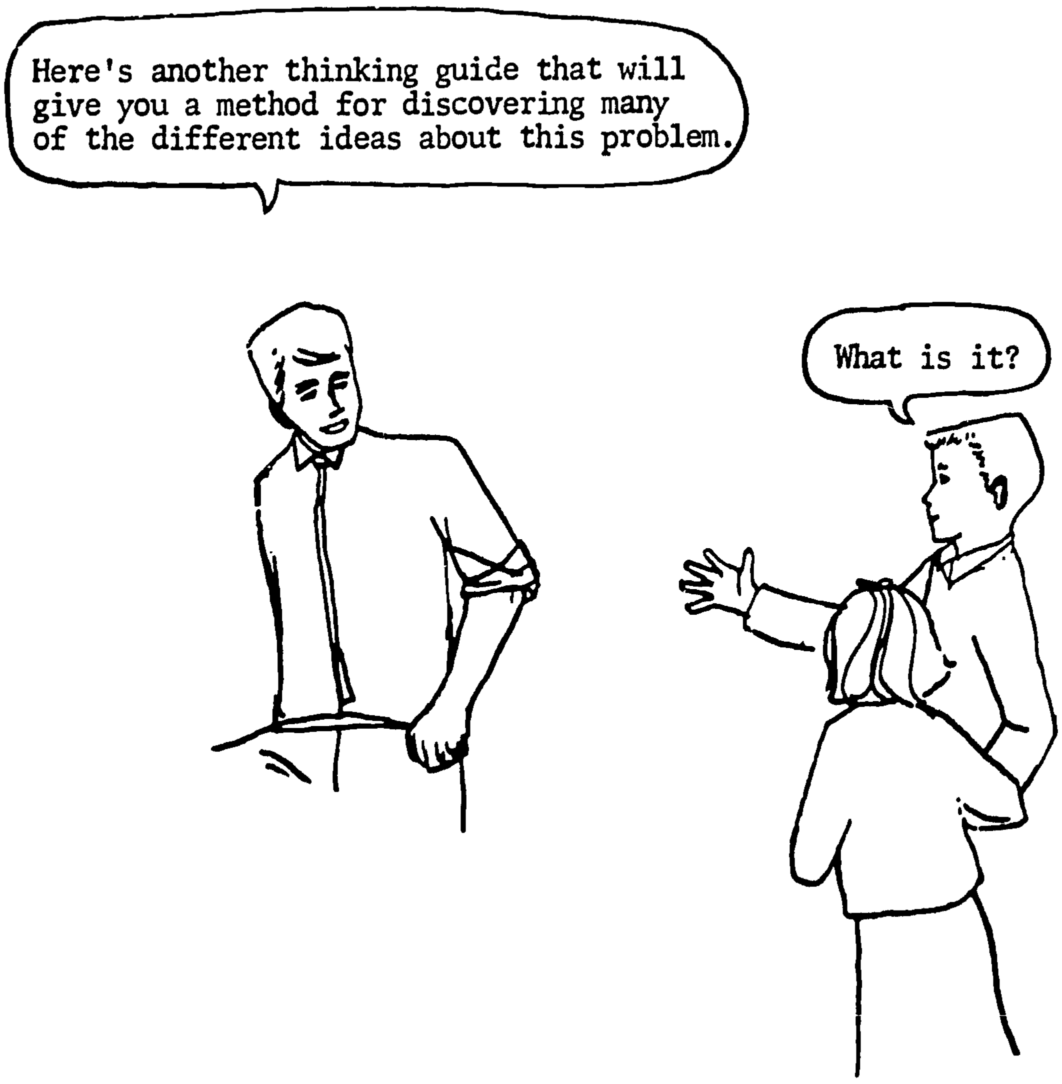
In contrast to the effect of IQ on test performance, there was no consistent pattern of interactions between IQ and Treatment; that is, the increments in thinking and problem-solving performance produced by the training program were approximately equal for all three levels of IQ. This finding agrees with the results of the earlier Covington and Crutchfield study. It should be pointed out, however, that although the materials themselves may be used effectively with students at many different levels of intellectual ability, the kind of teacher participation required is very different for various levels of ability. Students who have serious reading difficulties obviously require a great deal of assistance with the materials, while teachers whose classes are above average in general ability may devote most of their effort to showing the students how to extend and transfer the thinking principles to other curriculum topics. In an average class, the self-administering feature of the materials makes it possible for the teacher actively to assist a small group of students while the rest of the students work individually. Generally, though, the lower the level of student ability, the greater amount of teacher participation required for the effective use of these materials.

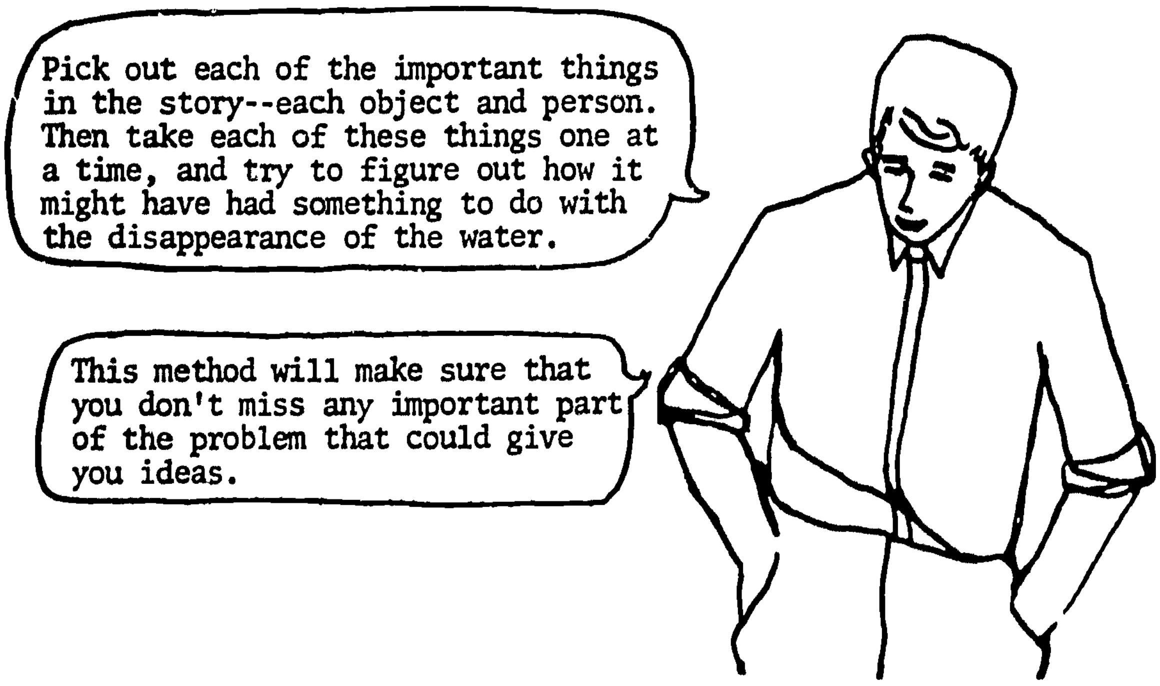
18

Uncle John notices Jim's silence:



Copyright © 1966 by Martin V. Covington, Richard S. Crutchfield, and Lillian B. Davies. Reprinted with permission.





Now, what will happen as Jim and Lila take Uncle John's advice? Turn the page to find out.


20
I'm not sure I understand how this would work....
Well, first think of some of the main things in the story....
How about the mule?
Yes, the mule. And there were some other things, too. What were they?
Let's make a list of them!

You try making a list, too. Go back to pages 8 and 9 and read the story again. Then pick out each of the main things in the story and write it down:

____

____

____

____

____

____

____

REFERENCES

Blount, N. S., Klausmeier, H. J., Johnson, S. L., Fredrick, W. C., & Ramsay, J. G. The effectiveness of programed materials in English syntax and the relationship of selected variables to the learning of concepts. Technical Report from the Research and Development Center for Cognitive Learning, University of Wisconsin, 1967, No. 17.

Covington, M. V. New directions in the appraisal of creative thinking potential. Journal of Educational Measurement, 1968, in press.

Covington, M. V., & Crutchfield, R. S. Experiments in the use of programed instruction for the facilitation of creative problem solving. Programed Instruction, 1965, 4, 3-10.

Covington, M. V., Crutchfield, R. S., & Davies, L. B. The Productive Thinking Program, Series One: General Problem Solving. Berkeley, California: Brazelton Printing Company, 1966. Distributed by Educational Innovation, Box 9248, Berkeley, California 94719.

Crutchfield, R. S. Creative thinking in children: Its teaching and testing. In O. Brim, R. S. Crutchfield, and W. Holtzman, Intelligence: Perspectives 1965. New York: Harcourt, Brace, 1966. Pp. 33-64.

Getzels, J. W., & Jackson, P. W. Creativity and intelligence: Explorations with gifted students. New York: Wiley, 1962.

Guilford, J. P., & Hoepfner, R. Creative potential as related to measures of IQ and verbal comprehension. Indian Journal of Psychology, 1966, 41(1), 7-16.

Klausmeier, H. J. Effects of accelerating bright older elementary pupils: A follow up. Journal of Educational Psychology, 1963, 54, 165-171.

Klausmeier, H. J., Harris, C. W., & Ethnathios, Z. Relationships between divergent thinking abilities and teacher ratings of high school students. Journal of Educational Psychology, 1962, 53, 72-75.

Klausmeier, H. J., & Teel, D. A research-based program for gifted children. Education, 1964, 85, 131-136.

Klausmeier, H. J., & Wiersma, W. Relationship of sex, grade level, and locale to performance of high IQ students on divergent thinking tests. Journal of Educational Psychology, 1964, 55, 114-119.

Ripple, R. E., & May, F. W. Caution in comparing creativity and IQ. Psychological Reports, 1962, 10, 229-230.

Torrance, E. P. Guiding creative talent. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

Wallach, M. A., & Kogan, N. Modes of thinking in young children: A study of the creativity-intelligence distinction. New York: Holt, Rinehart & Winston, 1965.

GPO 806-459-3