



Comparison of differences in measures of acquired information, anxiety, creativity, and student interest for two methods of teaching : lecture and problem solving
by Kitty Kienle Dick

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of Master of Science in Psychology
Montana State University
© Copyright by Kitty Kienle Dick (1973)

Abstract:

Two methods of teaching, problem solving and lecture, were used in an introductory psychology course. Students in the two groups were compared on four measures: 1) general knowledge in psychology, 2) situational anxiety, 3) creativity, and 4) interest in psychology.

It was hypothesized that: 1) the problem solving group would show greater achievement scores on a test that measures generalization and application of knowledge learned in psychology, 2) the problem solving group would have a higher level of situational anxiety than the lecture group, 3) the problem solving group would show greater scores on a test for creativity, and 4) the problem solving group would have greater interest in psychology than the lecture group.

Pre and post test scores were obtained on each of the measures and percentage of change scores were calculated for each student.

These scores were used in an analysis of variance. The results indicated that there were no significant differences between the groups and thus the hypotheses were not supported.

The author concluded that method of presentation of class material does not appear to make a difference in student acquisition of class material. This conclusion is limited to those classes which have time constraints as opposed to a programmed learning situation where the student learns at his own pace.

One factor that could affect acquisition of knowledge is the student's own desires and expectations. These could be of more importance than the method of presentation of class material.

It is suggested that future studies investigate the linkages between teaching and learning rather than the technology of teaching. Perhaps by studying the similarities between teaching methods some of these linkages can be clarified.

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Montana State University, I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by my major professor, or, in his absence, by the Director of Libraries. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature Kitty Dick
Date March 2, 1973

COMPARISON OF DIFFERENCES IN MEASURES OF ACQUIRED INFORMATION,
ANXIETY, CREATIVITY, AND STUDENT INTEREST FOR TWO METHODS
OF TEACHING: LECTURE AND PROBLEM SOLVING

by

KITTY DICK

A thesis submitted to the Graduate Faculty in partial
fulfillment of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

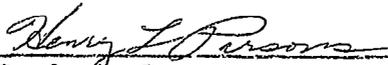
Approved:



Head, Major Department



Chairman, Examining Committee



Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

March, 1973

ACKNOWLEDGEMENTS

I am grateful to my thesis advisor Dr. M. Paul Willis and to my graduate committee for providing the guidance and assistance necessary for the completion of this study.

My sincere gratitude goes to the many friends, who gave me the encouragement I needed throughout this endeavor.

A special thanks goes to Sonja Holt who allowed me the use of her class and also provided me with her opinions and criticisms at the critical beginning stages of this study. Without her assistance, this study would not have been possible.

TABLE OF CONTENTS

	<u>Page</u>
VITA	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	v
LIST OF ABBREVIATIONS	vi
ABSTRACT	vii
I. INTRODUCTION	1
II. METHOD	16
Subjects	16
Design	17
Tests	17
Procedure	21
III. RESULTS	26
IV. DISCUSSION	30
The Problems and the Problem Solving Class	32
V. CONCLUSIONS	36
BIBLIOGRAPHY	39
APPENDICES	42
Appendix I - The General Knowledge Psychology Test	43
Appendix II - The Interest Scale	57
Appendix III - Problems Given to Group II	59
Appendix IV - Reference List Given to Group II	74

LIST OF TABLES

	<u>Page</u>
TABLE I - Means and Standard Deviations for the Percentage of Change Scores for Each Test and Each Group	28
TABLE II - Variance Table for Group I and Group II on All Measures	29

LIST OF ABBREVIATIONS

GKPT - General Knowledge Psychology Test

RAT - Remote Associates Test

STAI - State-Trait Anxiety Inventory

ABSTRACT

Two methods of teaching, problem solving and lecture, were used in an introductory psychology course. Students in the two groups were compared on four measures: 1) general knowledge in psychology, 2) situational anxiety, 3) creativity, and 4) interest in psychology.

It was hypothesized that: 1) the problem solving group would show greater achievement scores on a test that measures generalization and application of knowledge learned in psychology, 2) the problem solving group would have a higher level of situational anxiety than the lecture group, 3) the problem solving group would show greater scores on a test for creativity, and 4) the problem solving group would have greater interest in psychology than the lecture group.

Pre and post test scores were obtained on each of the measures and percentage of change scores were calculated for each student. These scores were used in an analysis of variance. The results indicated that there were no significant differences between the groups and thus the hypotheses were not supported.

The author concluded that method of presentation of class material does not appear to make a difference in student acquisition of class material. This conclusion is limited to those classes which have time constraints as opposed to a programmed learning situation where the student learns at his own pace.

One factor that could affect acquisition of knowledge is the student's own desires and expectations. These could be of more importance than the method of presentation of class material.

It is suggested that future studies investigate the linkages between teaching and learning rather than the technology of teaching. Perhaps by studying the similarities between teaching methods some of these linkages can be clarified.

I. INTRODUCTION

This study compared the differences between a lecture method and a problem solving method of instruction in a college level introductory psychology class. These methods were compared on, a) acquisition of general knowledge in psychology, b) anxiety in the class situation, c) creativity, and d) interest in the subject matter of the class.

For decades studies have been done comparing different methods of instruction. The results of these studies are not encouraging. In essence, these results are conflicting, for no one teaching method has emerged as a superior method in terms of acquisition of the class material by the students involved.

The following are seven examples of the work that has been done in the area of comparing methods of instruction. Scheidenmann (1927) compared a project course developed by Seashore, (1928) with a lecture-discussion method.¹ He found no significant difference between the average final examination scores of the students for both methods.

Greene (1928) studied the relative effectiveness of lecture and individual reading as methods of college teaching. One group of students received the class material through lecture and the other group read the

¹The project course consisted mainly of supervised individual study, conferences with the instructor, and writing of reports on eight projects which made up the course content.

material. Each class covered four topics in their respective methods. The results showed that there was little difference between the average scores on class tests for the two groups. These results agree with the results of Scheidenmann. However, Greene further concluded that the lecture class tended to bring both the best and the poorest students toward the class average. Greene summarized that the lecture method might be advantageous for poorer students while individual study might be better for good ones, since the good students in the reading group were more confident of their knowledge than were those in the lecture group. On a test given a few weeks later, however, the lecture group showed a greater retention of the material than did the reading group.

Guetzkow (1954) compared three methods of instruction at the college level: recitation, discussion, and tutorial. When the groups were compared on an objective final examination over the course material, he found no difference among the groups. The author generally concluded that his study confirms the conclusions that have been derived from experiments on instructional procedures since the 1920's: that there is no consistent difference among methods for acquisition of class content.

Rohrer (1957) studied the effect of class size on achievement and student attitude in a college level American Government class. His most significant finding was that amount of achievement as measured by

a standardized test, attitude of the students toward the difficulty of the class material, and their interest in it varied as a function of the course instructor and did not vary as a function of the size of the class. There are two characteristics of these classes that Rohrer felt should be mentioned when generalizing his results. First, the classes were introductory rather than advanced. Second, they required students to acquire general principles and processes rather than to develop skill in applying principles and processes.

Remmers (1933) studied the differences between examination scores on course content for students being taught in a large lecture method and a small recitation method. He found no significant difference between the two groups.

Marr, Plath, Wakeley, and Wilkins (1960) tested the assumption that something is being taught by the lecture method that the student cannot get for himself by reading the textbook. The two groups he used were a lecture method and a non-lecture group. This non-lecture group met once a week and the instructor answered specific questions the students raised concerning the reading assignment. However, no discussion was allowed in these groups. Both groups had the same text and were given exams. The criteria for comparison of the groups were a series of four objective tests, and an evaluation form that was filled out anonymously by all students. A separate analysis of variance was done on each test and it was found that those in the

lecture group did significantly better ($p < .01$) on the final examination than did those in the other group. Also, the students with higher grade point averages did better than students with lower grade point averages when all students received similar instruction. An analysis of variance was then done on the variables of instructor, methods, and ability level of student (as measured by grade point average). None of these variables were found consistently significant over groups except that students with higher grade point averages did better in each method. (Marr, et al, 1960) concluded that students who are taught by the lecture method do better on examinations than do the non-lecture students.

In the area of programmed learning, Roderick and Anderson (1968) studied the differences in achievement in an introductory psychology class for a programmed class and a class reading text book style summaries of the same class. The programmed class used the first four sets of the Holland-Skinner program (1961). The results revealed that college students who completed these sections of the program scored no higher on an achievement test given immediately after completion of the program than those who studied the summary. But the high school program group was superior to the summary group in retention scores on a delayed measure.

A recent trend in comparative teaching methods studies seems to be toward studying variables that occur in a teaching situation along

with the teaching method to see if certain variables that affect the learning situation can be identified. Many of these variables are classed as non-cognitive e.g. personality variables. Hall (1970) hypothesized that the congruence between the teacher's actual teaching style and student's ideal teaching style would correlate higher with the learning that occurred in the college classroom than with either of these constructs alone. He found that actual teaching style was generally a better predictor of learning than were the fit scores between actual and ideal teaching style. Lippman (1970) studied the unique properties of two types of instruction and the characteristics of students likely to perform best in each method. He found no significant difference between methods.

Spielberger (1970) investigated anxiety and drive theory in computer assisted learning. The findings supported the importance of a conceptual distinction between anxiety trait and anxiety state. Brown (1970) studied the effect of learner control in automated instruction. He found that learner control did not improve performance in comparison to the no control group.

The above studies represent individual efforts to compare different methods of instruction. Another approach that is taken by some investigators interested in teaching methods and their differences is to review many studies done over a period of years and to consider these results in addition to their own investigations. One such study

was done by Longstaff (1932). After considering the work that had been done by other investigators, he stated that: "The experimental evidence submitted to the present time tends to support the general conclusion that there is little difference in student achievement in large and small classes and, also that it makes little difference as to what method of presentation of materials of the course is used [pg. 337]."

Longstaff (1932) also studied two methods of instruction, a lecture-quiz and an all-lecture method. He compared them not to determine the relative value of the two methods but to determine whether the lecture-quiz method was superior under normal, practical conditions. He correlated class grades, as determined by a pre- and post-test, and the College Ability Test (CAT). With respect to final grades earned in the course, Longstaff found that one method was as efficient as the other. However, the lecture-quiz group did slightly better but this difference was not significant.

Ten years later, Wolfle (1942) reviewed investigations of teaching methods in psychology. He emphasized that elementary courses in psychology are usually taught in one of three ways: 1) a lecture method with a lecture in every class period given by one man or by several members of the department, 2) a lecture method that meets for most of the time, but once or twice a week the group is divided into small discussion or quiz sections, these sections are taught by junior members of the department, and 3) a class which is divided into

small discussion groups, each of which is taught by the same instructor every day. Added to any one of these methods, as a portion or elective supplement of the course, could be an elementary laboratory. In concluding the review of the studies, Wolfle said that all the studies done to that date on class size and method of instruction could be summarized by Longstaff's statement of 10 years earlier.

Birney and McKeachie (1955) made a survey of research that compared methods of instruction in psychology that had been published since 1942. They concluded that after three decades of research Longstaff's statement was still not outdated.

Dubin (1968) examined the data of over 80 studies comparing different teaching methods. He "added up" the data of these studies, rather than the conclusions of their authors, and reported no significant difference among the different teaching methods in learning of the class material. This learning was usually measured by a written final examination. Dubin concluded as did the three previous studies reviewing the literature, that no particular method of presentation of the class material is clearly superior. He suggested that one problem with these studies is that they have investigated the technology of teaching rather than the linkages between teaching and learning. However, one similarity among the studies examined by Dubin, as well as those examined by Longstaff, Wolfle, and Birney and McKeachie, which might explain the lack of apparent difference among methods of

instruction was the measure of learning: namely, the written final examination. A more appropriate measure of the acquisition of knowledge might be a test which measures not merely the recall of facts but rather the generalization and application of knowledge. This assumes that, in fact, different teaching methods should result in differences not measurable by a standard final examination. This assumption does not include possible differences, between the groups attributable to teaching method, on a measure of retention. Studies by McQueen (1962) and Coop (1970) suggest that a difference may be expected.

McQueen (1962) did an experiment on the teaching of psychology and compared two methods of teaching. A control group met three times a week for lecture-discussion and read one text book (the conventional method). The experimental group read the same text book, plus two additional volumes of selected readings in general psychology, and met one day a week for a lecture-discussion meeting. Eight bi-weekly examinations were given to both groups and these exams were the basis of analysis between the groups. It was found that the control group did significantly better ($p < .05$) on these exams but that the groups did not differ significantly in scholastic ability. Scholastic ability was determined by the accumulative grade point average of each student. McQueen makes two conclusions that might explain these results. First, that perhaps the experimental method could be improved if the students

were given guidance in their extra readings, for example if study questions were used to direct their reading. Secondly, he felt that the students in the experimental group were not accustomed to the responsibility and had not developed the study habits necessary to perform at their best in the experimental method. From the above conclusions, it is suggested that study questions used to direct students reading could improve the individual study method.

Coop (1970) investigated the effects of different student cognitive styles on achievement in two methods of teaching. The cognitive styles used in this study were those categories defined by Kagan, Moss and Segel (1963). These were descriptive analytic style, inferential categorical style, and relational-contextual style. The two methods of teaching consisted of a teacher-structured presentation and independent-problem solving. The students used in this study were education majors. The teacher-structured presentation class was not strictly a lecture type class because it also included lecture-discussion, group discussion, and the analysis of audio and video tapes of classes. The independent problem-solving class was given film scripts of typical problem situations faced by teachers in the classroom. Besides the scripts they were also provided with discussion guides which served as self-instructional materials in classroom-interaction analysis techniques. The students in this group could either work alone or work in a group in analyzing the problem and

obtaining the necessary information to answer the problem posed by the script. Each student then wrote his (her) solution and it was given to the teacher. These students were not required to attend class and were given 14 days for each unit of study. An analysis of variance was done on the variables. A significant difference was found between the two methods in terms of achievement in factual content material on an objective test: the teacher-structured presentation group was found to be superior to the individual problem solving group. But there was no significant difference found among the cognitive styles and no interaction between teaching methods and cognitive styles. Coop made a few comments about his finding that the teacher-structured presentation method was found superior both in factual content and conceptual generalization achievement. He felt one reason for these results could be that students are more used to this method of teaching because most of their college classes are similar to this method. In addition, the individual problem solving class is less structured than what students experience in other classes. Because of this the individual-problem-solving method appears to demand more internal motivation than the other class since students are not required to attend class every day. These two considerations might account for the poorer showing of the individual problem solving group.

Coop's (1970) study points to a possible solution to McQueen's (1962) suggestion for study guides. These are individual problems

that require the gathering of certain information and the eventual formulation of a solution. A variation on this type of problem made in the present study, is the situation problem. These situation problems are not scripts but require the practical application of information. In addition, because of the criterion used in developing these problems, the students have to show a certain degree of mastery or expertise on the topics of the problems. These two characteristics of these situation problems differentiate them from study questions. They require the student to utilize facts of psychology in order to find a solution to a unique situational problem.

The conclusions drawn by McQueen (1962) and Coop (1970) suggest that a factor which may have contributed to the individual study group's poorer performance was that the situation of the class was new to them. Coop (1970) suggests that the individual study group was not accustomed to the lack of structure experienced in the individual study situation. A factor that could be at work here is anxiety, propagated by two factors, the new situation and the lack of structure.

Pickrel (1958) found that anxiety can affect test performance. He studied the differential effect of manifest anxiety on test performance. He found that in simple measures, where the subjects are confronted with a simple task, there was no significant difference between an anxious group and a non-anxious group. But when the measures

became more complex, that is, when there were a number of competing alternatives in a task, the non-anxious group's performance was superior to the anxious group's performance. It appears then that anxiety can have an effect on test performance.

In considering the idea of the less structured atmosphere of the individual study group, McKeachie (1951) tells us that anxiety can have an effect on classroom performance. He found that anxiety is a motivating force in performance but that if it is not resolved or does not have an outlet, it can inhibit performance. McKeachie points out two sources of anxiety in the classroom. First, the student experiences anxiety on entering the class because of his dependency on the teacher for grades. Then second, the student's anxiety is heightened or reduced by the teacher's teaching behavior. Anxiety can best be controlled, says McKeachie, in a highly structured situation. If the instructor does not provide such structure, the typical student may be deprived of a mechanism for handling his anxiety.

From McKeachie's findings and discussion it appears important to determine the effect a new teaching method has on students in terms of anxiety. The study of a new teaching method also should involve an examination of the teacher's teaching behavior. If the teaching method creates excessive student anxiety, it may not be an optimum teaching method for acquisition of knowledge because of possible inhibiting effects on performance.

Many investigations have been done relating creativity and problem solving. In research, creativity and problem solving have been combined under the title of "creative problem solving". One trend in this area of research is to determine the effect of creativity on problem solving ability. An example of such a study is Eisenstandt's (1966) investigation of problem solving in creative and non-creative college students. His main hypothesis was that since creative subjects are supposed to have a heightened ability to observe accurately, they would then respond differently to a threat situation than would non-creative subjects. His results revealed that creative individuals do have faster solution times, probably resulting from their more accurate observations. But his results did not indicate that the groups differed in a threat situation. Eisenstandt (1966) concluded that what separated the creative from the non-creative individual is the efficient solution of the problems. Maier (1969) studied a slightly different aspect of creativity and problem solving. He investigated whether subjects superior in solving difficult problems having objectively correct solutions also achieve solutions rated as 'creative' for a problem with several possible answers. His results indicate that those individuals who reached an integrative (creative) solution to the Changing Work Procedure (CWP) problem were found to have solved significantly more of the objective problems than did the individuals who had other solutions not designated 'creative.' The object of the

CWP problem is to find a solution to the problem of changing the work assignments of three workmen so that the most efficient use of time is achieved. Maier (1969) concluded that those individuals who are innovative problem solvers, as tested by the CWP problems, are likely to be more successful in finding solutions to other problems. However, he cautions that they may not be successful on all other types of problems.

The common denominator between creativity and problem solving, according to Davis (1966), is that uncommon responses are required by both. Given that creative people are more efficient problem solvers, and assuming Davis is correct, it is reasonable to suggest that experience in finding solutions to unique problems might increase a person's creative ability.

Informal discussions with students and teachers involved in college level introductory classes suggest that these classes are often uninteresting and monotonous. It is possible that significant differences in instructional methods could be revealed by some appropriate measure of student interest. An assumption made by the present experimenter is that if a specific method of instruction bores the student involved, then student interest in the given subject matter will likely decrease. Similarly, if a teaching method is perceived by the student as 'interesting,' 'exciting,' etc., then it may very well lead to increased student interest in the given subject matter of the class.

The present study was designed to examine the variables discussed above: teaching method (specifically problem solving versus lecture), anxiety, creativity, and student interest. A number of hypothesis are proposed as to the differences that will result between the groups experiencing the different methods of teaching: 1) the problem solving group will show greater achievement scores on a test that measures generalization and application of knowledge learned in psychology, because they have been solving problems that required them to use facts they have gathered from many sources in psychology, 2) the problem solving group will have a higher level of situational anxiety than the lecture class because the problem solving method probably will be new to most of the students, 3) the problem solving group that had experience finding solutions to problems during the quarter will show greater scores on a test for creativity, and 4) the problem solving class will have greater student interest in the class and therefore greater interest in the subject matter of the class.

II. METHOD

Subjects

Two groups were randomly selected from an introductory psychology class containing 131 students, during the Spring quarter of 1972. Each student on the class rolls was assigned a number, then a table of random numbers was used to choose 48 students. There were twenty-four subjects in each group, 12 males and 12 females. Eighty-three students were not involved in either group. The 24 students in the lecture group were left in the class with the 83 non-participating students. The 24 students in the problem solving group met in another classroom. These classes began on March 28 and ended on June 9.

The average age for both groups was 19 years. The mean grade point average in Group I was 2.94 and in Group II it was 2.78. One student in Group I had three credits in psychology and one student in Group II had four credits in psychology. All other subjects had no previous credits in psychology.

In Group I, eleven students were taking the introductory class to meet a requirement of their curriculum. In Group II fourteen students were taking the class for this reason. The rest of the students were not filling a requirement in taking the class. Group I was made up of 19 freshmen, 3 sophomores, 1 junior, and one student who did not respond to this question. In Group II there were 20 freshmen, 1 sophomore, 1 junior, 1 senior, and 1 student who did not respond to this question.

Design

Group I was the lecture group, (controls) and Group II was the problem solving group, (experimentals). Pre- and post-test scores were used to calculate a "percentage of change" score for each student on each of four tests described below. "Percentage of change" scores were obtained from the following formula:

$$\frac{B - A}{T} \times 100 = \text{percentage of change}$$

A = first score on the test
 B = second score on the test
 T = total possible score for the test

These "percentage of change" scores were used as the dependent variable in an r(B)C unrepeated nested model of an analysis of variance. In this design, subjects were regarded as a random variable, teaching method (lecture vs. problem solving) was the fixed constant "nest" variable and the four "tests" were regarded as categories of a regular fixed-constant variable.

Tests

Four tests were used to measure different aspects of the learning situation: 1) general knowledge in psychology, 2) anxiety, 3) creativity, and 4) interest in psychology. These tests were given to all Ss in both groups.

To test for broad content learning in the two groups, a "General Knowledge Psychology Test" (GKPT) was constructed from two sources: 1) questions supplied by different publishers as sample questions for introductory psychology textbooks, and 2) questions submitted by staff members of the Montana State University Department of Psychology (c.f. Appendix I). The GKPT questions were selected, by the investigator, for their general character, i.e., they were not directed toward a specific text or class but rather attempted to cover the general field of psychology. The purpose of the GKPT was to provide a measure of knowledge of psychology, representing a content domain which was presumably of greater scope than the typical "text-oriented" final examination in introductory psychology classes.

To measure anxiety in the class situation, the State-Trait Anxiety Inventory (STAI) was used (Speilberger, 1970). Two forms of this test were used, the A-state and A-trait forms. A-state is defined, as being "a transitory emotional state or condition that is characterized by subjective consciously perceived feelings of tension and apprehension. A-state may vary in intensity and fluctuate over time and it measures a reaction taking place at a particular moment in time and at a given intensity" (Speilberger, 1970). This form of the test presumably measures the "situational anxiety" present in each group, reflecting possible group differences created by the two teaching methods. The trait anxiety (A-trait) form of the STAI was also given to both groups

in an effort to determine whether the groups differed in terms of a disposition toward anxiety. A-trait, on the STAI, refers to ". . . relatively stable individual differences in anxiety proneness. . . ." (Spielberger, 1970). The A-trait form, presumably, measures differences "between people in the tendency to respond to situations perceived as threatening" (Spielberger, 1970).

There are 20 items on each scale of the STAI. On the A-state the Ss are instructed to indicate how they feel right now in terms of the items. Or as in the experimental condition, how the students felt in their respective classes. For each item or statement they had a choice of four responses: 1) not at all, 2) somewhat, 3) moderately so, or 4) very much so. For the A-trait, Ss indicated how they "generally" felt. The items were rated as follows: 1) almost always, 2) sometimes, 3) often, or 4) almost never. Scores were then given to each response and a total score was derived.

College undergraduates were used as subjects in collecting test-retest reliability data for A-state and A-trait scales of the STAI. For the A-trait scale, test-retest reliability ranged from .86 to .73. The test-retest reliability for A-state ranged from .16 to .54. (Spielberger, 1968).

The STAI A-trait scale has been correlated with other measures of trait anxiety. These were IPAT (Cattell and Scheier, 1963), the Taylor Manifest Anxiety Scale (Taylor, 1953), and the Zuckerman Affect Affective

Checklist (Zuckerman, 1960). For college females the test showed correlations of .75, .80, and .52 respectively, and for college males they correlated .76, .79, and .58 respectively. (Speilberger, 1968).

The Remote Associates Test (RAT) was given to both Group I (lecture) and Group II (problem solving) (Mednick and Mednick, 1967). This test was used in an attempt to determine whether the different instructional methods (group treatments) resulted in differences on a measure of "creativity". The RAT purportedly measures a specific aspect of creativity, an associative aspect. This is the ability to recognize relationships between "seemingly remote ideas and forming them into new associative combinations...the more mutually remote the elements of the new combinations, the more creative the process or solution" (Mednick et al, 1967).

The RAT consists of 30 groups of words. The Ss are instructed to find a fourth word that is related to the other three in each group. The fourth word can be related to the others for various reasons. The scoring key consists of the words that are correct responses for each group. Those responses that are correct are graded, those that are not correct are not scored.

From normative data collected on a number of college undergraduate samples, odd-even reliability was calculated with the Spearman-Brown odd-even coefficient of correlation. For male undergraduates at the University of Michigan, correlation was .91 for test form 1. Female

students at Bennington College had a correlation of .92 for test form 1, and for University of Maryland undergraduates the correlation was .86 for test form 2 (Mednick et al, 1967). Test form 1 and 2 were given to a group of undergraduates and a correlation of .81 was found between the two forms (Mednick et al, 1967).

The RAT scores were also correlated with faculty ratings of 20 architectural students for creativity. These raters had been advising and evaluating these students in their designs and models for at least one year. This correlation was found to be .70 ($p < .01$) (Mednick et al, 1967).

To determine the amount of interest in psychology that existed in both groups before the class began, the students were asked to rate themselves on a scale from 0 to 100 as to how much interest they felt they had in psychology compared to other college students in general (c.f. Appendix II). "Interest" was defined on this scale as an active interest, i.e. planning to do some work in the subject of psychology such as, take another psychology class and/or read a psychology book on their own.

Procedure

When the students were tested on the four measures described above, each test was given to each student twice to get a pre- and post-test score. The learning test, the creativity test, and the interest scale were given at the beginning and at the end of the course.

The anxiety scale was given at the beginning of the course and at mid-quarter. This mid-quarter testing for anxiety was considered desirable in order to get a measure of anxiety during the period of teaching. The experimenter felt that if this test, which purports to test a transitory situational anxiety, were given at the end of the quarter, the anxiety possibly created by the teaching methods might no longer be of sufficient strength for measurement. For, at the end of the quarter, the student might well consider himself as "out of", or nearly "out of" the situation.

Group I was taught introductory psychology in a conventional lecture method. These Ss attended class four times a week for four academic credits, three days were spent on lecture and one day a week was reserved for a film on some aspect of psychology. These students were held responsible for reading the text and for taking five exams during the course.

Group II was an individual study, problem solving group. This group was given nine problems during the quarter, one problem a week (c.f. Appendix III). Each problem covered one or two aspects of psychology and these topics corresponded to those being lectured on in Group I. At the end of the quarter, both groups had covered the same material but in their respective ways.

The problems for Group II were developed by the investigator using some introductory texts as sources (e.g. CRM Books, 1970). A

number of criteria were used in problem development. First, each problem covered a different but specific aspect of psychology each week. Second, the problems allow the students a wide range of solutions, because there are no objectively correct solutions. The students could then have a choice of perspective for their solutions. Third, the problems were to involve practical application of knowledge and were designed to challenge the student's imagination when he was seeking a solution. These problem situations were intended to differ from those in which common sense alone would be sufficient for an effective solution. If the student did not show in his solution that he had adequately mastered outside readings on the topics of the problem, he did not receive a good grade on the problem.

The grading system for the problems considered in Group II was based on two criteria. The first criterion was apparent coverage of the expected minimum outside reading material identified for each problem. Specifically, each problem had a list of topics handed out along with it, and the student, in finding a solution, was required to touch upon these relevant topics. The second criterion was adequate "depth" of the solution, i.e., did the student merely make a brief outline of what he had covered in the outside readings or did he, in his solution, show that he was applying the information correctly to the problem. In other words, did the solution show that the student understood what he was talking about. Since no one solution was "the"

correct answer for these problems, grading was probably not as reliable as it would have been with problems having only one correct solution. Each solution was graded on its individual merit. The scoring of the problems was done by the investigator and a student assistant. The first problem was graded exclusively by the investigator. During the second week, a student assistant was given the above criteria and the rest of the problems were graded by this assistant. To keep the investigator up to date on the students' progress with the problems, the investigator and the assistant consulted often during the quarter. No effort was made to determine interrater agreement between the grades on the first problem and the grades on the problems that followed. The graded solutions were handed back to the students each week.

During the week the students in Group II would do individual readings and study on the topic for the week. They collected their psychology information from individual readings of the text, from the references they were given, and from the references they had found on their own. These Ss were instructed to find a solution to the problem using whatever relevant information they had gathered from these readings. They were assigned the same textbook as Group I (Hilgard, Atkinson and Atkinson, 1971) but they were also given a supplementary list of references to use as additional sources of information to solve the problems (c.f. Appendix IV). These references were on reserve in the library and consisted mostly of an assortment of chosen

texts, introductory and upper level psychology texts, that provided a presumably wider view of the different aspects of psychology than that covered in a typical introductory class.

Group II was required to attend class only once a week, on Fridays. On this day, these students handed in their individual solutions to the week's problem. In this class, the students discussed, with as little direction as possible from the instructor, problems and questions that had come up during their individual problem solving. This interaction among students helped supplement their individual readings. The students in this class had no lecture periods and were largely responsible for finding the facts necessary to solve the problems.

III. RESULTS

The mean percentage of change score for Group I (lecture) on the GKPT, is 14.14 as compared to the Group II (problem solving) mean of 7.45 (see Table I). This difference, while not significant at the .05 level, is in the opposite direction from that which was hypothesized.

The mean percentage of change score for Group II on the A-state scale of the STAI is 6.30 while it is -.37 for Group I (see Table I). This difference, while in the hypothesized direction, is not significant at the .05 level. This suggests that the problem solving situation created more situational anxiety for Group II than the lecture situation did for Group I.

The two groups did not differ much in their mean percentage change scores for the RAT. The mean for Group I is 6.25 and 5.83 for Group II (see Table I). These results are not significant at the .05 level, and show a difference between the groups that is contrary to the hypothesis that Group II would have higher percentage of change scores for the creativity scale.

Also, there was not much difference in the mean percentage of change score for the groups in interest, with -1.04 for Group I and -2.04 for Group II (see Table I). These results are contrary to the hypothesis that Group II would have more interest in the subject matter of the class, although the difference is not significant at the .05 level. These results show that both groups had a negative mean

percentage of change score, i.e. both groups decreased in interest for the subject matter of the class from the beginning to the end of the course, but that Group II had a slightly greater drop than Group I.

The mean scores for the A-trait scale of the STAI did not reveal a significant difference between the groups. Group I had a mean score on this test of 39.70 and Group II had a mean of 35.08. A t-test was run on these scores and the group difference was not significant at the .05 level. This result shows that neither group revealed a significantly greater tendency to be anxious on this test.

An r(B)C model of an analysis of variance was done, with "r" representing the number of individuals randomly selected for each "B" group nest. The "C" variable represented the four tests given to each group. The results of the analysis of variance revealed no significant variation at the .05 level, between the two teaching methods for the measures used (see Table II). The "tests" variable did reveal significant differences ($p < .001$) in mean percentage of change scores. The "tests x groups" interaction, however, was not significant ($p > .05$) (see Table II). Thus, the method of instruction variable did not produce significant differences in over-all (across tests) means nor in the pattern or trend of the separate test means.

